

Partie III

***Modalité d'interaction
et Multimodalité***

Modalité d'interaction et Multimodalité

Laurence Nigay

Mémoire présenté en vue de l'habilitation à diriger des recherches, Décembre 2001.

Préambule

Ce mémoire décrit les travaux de recherche que j'ai effectués à l'IMAG, au laboratoire CLIPS (Communication Langagière et Interaction Personne-Système) dans l'équipe IIHM (Ingénierie de l'Interaction Homme-Machine), depuis que je suis devenue Maître de Conférences à l'Université Joseph Fourier en septembre 1994.

J'ai mené ces travaux en collaboration avec des collègues d'autres laboratoires de l'Université Joseph Fourier (LSR-IMAG et TIMC-IMAG), et d'autres Universités françaises ou étrangères dans le cadre de projets de recherche, ainsi qu'avec les étudiants en Thèse que j'ai encadrés ou que j'encadre

1 Introduction

Nos travaux s'inscrivent dans le domaine de l'Interaction Homme-Machine et ont trait à la conception et à la modélisation logicielles des systèmes interactifs. Dans notre étude, nous adoptons une approche centrée sur l'interaction et nous visons à concevoir et à mettre en œuvre de nouvelles modalités d'interaction et formes de multimodalité, en sortie (du système vers l'utilisateur) comme en entrée (de l'utilisateur vers le système). Ce faisant, il ne s'agit pas seulement d'exploiter au mieux les ressources techniques, mais de développer des techniques d'interaction qui soient conformes aux attributs de la communication humaine. Et l'être humain bouge, parle, manipule, voit, entend. Dans ce contexte, la multimodalité prend son sens aussi bien dans la variété des canaux d'interaction à destination de l'utilisateur (interface en sortie) que dans les diverses manières d'utiliser ces canaux (interface en entrée).

Dès 1980, la multimodalité est étudiée aux États-Unis : en particulier Richard Bolt propose le paradigme du "mets ça là" dans le cas de l'intégration de la parole et du geste. En 1983, Nicholas Negroponte fonde le MediaLab au MIT avec comme objectif l'étude de l'intégration de ces nouvelles technologies. En France, la prise de conscience a lieu plus tard avec la création en 1990 d'un groupe d'étude au sein du GDR-PRC CHM. Le concept d'interaction multimodale est né.

- Dès lors, les travaux se sont axés sur la multimodalité en entrée, mettant en œuvre plusieurs dispositifs physiques d'entrée (Bourguet & Caelen 92) (Gourdol et al. 92) (Nigay & Coutaz 93) (Mignot et al. 93) (Martin 94), pour ne citer que certains travaux français.
- Peu de travaux se sont intéressés à la conception de la multimodalité en sortie. Les travaux existants n'adoptent pas un point de vue globale de l'interaction multimodale en sortie. En effet certains étudient la pertinence de l'information véhiculée par les modalités, d'autres l'accès aux documents et aux postes de travail pour les handicapés (en particulier visuels) et enfin certains étudient la

génération automatique d'énoncés multimodaux du point de vue des supports de l'information, le terme multimédia y étant le plus souvent retenu.

Aujourd'hui, les progrès technologiques comme la miniaturisation des microprocesseurs, l'explosion des technologies communicantes, aiguillonnée par l'objectif "*d'accès à l'information par tous, n'importe où, n'importe quand, pour des besoins personnalisés*", en d'autres termes, l'ubiquité instantanée et universelle de l'information, ouvrent des perspectives nouvelles pour l'interaction tout en augmentant la complexité de la conception par l'ajout de paramètres. Face à l'éventail des possibilités et à la complexité de conception, l'objectif de notre recherche est l'identification de concepts, l'élaboration de modèles et la mise en œuvre d'outils logiciels visant l'invention de nouvelles techniques d'interaction (modalités d'interaction et formes de multimodalité) et permettant de comprendre la nature de l'interaction. Ces objectifs s'appuient sur la démarche scientifique suivante :

- Identification de caractéristiques motivées par les sciences cognitives et le génie logiciel
- Organisation des caractéristiques en espaces conceptuels permettant de comprendre la nature d'un phénomène, et/ou de raisonner sur des choix de conception
- Intégration logicielle des éléments conceptuels sous forme d'architectures logicielles, d'outils ou de nouvelles techniques d'interaction.

Ce faisant, les concepts, modèles et outils produits interviennent sur l'ensemble du processus de développement d'un système interactif : définition des besoins, conception ergonomique, conception logicielle (architecture), réalisation logicielle et évaluation. Pour chacune de ces phases, nous cernons notre étude des modalités d'interaction et des formes de multimodalité en considérant deux axes de recherche :

- Exploration visuelle d'espaces informationnels :
 - Visualisation de grandes quantités de données
 - Visualisation de données temporelles
 - Navigation collaborative dans un espace informationnel
- Réalité augmentée :
 - Réalité augmentée et Virtualité augmentée
 - Réalité augmentée en situation de mobilité
 - Réalité augmentée collaborative

Nous justifions l'importance de ces deux axes vis-à-vis des avancées technologiques ainsi que leurs pertinences par rapport aux modalités d'interaction et à la multimodalité comme suit :

Exploration visuelle d'espaces informationnels :

La conception de techniques d'interaction qui permettent aux utilisateurs d'explorer, comprendre et appréhender une grande quantité d'information définit un sujet d'étude d'importance en réponse aux avancées des technologies de la communication. En particulier, les autoroutes de l'information permettent une diffusion instantanée et ubiquitaire d'une grande quantité d'information. Face à ce déploiement technologique, il convient cependant de s'interroger sur la qualité de l'information et sa maîtrise intellectuelle. Cette abondance à l'échelle planétaire nourrit en même

temps le risque de détruire l'information : cette dernière existe, mais elle devient inaccessible. Le problème n'est plus tant la disponibilité de l'information mais la capacité de sélection d'une information qui réponde au besoin. Par exemple, les travaux sur la télévision interactive illustrent le problème : une grande quantité d'information doit être accessible à tout public et tenir compte de la variété des conditions d'usage à domicile. De plus la quantité d'information numérique disponible ne fait que croître ; cette augmentation ne peut pas être maîtrisée uniquement par la croissance en puissance des calculateurs. Un projet à l'Université de Berkeley a estimé à un exa-octet (1 million de tera-octets) la quantité de données générées annuellement de par le monde. Parmi ces données, 99,997 % sont disponibles sous forme numérique (Keim 01).

À l'évidence, il convient d'assister les utilisateurs dans leur quête d'information pour des situations chaque fois particulières. Plusieurs approches sont envisageables : améliorer les moteurs de recherche ou encore déléguer certaines tâches de recherche à des agents logiciels. Toutes ces pistes sont assurément valides et fonctionnent de manière complémentaire. Dans notre étude, nous adoptons une approche centrée sur l'interaction et nous visons à concevoir de nouvelles modalités graphiques informationnelles ou techniques de visualisation ainsi que de nouvelles formes d'interaction multimodale et collaborative qui aident l'utilisateur à appréhender une grande quantité de données.

Le chapitre 4 expose notre approche de travail et les résultats obtenus concernant l'exploration visuelle d'espaces informationnels.

Réalité augmentée :

Nous constatons que les travaux à l'origine de nouvelles modalités d'interaction et formes de multimodalité reposent sur un ensemble de dispositifs physiques dédiés à l'interaction et centrés autour d'un même **espace d'interaction : l'ordinateur de bureau**. Ainsi, un microphone combiné avec une souris (Nigay & Coutaz 93), deux souris manipulées simultanément (Chatty 94), une caméra couplée à une souris (Bérard 99), définissent des formes d'interaction multimodale visant à offrir une communication plus efficace et plus naturelle avec l'ordinateur. Or les formidables progrès accomplis dans la miniaturisation des microprocesseurs et dans les réseaux informatiques sans fil permettent d'envisager que la "boîte grise" du ordinateur personnel soit "condamnée à disparaître" (Alberganti 98) ou du moins à ne pas être le seul lieu d'interaction entre l'utilisateur et le monde numérique. Ceci s'inscrit dans le mouvement de l'ordinateur ubiquitaire et évanescent. La multiplication des ordinateurs de poche et des assistants personnels, la généralisation de l'utilisation des cartes à puces et la multiplication des systèmes embarqués dans des objets d'usage courant (automobile, télévision, etc.) constituent des témoins de cet essor. Nous retenons de cet essor que **l'espace d'interaction devient plus vaste, il comprend l'environnement physique et ne se limite plus seulement à un ordinateur sur un bureau**. Ainsi les objets physiques deviennent des supports à l'interaction. Par exemple les Media Blocks (Ullmer et al. 98) reposent sur l'utilisation d'icônes physiques (des cubes) pour manipuler des fichiers informatiques. Quant aux systèmes de chirurgie assistée par ordinateur, ils utilisent des localisateurs 3D pour localiser les outils chirurgicaux dans des images médicales acquises en phase préopératoire. L'espace d'interaction sort alors définitivement de l'espace composé par le bureau et l'ordinateur.

Dans ce contexte, la réalité augmentée (RA) est un paradigme d'interaction qui est né de cette volonté de fusionner les capacités de traitements informatiques et l'environnement physique. En réponse à l'avancée des nouvelles technologies d'interaction, les applications reposant sur le paradigme de RA sont de plus en plus diversifiées. Nous distinguons deux classes d'application, selon que le monde réel est support à l'interaction, ou qu'il est objet de la tâche. Dans les deux cas, l'objectif est de combiner les deux mondes, réel et virtuel (informatique). Pour notre étude, nous considérons donc d'une part les modalités d'interaction en entrée et en sortie reposant sur des objets physiques de l'environnement : l'environnement est alors support à l'interaction avec un ordinateur. Dans le cas de tâches dans le monde réel, nous étudions la composition de modalités de sortie (à destination de l'utilisateur) avec le monde réel perçu par l'utilisateur. Ainsi nous faisons l'analogie entre percevoir un énoncé multimodal, et percevoir le monde réel augmenté par l'ordinateur grâce à la mise en œuvre de modalités de sortie. Enfin nous visons la conception de modalités qui permettent à l'utilisateur de manipuler des entités à la fois réelles et virtuelles afin de favoriser la fluidité d'actions entre les deux mondes, réel et virtuel. Le chapitre 5 est dédié à nos travaux sur la réalité augmentée.

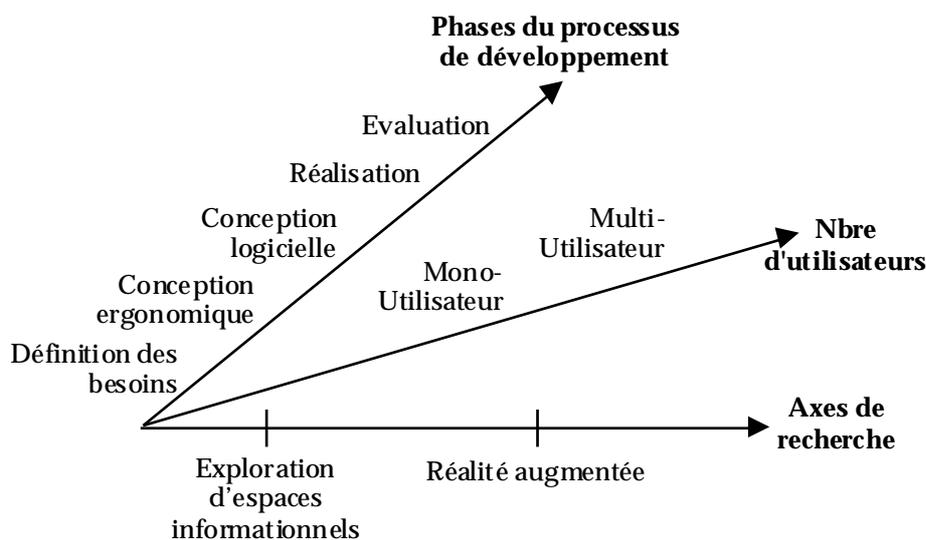


Figure 1 : Contours de nos travaux sur les modalités d'interaction et la multimodalité.

La Figure 1 dessine les contours de nos travaux, selon les deux axes de recherche ci-dessus, les phases du processus de développement logiciel, et le fait que le système est mono ou multi-utilisateur. Plusieurs domaines d'application servent à la fois de guide et de terrain de validation pour ces travaux. Nous listons les domaines et les systèmes réalisés, multimodaux en entrée et/ou en sortie ; la Figure 2 les situe par rapport à notre cadre de recherche.

- Visualisation de grands espaces d'information comme les résultats des moteurs de recherche sur la Toile (système VITESSE), une hiérarchie de fichiers (système PARENT), une table (système MulTab),
- Visualisation de données temporelles pour les Systèmes d'Information Géographique (système INVEST),

- Navigation collaborative sur la Toile (système CoVITESSE),
- Geste médical assisté par ordinateur (système CASPER),
- Situation collaborative mobile lors de fouille archéologique (Plate-forme MAGIC).

Le caractère mobile de notre plate-forme MAGIC est à souligner. La mobilité des utilisateurs est un trait de classification des systèmes interactifs. Nous ne l'avons pas fait apparaître aux Figures 1 et 2 car seule la plate-forme MAGIC constitue un exemple de système sur support mobile. En conclusion générale à nos travaux, nous exposons une étude de la multimodalité sur support mobile que nous venons d'entreprendre.

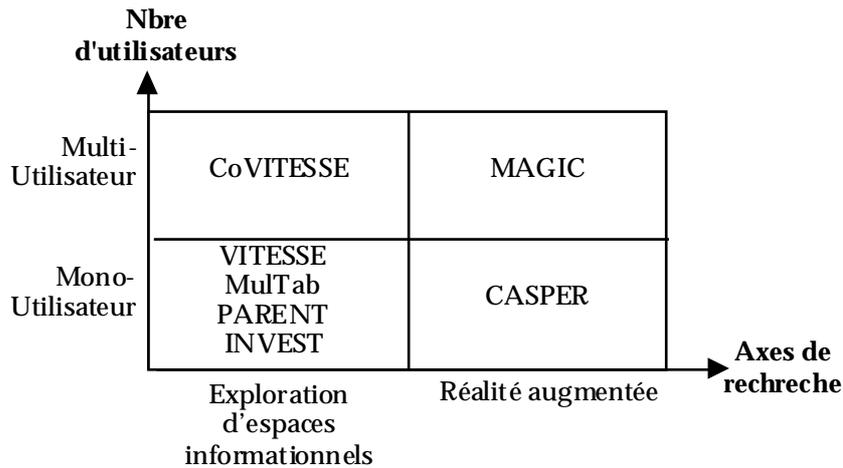


Figure 2 : Systèmes démontrant les résultats conceptuels.

La structure en 5 chapitres de ce mémoire est la suivante : nous situons d'abord nos travaux à la croisée des sciences cognitives et du génie logiciel, puis au chapitre 3, nous asseyons la terminologie et détaillons notre cadre général de conception de la multimodalité. Nos résultats concernant les deux axes de recherche décrits ci-dessus sont ensuite présentés aux chapitres 4 et 5. Ces deux derniers chapitres ont une structure commune, une introduction au domaine, nos contributions et nos perspectives.

2 A la croisée des sciences cognitives et du génie logiciel

Nos travaux concernent l'ingénierie de l'Interaction Homme-Machine ; deux disciplines essentielles de l'IHM sont donc concernées : les sciences cognitives et le génie logiciel. Dans (Bellotti et al. 96), nous avons souligné l'apport d'une approche pluridisciplinaire pour enrichir l'espace de conception, en appliquant des modèles analytiques des deux disciplines comme CTA (Cognitive Task Analysis) (Barnard & May 93) et PUM (Programmable User Models) (Young et al. 89) ainsi que notre modèle d'architecture logicielle PAC-Amodeus (Nigay 94). Le principe directeur de nos activités est la prise en compte, par les logiciels, des résultats exploitables des sciences cognitives. Pour cela, s'appuyant sur notre démarche de travail exposé au chapitre 1, nos travaux sur la conception et la modélisation logicielles des systèmes interactifs visent un triple objectif :

1. définir des propriétés issues des sciences cognitives et du génie logiciel. Dans notre livre collectif (Gram & Cockton 96), nous définissons des propriétés externes caractérisant l'interaction qui sont issues de la psychologie cognitive et de l'ergonomie cognitive ainsi que des propriétés dites internes décrivant le logiciel et appartenant au génie logiciel. Pour illustrer ces deux types de propriétés, nous considérons les exemples suivants :
 - Propriété externe : Observabilité
L'observabilité caractérise la capacité du système à garantir que l'utilisateur peut percevoir l'état interne du système au travers des informations qu'il fournit.
 - Propriété interne : Extensibilité
L'extensibilité caractérise la facilité avec laquelle un logiciel se prête à une modification ou à une extension des fonctions qui lui sont attribuées.
2. comprendre et structurer l'espace problème des interfaces en des termes qui soient communs aux deux disciplines : les sciences cognitives et l'ingénierie des logiciels. Les modèles visés doivent permettre aux concepteurs ergonomes et psychologues, tout comme les concepteurs de logiciel, de raisonner sur le domaine des interfaces concernées ;
3. fournir un espace solution à la fois conceptuel et opérationnel, en accord avec les besoins de l'espace problème. Bien que nous visions l'intégration de recommandations issues des sciences cognitives, les modèles solutions s'adressent, pour l'essentiel, aux besoins de l'ingénierie des logiciels. Nous voulons des modèles conceptuels, c'est-à-dire indépendants d'une technologie particulière, mais aussi opérationnels, applicables par heuristique au cas de la conception logicielle d'une interface utilisateur donnée. Nous établissons des modèles conceptuels d'architecture logicielle, aides à l'organisation modulaire du logiciel de l'interface d'un système, comme PAC-Amodeus (Nigay 94) (Nigay & Coutaz 95) et PAC* (Calvary et al. 97). Dans (Coutaz & Nigay 01), nous proposons une revue des modèles d'architecture applicables aux systèmes interactifs.

Chacun de nos objectifs se situe à la croisée des pratiques des sciences cognitives et du génie logiciel. Tandis que l'apport des sciences cognitives pour la conception de l'IHM est certes évident (Objectif 1 : propriétés ergonomiques, et Objectif 2 : espace de conception), l'établissement de liens entre les propriétés externes de l'IHM et l'architecture logicielle (Objectif 3) constitue un point d'ancrage des sciences cognitives dans le processus de développement qui est étudié depuis les années 80 mais qui est encore mal cerné. Dans (Nigay & Coutaz 97), nous avons souligné l'importance de ce point d'ancrage : un modèle d'architecture logicielle constitue un outil dédié à la phase de spécification globale du logiciel, phase de transition dans le processus de développement, de l'espace Interface Homme-Machine (IHM) vers l'espace logiciel. En effet, il est classique de répartir les étapes du processus de développement d'un système en deux espaces : l'espace IHM et l'espace logiciel. Le premier se caractérise par la priorité qu'il faudrait accorder aux aspects ergonomiques, le second par l'accent mis sur les techniques d'implémentation logicielles. L'espace IHM inclut en particulier l'analyse des besoins et la conception de l'interface consignée dans le document de Spécifications Externes. A cet étape de conception, intervient l'ergonomie pour la définition et la spécification de l'interface utilisateur. L'espace logiciel laisse la place aux compétences informatiques avec les

conceptions globales et détaillées, le codage et les tests unitaires et d'intégration. Aussi avec la conception globale du logiciel reposant sur un modèle d'architecture conceptuel, nous quittons l'espace de conception proprement dit de l'Interface Homme-Machine pour pénétrer dans l'espace logiciel réservé aux informaticiens. Cette situation particulière dans le processus de développement se traduit par une tension permanente entre d'une part la satisfaction des requis de l'utilisateur, des spécifications externes de l'IHM et donc des propriétés externes de l'IHM et d'autre part les contraintes techniques imposées par les outils de mise en œuvre, la satisfaction des requis de qualité logicielle et donc des propriétés internes.

De cette situation charnière entre l'espace IHM et l'espace logiciel, nous concluons l'importance d'établir des liens entre les propriétés externes et l'architecture logicielle. Comme souligné dans (Bass & John 01a), cette conclusion va à l'encontre des nombreux travaux sur les architectures logicielles des années 80 et 90. Certes, établir des liens entre l'utilisabilité et l'architecture logicielle a été le sujet d'étude de nombreux travaux. Néanmoins l'approche adoptée a été de traduire l'utilisabilité en terme de modifiabilité du code, incontournable dans le cadre d'un processus de conception itérative centrée sur l'utilisateur. Ainsi les modèles de référence répondent à cette exigence de modifiabilité avec le principe de séparation fonctionnelle. Un premier grain de séparation est la distinction entre les services d'une application (noyau fonctionnel) et les fonctions chargées d'assurer l'interaction avec l'utilisateur. Le modèle de référence Seeheim (Pfaff 85) fournit une structure fonctionnelle canonique à gros grain qui repose sur cette distinction entre le noyau fonctionnel et le logiciel de l'IHM.

Dans (Nigay & Coutaz 97), l'approche adoptée va au-delà de la traduction réductrice de l'utilisabilité en terme de modifiabilité du code, et nous présentons des motifs architecturaux PAC-Amodeus (Nigay 94) directement dérivés de propriétés externes. Par exemple le motif "Vue multiple" correspond à la propriété externe "Multiplicité des représentations", cette propriété participant à la flexibilité de l'IHM. A plus long terme, notre objectif lié au modèle d'architecture logicielle PAC-Amodeus est de coupler des patrons de conception d'IHM (Bayle 97) avec des patrons de conception logicielle PAC-Amodeus. Nous avons esquissé cette approche dans la Thèse de F. Vernier (Vernier 01) pour le cas de la visualisation de grandes quantités d'information. Le modèle PAC-Amodeus n'est certes pas le seul modèle d'architecture avec lequel nous pouvons entreprendre ces travaux. Néanmoins, dans le cas de la visualisation, ce modèle hybride semble être particulièrement adapté car il localise les étapes du processus de visualisation dans des composants logiciels distincts et supporte un affinement en agents logiciels des techniques de visualisation et de leurs compositions.

Dans (Bass & John 01a), nous trouvons une nouvelle approche pour coupler les aspects d'utilisabilité liés aux propriétés externes et les aspects d'architecture logicielle. Cette approche est générale car elle ne dépend pas d'un modèle d'architecture particulier. Elle repose sur des scénarios compréhensibles par les concepteurs ergonomes et psychologues, tout comme les concepteurs de logiciel. Ces scénarios constituent le point d'ancrage entre les aspects d'utilisabilité et les mécanismes architecturaux. Par exemple, un scénario (Bass & John 01b) est le suivant.

2.21 Supporting Undo

A user performs an operation, then no longer wants the effect of that operation. For example, a user may accidentally delete a paragraph in a document and wish to restore it. The system should allow the user to return to the state before that operation was performed. Furthermore, it is desirable that the user then be able to undo the prior operation (multi-level undo).

Du point de vue utilisabilité, ce scénario est mis en relation avec les bénéfices escomptés d'utilisabilité, que nous assimilons aux propriétés externes. Ces bénéfices incluent par exemple :

Increases individual effectiveness

Du point de vue architectural, le scénario ci-dessus est mis en relation avec le mécanisme architectural suivant :

Recording: This mechanism records system state periodically for further use. Some of the variables that are dependent upon the particular application of the mechanism are

- *the frequency with which the state is recorded*
- *the actual state recorded*
- *the use to which the recorded state is put*
- *the persistence of the data recorded*
- *the consistency of the data recorded*

La différence essentielle entre cette approche (Bass & John 01a) et celle que nous avons adoptée (Nigay & Coutaz 97) réside dans l'indépendance par rapport à un modèle d'architecture logicielle. Tandis que nos travaux sont dépendants du modèle PAC-Amodeus, (Bass & John 01a) définissent un ensemble de mécanismes architecturaux généraux indépendants d'un modèle particulier. Les deux approches se complètent :

- D'une part, une extension à nos travaux serait de traduire les mécanismes architecturaux proposés au sein d'une architecture PAC-Amodeus.
- Réciproquement, certains mécanismes, comme notre moteur de fusion pour la multimodalité en entrée, constituent des extensions à la liste des mécanismes architecturaux fournie dans (Bass & John 01a). En effet dans le chapitre "From HCI properties to Software Design" (Nigay & Coutaz 95b), nous soulignons l'articulation entre les propriétés ergonomiques CARE et le moteur de fusion ancré dans une architecture PAC-Amodeus.

Nous venons d'exposer les objectifs de notre recherche, à la croisée des pratiques des sciences cognitives et du génie logiciel. Ce faisant, nous avons justifié notre démarche de travail exposé au Chapitre 1. Les objectifs cernés, nous présentons maintenant les fondements de nos travaux.

3 Modalité d'interaction : définition et cadre de conception

3.1 Définition : modalité d'interaction et multimodalité

Dans (Nigay & Coutaz 96), nous proposons une revue terminologique des termes modalité, média et mode. Parce que nos travaux sont dédiés à la conception d'interfaces, il convient d'adopter une définition de la modalité qui couvre les deux angles d'étude de la conception, l'utilisateur et le système interactif. A cette fin, nous proposons la définition d'une modalité comme le couple $\langle p, r \rangle$ où :

- p désigne un dispositif physique (par exemple, une souris, une caméra, un écran, un haut-parleur),
- r dénote un système représentationnel, c'est-à-dire un système conventionnel structuré de signes assurant une fonction de communication (par ex., un langage pseudo naturel, un graphe, une table).

L'expression suivante fournit une définition complète de la notion de modalité :

modalité ::= <p, r> | <modalité, r>

Par exemple, dans le système MATIS (Nigay 95), un utilisateur peut formuler des requêtes de renseignements sur des horaires d'avion en saisissant au clavier une phrase en langage naturel comme "Flights from Pittsburgh to Boston", ou bien énoncer oralement cette même phrase ou encore remplir un formulaire par manipulation directe. Les couples <clavier, langage naturel>, <microphone, langage naturel> et <souris, formulaire graphique> décrivent ces trois possibilités. Si nous considérons les systèmes présentés dans (Harrison et al. 98) (Fishkin et al. 98) illustrant des interfaces manipulatoires ou interfaces incorporées ("embodied user interface"), l'utilisateur tient dans sa main un ordinateur de poche augmenté de capteurs. Pour parcourir une liste séquentielle, l'utilisateur incline l'ordinateur de poche. La vitesse de défilement est fonction de l'inclinaison c'est-à-dire de l'amplitude du mouvement du poignet. L'utilisateur arrête le défilement en relevant l'ordinateur de poche vers sa position neutre, l'horizontale, ou en exerçant une pression des doigts sur le côté. La modalité d'interaction est alors décrite par le couple <ordinateur de poche, langage gestuel>. Dans notre système de gestion de courriers électroniques (Calvet et al. 01), illustré à la Figure 3, le langage gestuel défini est simple et correspond à la spécification de huit commandes comme l'ouverture d'un message par un mouvement vertical du bas vers le haut de l'ordinateur de poche, ou le défilement de la liste de messages reçus par inclinaison de l'ordinateur de poche.

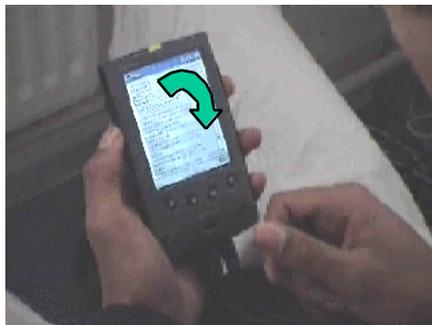


Figure 3 : Interface incorporée. Modalité = <ordinateur de poche, langage gestuel>.

Cette définition d'une modalité constitue une façon de caractériser les échanges entre un système et son utilisateur car elle identifie et met en relation deux niveaux d'abstraction pertinents tant du côté humain que du côté système. Du côté humain, nous distinguons le niveau perception/action lié au dispositif, du niveau cognition en relation avec le système représentationnel. Du côté système, nous différencions le dispositif physique mis en jeu et le format des expressions échangées. Aussi une modalité définie comme le couple <p, r> nous permet d'envisager la définition d'un cadre de conception qui recouvre à la fois les perspectives utilisateur et système en identifiant deux niveaux d'abstraction, physique et logique.

Pour appréhender la multimodalité en entrée (Nigay & Coutaz 95a) comme en sortie (Vernier & Nigay 00a), nous nous basons sur notre définition d'une modalité (le couple $\langle p, r \rangle$), le préfixe multi traduisant le nombre. L'étude de la multimodalité implique donc de considérer le nombre de dispositifs et de systèmes représentationnels distincts, constitutifs des passerelles logiques et physiques entre le système et l'utilisateur. La multimodalité revêt alors les trois aspects suivants :

- Multi-dispositif et mono-système représentationnel
- Multi-dispositif et multi-système représentationnel
- Mono-dispositif et multi-système représentationnel

La composition de modalités graphiques constitue un exemple d'interface multimodale en sortie qui est mono-dispositif (un écran) et multi-système représentationnel (comme un tableau et un histogramme). Notons que dans ce cas, certains auteurs (Maybury 93) préfèrent le terme multicode à multimodal, le terme multimodal étant alors associé au fait qu'il y ait plusieurs dispositifs sollicitant des capacités sensorielles humaines différentes (plusieurs modalités sensorielles).

3.2 Cadre de conception

Nous abordons la conception d'interfaces multimodales en adoptant les deux niveaux d'abstraction d'une modalité : dispositif et système représentationnel. Nous discernons deux volets dans la conception : le choix des modalités et donc des dispositifs et des systèmes représentationnels et leurs compositions. Les propriétés CARE (Complémentarité, Assignation, Redondance et Equivalence) (Nigay & Coutaz 95b) et (Coutaz et al. 95) traduisent ces deux volets de la conception. L'existence de choix de modalités se traduit par l'équivalence, tandis que l'assignation désigne l'absence de choix. Complémentarité et Redondance traduisent la composition de modalités. De nombreux travaux se sont intéressés à caractériser plus finement la composition, c'est-à-dire les relations entre les modalités, comme l'espace UOM (Nigay 94), l'espace MSM (Coutaz et al. 96), et l'espace TYCOON (Martin 99). Nous avons récemment proposé un espace intérateur de ces approches dans (Vernier & Nigay 00a). Notre espace repose sur des aspects et schémas de composition. Nous considérons d'abord cinq aspects d'une composition :

- Composition temporelle de modalités
- Composition spatiale de modalités
- Composition articulatoire de modalités
- Composition syntaxique de modalités
- Composition sémantique de modalités

Les quatre premières compositions retiennent le niveau physique (temporel, spatial et articulatoire) et le niveau logique (syntaxique) d'une modalité : une modalité désigne alors un véhicule de l'information. A l'opposé, la composition sémantique concerne les relations entretenues entre les informations véhiculées par les modalités. Dans une seconde étape, nous considérons ces cinq aspects de composition (temporel, spatial, articulatoire, syntaxique et sémantique) selon les cinq schémas issus des relations d'Allen (Allen 83) :

- Composition de modalités éloignées
- Composition de modalités avec un point de contact
- Composition de modalités avec une intersection non vide
- Composition de modalités dont l'une est plus étendue et englobe l'autre
- Composition de modalités de même étendue

Par exemple, la composition de modalités de même étendue déclinée selon les cinq aspects de composition définit des caractéristiques d'une composition que nous intitulons :

Aspects d'une composition					
	temporel	spatial	articulatoire	syntactique	sémantique
Schéma : Composition de modalités de même étendue	Parallèle Simultanée	Recouvrant	Capacité sensorielle identique	Jumelle	Totalement redondante

Dans (Vernier & Nigay 00a), nous décrivons l'espace complet et l'illustrons au cas de la multimodalité en sortie.

3.3 Perspectives de recherche

L'espace présenté dans (Vernier & Nigay 00a) complète et unifie les espaces existants et s'applique à la multimodalité en entrée comme en sortie. Cet espace s'appuie sur cinq aspects de la composition (temporel, spatial, articulatoire, syntaxique et sémantique). Ces aspects correspondent d'une part au niveau dispositif d'une modalité (aspects temporel, spatial et articulatoire), et d'autre part au niveau système de représentation (aspect syntaxique). Une extension consiste à affiner notre définition d'une modalité en détaillant les caractéristiques du niveau dispositif. Jusqu'à alors, il était classique de distinguer les dispositifs selon les capacités sensorielles mises en jeu (aspect articulatoire), deux autres propriétés des dispositifs seraient donc liées au temps et à l'espace. Dans la taxonomie des dispositifs d'entrée de (Mackinlay et al. 90), nous retrouvons par exemple la notion d'espace, par un opérateur de composition, la disposition, traduisant une composition spatiale de dispositifs. Nous revenons sur ce point dans l'exposé de notre étude de la Réalité Augmentée au paragraphe 5.

Pour la complétude des travaux liés à notre espace de conception, il convient aussi d'étudier les impacts des compositions identifiées à la fois sur la conception logicielle du système et sur l'utilisabilité. Ce travail est en cours, en collaboration avec des collègues du réseau européen TACIT et en particulier des psychologues de l'Université de Sheffield. En effet, cet espace traduit les possibilités offertes en termes de compositions. Néanmoins il ne peut servir de guide à la conception tant que les types de composition ne sont pas caractérisés en terme d'utilisabilité. Nous présentons dans les deux chapitres suivants (Exploration Visuelle d'espaces informationnels et Réalité Augmentée) des résultats concernant certaines compositions. L'espace des possibilités ainsi défini est vaste (5^4 types de composition) et de nombreuses compositions restent à étudier.

Plus généralement, comme nous l'avons annoncé dans le Chapitre 1, la multimodalité a été principalement étudiée pour l'interface en entrée : définition, espace de conception, architecture logicielle et plate-forme de réalisation logicielle comme notre moteur de fusion (Nigay & Coutaz 95a). Renforçant ce constat, les 10 mythes sur la multimodalité (Oviatt 99) ne traitent que de la multimodalité en entrée. De plus il est important de souligner que la plupart des travaux ont focalisé sur la

combinaison de la parole avec un geste de désignation (Oviatt et al. 00) et que de nombreuses autres combinaisons de modalités restent à étudier. Ainsi comme perspectives de recherche sur la multimodalité, (Oviatt et al. 00) soulignent l'intérêt grandissant dans les systèmes multimodaux intégrant des techniques de vision par ordinateur comme le suivi du regard. Ces techniques sont dites passives car elles ne requièrent pas de commandes explicites de la part de l'utilisateur ("perceptual user interfaces" (Turk & Robertson 00)). Combiner les techniques actives de la multimodalité avec celles dites passives permettrait par exemple de rendre le mécanisme d'interprétation de l'énoncé multimodal plus robuste. Outre ces perspectives de recherche sur la multimodalité en entrée, un vaste champ d'études de la multimodalité en sortie reste à explorer. Dans le chapitre suivant, nous présentons nos résultats concernant la multimodalité en sortie pour le cas de modalités graphiques. Par rapport aux interfaces en entrée, la conception et la réalisation d'interfaces multimodales en sortie constituent un problème plus complexe à traiter. Nous expliquons cette complexité accrue en considérant les acteurs du choix des modalités et de leurs usages pour une tâche donnée : les trois cas suivants ne sont pas incompatibles au sein d'une même interface :

1. **Assignment** par le concepteur : les modalités et leurs usages ont été figés par le concepteur.
2. **Adaptativité** : les choix des modalités et de leurs usages sont réalisés par le système grâce à un mécanisme de sélection.
3. **Adaptabilité** : les modalités et leurs usages sont choisis par l'utilisateur. Le système offre alors un ensemble de possibilités quant aux choix des modalités et un mécanisme permettant à l'utilisateur d'effectuer ces choix.

Dans le cas 1, les choix des modalités et de leurs usages sont effectués lors de la phase de conception, tandis que dans le cas 2 les choix sont réalisés à l'exécution pour tenir compte de l'utilisateur en situation réelle. Le mécanisme de sélection repose alors sur des données collectées dynamiquement comme la tâche courante entreprise par l'utilisateur ou le contexte physique, données sur lesquelles se basent les choix des modalités et de leurs usages. Dans le cas 3, les choix des modalités sont aussi accomplis à l'exécution, mais reviennent à l'utilisateur.

Pour les entrées, les choix des modalités et de leurs usages reviennent couramment à l'utilisateur, ce dernier étant le plus apte à choisir le ou les modalités les plus efficaces et adéquates à son contexte courant. Les enjeux de conception résident alors dans la définition d'une panoplie de possibilités et dans la mise en place d'un mécanisme d'analyse des énoncés multimodaux d'entrée pour en déduire l'intention de l'utilisateur. A l'opposé, pour l'interface en sortie et par analogie avec l'interface en entrée, les choix des modalités (génération) et leurs mises en œuvre (présentation) reviendraient au système pour définir un énoncé multimodal en sortie. L'objectif serait alors que le système ait assez de connaissances sur le contexte courant pour choisir le ou les modalités de sortie adéquates. C'est dans l'ampleur des connaissances que le système devrait acquérir et maintenir que réside la complexité de la multimodalité en sortie. Ainsi les systèmes multimodaux en sortie comme COMET (Feiner & McKeown 91) et WIP (André et al. 93) réduisent le contexte à l'état du dialogue. La génération de l'énoncé multimodal s'appuie alors sur la notion d'acte de dialogue. Bien d'autres aspects du contexte peuvent être pris en compte : par exemple le développement récent de l'informatique mobile et ubiquitaire met en

exergue le problème de définition, de capture et de modélisation de l'environnement physique de l'utilisateur qui va bien au-delà de la localisation de l'utilisateur (Gray & Salber 01). Tandis que l'on associe la multimodalité en entrée à l'adaptabilité et la multimodalité en sortie à l'adaptativité, il convient néanmoins de noter que l'adaptabilité en sortie a un sens. Prenons pour preuve le choix laissé à l'utilisateur sur son téléphone mobile quant aux modalités de sortie, entre une modalité sonore ou tactile.

Nous avons précisé dans ce chapitre notre point de vue sur une modalité d'interaction et la multimodalité :

- en définissant une modalité selon deux niveaux d'abstraction, le dispositif et le système représentationnel et,
- en introduisant notre espace de composition de modalités.

La terminologie adoptée et le cadre de conception étant exposés, nous présentons dans les deux chapitres suivants nos résultats pour les deux axes de recherche étudiés : le chapitre suivant est dédié à l'exploration visuelle d'espaces informationnels.

4 Exploration visuelle d'espaces informationnels et modalités graphiques

4.1 Axe de recherche

L'exploration visuelle d'un espace informationnel repose sur une visualisation interactive de l'espace. La visualisation d'information désigne le processus qui transforme des données, informations et connaissances dans une forme qui repose sur les capacités de traitement du système visuel humain pour être perçue et interprétée. Le but est d'aider l'utilisateur à observer, analyser et comprendre un espace informationnel en lui offrant une représentation visuelle de l'espace et des techniques d'interaction.

Axe de recherche récent, il combine des aspects de l'imagerie, du graphisme, de la visualisation scientifique, de la recherche d'information, de la fouille de données, comme de l'Interaction Homme-Machine (Gershon & Page 01). Tandis que la visualisation scientifique, axe de recherche plus ancien, vise à définir des représentations visuelles d'un objet ou d'un phénomène réel mais invisible ou difficilement reproductible à l'aide des techniques d'images traditionnelles comme la trajectoire d'atomes (Huitema & Van Liere 00, session "Scientific applications"), la visualisation d'information focalise sur l'information qui est souvent abstraite (Robertson et al. 93) et non spatiale. Le problème clé de la visualisation d'information est alors d'inventer des métaphores visuelles, une représentation spatiale interactive des informations. Au sein du domaine de l'IHM, la visualisation d'information vise à définir de nouveaux paradigmes de présentation et d'interaction (autre que le traditionnel WIMP-Window Icon Menu Window, Icon, Menu, Pointing device ou Pull-down menu), paradigmes souvent destinés à de grands espaces informationnels.

Axe de recherche important, il concerne de très nombreux domaines d'application, comme la vente, les banques et les assurances. De plus, face à la croissance continue

des quantités d'information manipulées, il convient d'assister les utilisateurs dans leur quête d'information pour des situations chaque fois particulières. En réponse au problème actuel qui n'est plus tant la disponibilité de l'information mais la capacité de sélection d'une information qui réponde au besoin, nous constatons une profusion de techniques de visualisation d'information (Card et al. 99) (Hascoët & Beaudouin-Lafon 01). A titre d'exemple, nous présentons à la Figure 4 notre technique notée Fieldmaps, reposant sur une extension de l'algorithme Treemaps (Vernier & Nigay 00b) pour visualiser un arbre. Cette multiplicité des techniques de visualisation témoigne du dynamisme certain de cet axe d'étude.

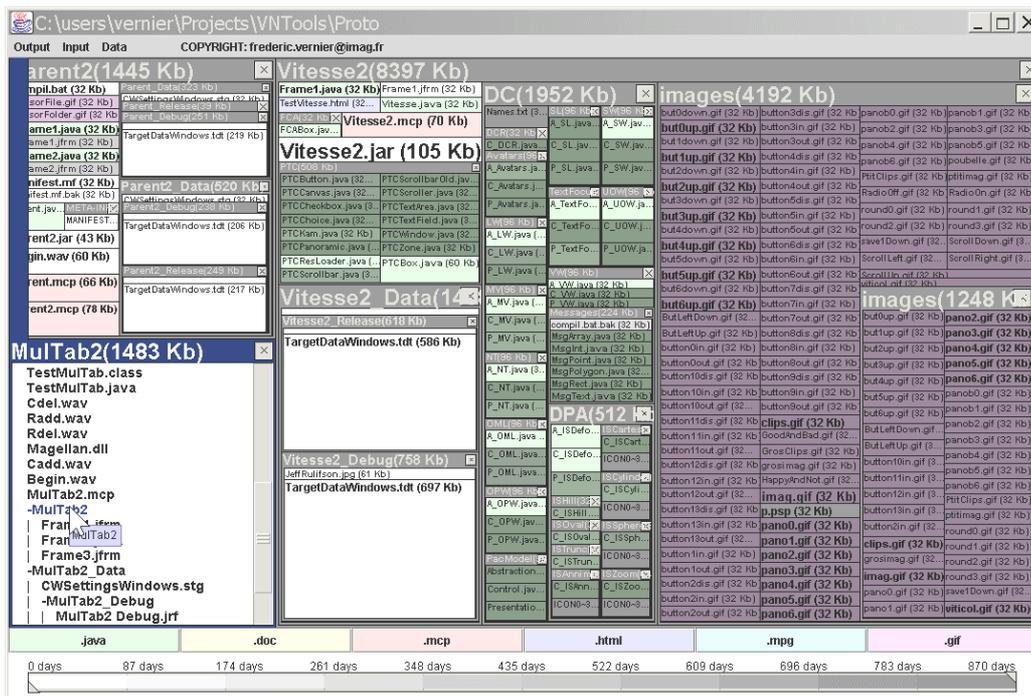


Figure 4 : 325 fichiers et répertoires visualisés par l'algorithme Fieldmaps.

Il est néanmoins important de noter qu'outre la conception de techniques de visualisation dont l'objectif est d'améliorer la qualité perceptuelle et interactionnelle de l'espace d'information, plusieurs autres approches sont envisageables pour assister l'utilisateur dans sa quête d'information : améliorer les moteurs de recherche dans un espace d'information, coupler les termes d'une requête à la représentation des résultats (Mulhem & Nigay 96), permettre l'expression fine des requêtes en langue naturelle augmentée de dialogues de clarification (Rouillard & Caelen 99), déléguer certaines tâches de recherche à des agents logiciels, faire appel au savoir collectif. A l'évidence, toutes ces pistes sont valides et fonctionnent de manière complémentaire pour aider l'utilisateur à explorer un espace informationnel.

Pour nos travaux, nous adoptons une approche centrée sur l'interaction et nous visons à concevoir de nouvelles formes d'interaction multimodale (paragraphe 4.2.1, et 4.2.2) et collaborative (paragraphe 4.2.3) qui aident l'utilisateur à appréhender un espace d'information.

4.2 Contribution à cet axe de recherche

4.2.1 Visualisation de grands espaces d'information

Une grande quantité d'information est constituée d'un grand nombre d'éléments où chaque élément possède des attributs variés : (Shneiderman 96), "I assume that users are viewing collections of items, where items have multiple attributes". Comme le souligne Lisa Tweedie (Tweedie 97) qui se base sur les travaux de Jacques Bertin (Bertin 81), deux formes de données peuvent être représentées : les valeurs des données et la structure des données (relations entre les données de l'ensemble). Le processus d'exploration dans une grande quantité d'information visuelle s'articule en trois étapes énoncées par le "information seeking mantra" (Shneiderman 96) : "Overview first, zoom and filter, then details-on-demand". "overview" signifie avoir une vue d'ensemble de l'espace d'information, tandis que "details-on-demands" correspond à l'obtention par l'utilisateur de détails sur un élément. Les deux verbes, "zoom and filter" correspondent à zoomer et filtrer l'espace d'information, deux classes de techniques d'interaction pour manipuler un grand espace d'information (Hascoët & Beaudouin-Lafon 01). Ces techniques permettent de définir une sous-partie de l'espace d'information, centre d'intérêt de l'utilisateur. De nombreuses techniques de visualisation permettent non seulement à l'utilisateur d'effectuer ces trois étapes du "information seeking mantra", mais aussi visent à intégrer de façon continue et harmonieuse ces trois étapes. Par exemple, comme souligné dans (Keim 01), il est important que l'utilisateur ait une vue d'ensemble de l'espace tout en pouvant analyser une sous-partie de l'espace, le focus.

Dans cet axe de recherche aux avancées disparates, se traduisant par de nombreuses techniques de visualisation présentées isolément, notre contribution concerne l'établissement de référentiels applicables à la conception afin de mieux cerner et comprendre l'espace de conception des interfaces dédiées à de grandes quantités d'information. A ces fins, nous adoptons une démarche d'analyse nouvelle consistant à considérer une technique de visualisation comme une modalité d'interaction de sortie. Nous jetons donc un nouveau regard sur les techniques de visualisation en adoptant le point de vue de la multimodalité en sortie. Notre approche de conception est descendante, centrée sur les tâches-utilisateur (Nigay & Vernier 98a). Nous partons des tâches-utilisateur pour identifier des critères d'ergonomie que doivent satisfaire les techniques d'interaction à concevoir. A partir de ces critères, nous énonçons des règles de conception (Nigay & Vernier 98a) et des patrons de conception (Vernier 01). Au paragraphe 4.2.1.1, nous détaillons notre analyse des tâches. Pour répondre à la diversité des tâches, la composition de techniques de visualisation d'une grande quantité d'information vise à mettre en évidence différentes caractéristiques de l'espace. La composition de techniques de visualisation, présentée au paragraphe 4.2.1.2, est une forme de multimodalité en sortie, mono-dispositif et multi-système représentationnel.

4.2.1.1 Tâches-utilisateur d'exploration de grands espaces informationnels

A plus haut niveau d'abstraction, nous trouvons des tâches qui sont dépendantes du domaine d'application. Ainsi la taxonomie TTT (Shneiderman 96) repose à plus haut niveau d'abstraction sur des tâches propres au domaine d'application. De plus dans (Ignatius et al. 96), il est souligné que l'apport d'une technique de visualisation

s'évalue par rapport aux tâches du domaine d'application. Appliquant une approche d'analyse descendante, les tâches qui sont dépendantes du domaine peuvent être détaillées en tâches non dépendantes du domaine. Nous trouvons ainsi dans la taxonomie TTT (Shneiderman 96) sept tâches : "overview, zoom, filter, details-on-demand, relate, history and extract". Une autre liste de tâches est présentée dans (Gershon et al. 97). Néanmoins si l'on vise à l'exhaustivité, identifier les tâches-utilisateur sans considérer un domaine particulier semble très ardu. L'analyse commencée lors d'une réunion du groupe de travail FADIVA confirme la difficulté de l'entreprise (Ioannidis 96). Des exemples de tâches sont les suivants :

- trouver un ensemble d'éléments par exploration,
- retrouver un ensemble d'éléments,
- identifier une tendance dans l'ensemble,
- comparer des éléments.

Notre approche consiste alors à raisonner au niveau de l'interaction, et à considérer que quelle que soit la tâche à accomplir, l'utilisateur devra naviguer dans l'espace (Nigay & Vernier 98b). Nous avons cerné les tâches de navigation en nous appuyant sur le modèle d'exploration d'information de (Waterworth & Chignell 91). Nous étudions la navigation pour identifier des critères d'ergonomie que doivent satisfaire les techniques d'interaction à concevoir. Dans (Nigay & Vernier 98), nous présentons des critères issus de notre étude de la navigation, comme l'observabilité du focus et du contexte.

Notre approche de conception basée sur les tâches-utilisateur est une vue réductrice de l'exploration, puisque nous considérons que la tâche élémentaire de navigation dans l'espace, tâche de manipulation de l'espace. Par exemple est exclu le fait d'évaluer la pertinence d'un élément. De plus, une autre approche complémentaire s'inscrit dans la cognition écologique, qui souligne l'adaptation du comportement humain, comme la théorie "Information Foraging" (Pirolli & Card 95). Cette théorie couvre les activités de recherche, de manipulation et d'évaluation de l'information. Le comportement humain n'est plus expliqué uniquement par la tâche à réaliser ; il est décrit comme adaptatif et expliqué par différentes stratégies. L'adaptation est guidée par le coût d'accès et d'extraction de l'information par rapport à la valeur escomptée de l'information. Récemment dans (Blafard & Stelmazewska 01), il a été montré des cas d'adaptation du comportement humain qui ne peuvent être décrits par les théories existantes. Un exemple est le comportement réactif de l'utilisateur lorsque l'interaction est trop lente, qui consiste à changer de stratégie et à travailler dans plusieurs fenêtres. La charge cognitive est conséquemment plus élevée, l'utilisateur passant d'une fenêtre à l'autre. Une implication pour la conception de l'interface consiste alors à maximiser l'observabilité des différents états liés aux fenêtres.

4.2.1.2 *Composition de modalités graphiques informationnelles*

Nous avons identifié deux types de composition de modalités graphiques : les vues multiples proposant plusieurs techniques de visualisation pour un espace informationnel et la composition par complémentarité de deux modalités, l'une pour le centre d'intérêt et l'autre pour la vue d'ensemble de l'espace d'information.

- La représentation multiple d'une grande quantité d'information vise à mettre en évidence différentes caractéristiques de l'espace pour répondre à la diversité des

tâches. Plusieurs présentations de la même information permettent d'effectuer des tâches différentes dans les meilleures conditions. ["Different presentations of the same information best support different tasks". (Tweedie 97)]. Dans (Vernier 01), une typologie de la représentation multiple est proposée ; quatre types de représentation multiple sont définis au regard des quatre étapes structurant le cadre d'étude fédérateur du processus de visualisation.

- Une autre forme de multimodalité en sortie réside dans la composition d'une modalité dédiée au focus de l'utilisateur, où les données sont détaillées, avec une modalité assignée à présenter la vue d'ensemble, où les données sont moins détaillées. En nous basant sur notre cadre de conception (paragraphe 3.1), cette forme de composition est complémentaire ou complémentaire/redondante au niveau sémantique, chaque modalité étant dédiée à la présentation d'une partie de l'espace informationnel. L'aspect spatial de cette composition est crucial. Il convient de distinguer les systèmes où le focus est spatialement intégré à la vue d'ensemble comme les techniques déformantes des systèmes Mur en Perspective (Mackinlay et al. 91) ou VITESSE (Vernier & Nigay 97), des systèmes où le focus n'est pas intégré au contexte global comme le système DragMag (Ware & Lewis 95). Nous considérons enfin la composition d'une modalité pour la vue d'ensemble avec plusieurs modalités pour présenter différemment plusieurs foci (Leung & Apperley 94).

Nous proposons des règles et patrons de conception liés à la composition de modalités graphiques (Nigay & Vernier 98a) (Vernier 01), comme la continuité temporelle, préconisant une continuité visuelle lors du passage d'une représentation à une autre. Nous montrons enfin comment vérifier ces patrons de conception lors de la réalisation logicielle selon le modèle PAC-Amodeus (Chapitre 2, A la croisée des sciences cognitives et du génie logiciel). Plus généralement, la mise en correspondance des étapes du processus de visualisation avec les composants du modèle d'architecture logicielle PAC-Amodeus permet de proposer une solution générique à la conception logicielle du processus de visualisation (Vernier 01). Plusieurs systèmes ont été développés comme terrain de validation pour nos travaux, chacun manipulant des données de types différents (1D, 2D et arbre) :

- VITESSE : Visualisation des réponses à une requête soumise à un moteur de recherche (Vernier & Nigay 97)
- MulTab : Visualisation de tableaux avec plusieurs foci (Vernier & Nigay 00a)
- Parent : Visualisation d'une hiérarchie de fichiers et de répertoires (Figure 4) (Vernier & Nigay 00b).

4.2.2 Visualisation de données temporelles

S'appuyant sur nos travaux sur la visualisation de grands espaces informationnels qui sont indépendants du type des données manipulées, nous étudions la visualisation de données temporelles. Plusieurs techniques de visualisation dédiées aux données temporelles ont été développées comme LifeLines (Plaisant et al. 96) ou la technique en spirale de (Carlis & Konston 98). Les données temporelles sont présentes dans un grand nombre de systèmes utilisant des bases de données, comme la finance et la comptabilité. Parmi ces systèmes, les systèmes d'information géographique constituent notre domaine d'application.

Notre démarche de conception repose sur le type particulier des données manipulées et les tâches de l'utilisateur géographe. Nous étudions d'une part les caractéristiques des données temporelles, et d'autre part les besoins et les tâches d'un utilisateur dans un processus d'analyse de données géographiques. Un exemple de tâche identifiée est : Etudier la corrélation de l'évolution des données. Nous traduisons ces besoins en termes de propriétés ergonomiques, ce qui constitue un premier pas vers la conception. Nous avons conçu et développé plusieurs techniques de visualisation de données temporelles présentées dans (Daassi et al. 01) comme la technique des cercles concentriques. L'objectif est d'intégrer ces techniques, au sein d'une plateforme logicielle extensible, INVEST, afin d'offrir à l'utilisateur un ensemble de modalités graphiques (représentation multiple de l'ensemble des données) répondant à la variété des tâches. Comme souligné dans (Keim 01), l'outil visé ne se substitue pas aux méthodes d'analyse statistique couramment utilisées par les géographes, mais au contraire les complète.

Complémentaire à INVEST, nous avons conçu une technique d'exploration dans une base de données orientées objets temporels, nommée "Pointwise Temporal Object Database Browsing" (Daassi et al. 01). La technique permet de passer d'un objet à l'autre selon les relations existantes mais aussi de naviguer dans le temps. Cette technique est complémentaire à celles d'INVEST. En effet INVEST offre plusieurs techniques interactives de visualisation d'historiques d'attributs, ou ensembles de données temporelles, résultats d'une requête soumise à la base de données.

4.2.3 Exploration collaborative d'espaces d'information

Travaux complémentaires à la visualisation d'information, nous visons la conception de formes de collaboration synchrone pour explorer un espace. La collaboration en phase d'exploration peut présenter a priori beaucoup d'intérêts. Par exemple, elle peut permettre aux utilisateurs :

- de gagner en qualité du résultat obtenu en bénéficiant des connaissances et compétences des autres participants, que ce soit sur le domaine de la recherche ou sur la technique d'interaction ;
- de gagner du temps en se partageant le travail d'exploration ;
- de gagner en satisfaction, le travail collectif étant souvent plus valorisant et plus agréable qu'un travail individuel.

Exploitant nos résultats sur la navigation dans un espace informationnel (Paragraphe 4.2.1.1), nous proposons un modèle de navigation, qui se décline en quatre modes de navigation, pour mettre en œuvre la navigation collaborative synchrone. Ce modèle est présenté dans (Laurillau & Nigay 00). L'activité collaborative offre une nouvelle dimension à l'information, la valeur sociale de l'information. Cette forme de navigation collaborative est qualifiée de navigation sociale (Dourish & Chalmers 94). Dans (Dieberger 97), nous trouvons deux formes de navigation sociale :

- La navigation sociale indirecte : celle-ci est accomplie en laissant son empreinte, une marque de son passage. Sa réalisation est aussi influencée par les traces des autres qui ont emprunté le même passage. Ces traces peuvent prendre des formes multiples. Par exemple, cela peut être un compteur qui recense le nombre d'individus ayant visité un lieu dans l'espace.

- La navigation sociale directe : celle-ci se réalise par la présence simultanée de plusieurs utilisateurs et l'emploi de moyens de communication permettant une interaction instantanée. Cette interaction n'existera que s'il se crée une communauté ayant des centres d'intérêts proches dans un même espace informationnel.

Nos travaux focalisent sur l'aspect synchrone de la navigation, la navigation sociale directe, car nous agissons au quotidien collaborativement et de façon synchrone : par exemple, en échangeant des informations autour de la fameuse machine à café. Selon Twidale (Twidale et al. 97), "*Browsing is a Collaborative Process*". Pourtant nous constatons que la majorité des moyens de navigation collaborative sont asynchrones (navigation sociale indirecte) : le courrier électronique, le filtrage collaboratif ou l'actuelle forme de navigation sur le WWW (par exemple, la rubrique "mes liens préférés" est un élément de navigation collaborative asynchrone).

La Figure 5 schématise l'approche adoptée pour définir notre modèle de navigation sociale directe. La dernière étape, en bas de la Figure 5, a consisté à instancier notre modèle de navigation au cas de l'espace des résultats d'une requête soumise à un moteur de recherche sur le WWW : le système Co-Vitesse (Laurillau 99). Reposant sur le système de visualisation VITESSE (Paragraphe 4.2.1.2), le système CoVITESSE intègre les quatre formes de navigation synchrone de notre modèle. CoVITESSE a fait l'objet d'une évaluation empirique préliminaire dont les résultats sont décrits dans (Laurillau & Nigay 00).

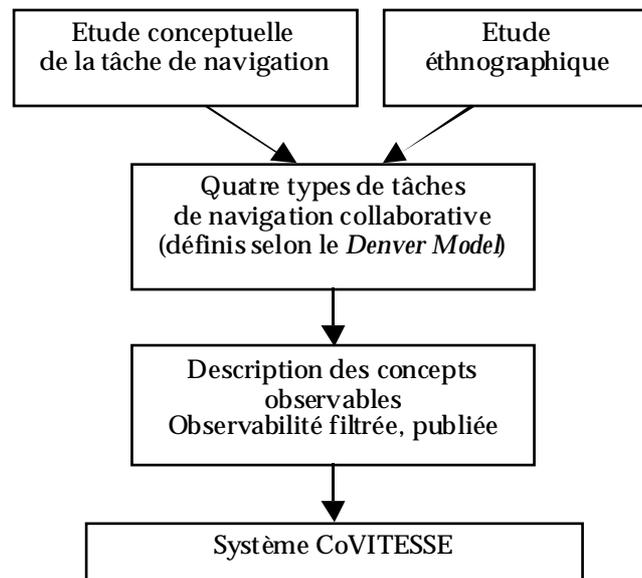


Figure 5 : Exploration collaborative d'un espace d'information : démarche de travail.

4.3 Perspectives de recherche

Les principales perspectives de recherche liées à nos travaux sur l'exploration visuelle multimodale et collaborative d'espaces informationnels concernent l'évaluation. Il convient en effet de conduire des évaluations empiriques des techniques développées. VITESSE a déjà fait l'objet d'une évaluation qui a souligné l'apport de la représentation multiple. Les résultats sont décrits dans (Nigay &

Vernier 98). Parent, Multab, INVEST et CoVITESSE ont fait l'objet d'évaluations informelles. Pour MulTab, nous envisageons de tester aussi l'adéquation des modalités d'entrée aux modalités de sortie. Les paramètres de l'évaluation, les modalités d'entrée/sortie, sont présentés à la Figure 6. Les hypothèses qui motivent l'expérimentation sont les suivantes :

- Le Magellan est plus adapté à la manipulation de l'espace vu en colline (dispositif physique qui rappelle la forme manipulée à l'écran) : facilité d'apprentissage et performance accrues.
- Le Magellan et la souris ou le clavier et la souris peuvent faire l'objet d'usage synergique pour se déplacer tout en sélectionnant (Main " faible " avec le Magellan ou le clavier et main " forte " avec la souris). Un gain de performance est escompté grâce à l'usage synergique.

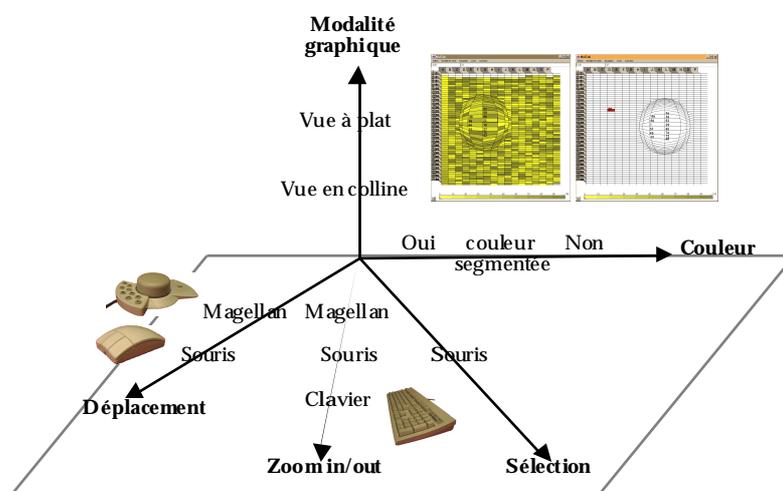


Figure 6 : Modalités d'entrée et de sortie de MulTab.

L'évaluation de techniques de visualisation, qu'elle soit analytique ou empirique, constitue une entreprise difficile ; or nous constatons de nombreuses techniques de visualisation présentées isolément et le besoin de méthodes d'évaluation systématique est clairement apparent. Les rares publications relatant une évaluation commencent toutes par souligner le besoin d'évaluation et de méthodes systématiques (Stasko et al. 00) (Plénacoste et al. 01). "Clearly, further work evaluating information visualizations is necessary. New visualization techniques, no matter how innovative, are not valuable unless they provide true utility and assist people with real tasks." (Stasko et al. 00). Certaines évaluations conduites comparent une version visuelle du système à une version non visuelle comme celle présentée dans (Turo & Johnson 92), tandis que d'autres comparent deux techniques de visualisation comme les cartes d'arbres ("treemaps") et les disques d'arbres (Stasko et al. 00). Enfin dans (Hascoët 98), nous trouvons deux évaluations, l'une analytique et l'autre empirique de deux visualisations qui de par leurs résultats contradictoires ont permis d'identifier un critère analytique. Hascoët explique l'absence de méthodes systématiques (critères, tâches, données, etc.) par le peu d'évaluations conduites. Nous pensons que notre approche basée sur les tâches et les critères d'ergonomie peut fournir des éléments pour l'évaluation. Enfin les deux patrons de conception liés à la composition de modalités graphiques (continuité visuelle et temporelle) définissent des propriétés

souhaitables d'une visualisation interactive indépendante de la tâche, au même titre que l'occlusion et le croisement de lignes (Keim 01).

L'évaluation de techniques collaboratives soulève aussi des problèmes. Par exemple l'approche consistant à observer un groupe d'utilisateurs est très difficile et coûteuse à appliquer (Baker et al. 01), comme nous l'avons constaté dans nos tests informels de CoVITESSE. Une approche nouvelle proposée par (Baker et al. 01) repose sur des heuristiques. L'approche elle-même n'est pas nouvelle, mais les heuristiques proposées sont nouvelles et adaptées au travail de groupe. Nous envisageons d'appliquer cette méthode à CoVITESSE.

Au-delà de l'évaluation des techniques développées, nous avons entamé des travaux sur l'architecture logicielle des systèmes collaboratifs, dans le cadre du développement de CoVITESSE, complétant nos résultats précédents PAC* (Calvary et al. 97). Tandis que PAC*, modèle d'architecture conceptuelle logicielle dédié aux collecticiels, étend le modèle PAC-Amodeus en structurant le composant Contrôleur de Dialogue, nos travaux actuels se concentrent sur le composant Noyau Fonctionnel. L'implémentation en cours de ce modèle d'architecture définit une plate-forme de développement et d'exécution extensible d'un collecticiel.

5 Réalité augmentée et modalités d'interaction

5.1 Axe de recherche

La réalité augmentée (RA) est un paradigme d'interaction qui est né de la volonté de fusionner les capacités de traitements informatiques et l'environnement physique. L'un des systèmes pionniers de RA est le Bureau Digital (Wellner 91) développé au centre de recherche de Xerox. Ce système conserve le document papier et le crayon et permet par exemple d'effectuer des opérations comme le copier/coller que l'on effectue habituellement avec un éditeur de texte. Un numéro spécial de la revue Communications of the ACM consacré à la RA a été édité en 1993 par P. Wellner, R Gold et W. Mackay (Wellner et al. 93) : un courant de recherche démarre.

Une première approche pour cerner ce paradigme de RA consiste à le comparer avec celui de réalité virtuelle (RV). En RV, l'utilisateur est dans un monde entièrement reconstitué par des données informatiques : l'utilisateur est coupé du monde réel. La RV vise donc à immerger l'utilisateur dans un monde artificiel, à l'opposé de la RA dont l'essence même est de maintenir l'utilisateur au contact de l'environnement réel. Une approche originale (Koleva et al. 00) vise à combiner la RA et la RV en étudiant les frontières possibles entre le monde réel et le monde virtuel, caractérisées d'interfaces traversables.

Tandis qu'en IHM, on s'accorde à opposer la RA et la RV, à la même époque au début des années 90, un autre courant de recherche aussi appelé RA était issu de la RV. Courant de recherche appartenant à la synthèse d'images, l'objet d'étude est la fusion harmonieuse d'images de synthèse avec des images du monde réel. De "vrais" pixels d'origine vidéo sont combinés à des pixels de synthèse (Azuma 93). Par exemple, des meubles de synthèse sont agencés par manipulation directe dans le film vidéo d'une cuisine réelle. L'utilisateur manipule des objets purement électroniques

dont certains représentent un monde qui existe quelque part (la cuisine en question), mais qui n'appartiennent pas au lieu physique actuel de l'interaction, l'agencement de la cuisine se faisant, par exemple, dans un magasin.

Aujourd'hui, dans la communauté IHM, la RA est un axe de recherche en plein essor qui s'inscrit dans le mouvement de l'ordinateur ubiquitaire et évanescent : "La boîte grise est condamnée à disparaître." (Alberganti 98) (Chapitre 1). Les publications de recherche présentées à la dernière conférence ACM CHI (Jacko et al. 00) confirment ce constat, avec une session dédiée aux interfaces tangibles, une aux écrans publiques, une à l'informatique mobile, une à l'informatique à domicile. De plus il est à noter la création d'une conférence dédiée à la RA, DARE (Designing Augmented Reality Environment) (MacKay 00a).

En parallèle des travaux en IHM, de nombreux domaines d'application ont vu un intérêt dans l'application de ce paradigme : de nombreux Gestes Médico-Chirurgicaux Assistés par Ordinateur (GMCAO) reposent sur ce paradigme de RA (Troccaz et al. 97), et les graphistes tentent de l'utiliser également (Whitaker et al. 95). Chaque domaine d'application adapte la RA à ses besoins et développe des principes de conception ad hoc. La réalité augmentée constitue donc un domaine évolutif tiraillé entre la pratique créative dirigée par le domaine d'application et aiguillonnée par les nouvelles technologies ("Technology Push") d'une part, et l'approche conceptuelle et mesurée de la théorie d'autre part. Pratique et théorie sont les deux approches complémentaires qui servent de piliers aux démarches de techniques d'interaction ouvrant des voies nouvelles dont on ne cerne pas les limites. Prenons pour preuve le flou d'une terminologie mal assise qui fléchit au gré de la mode : interfaces tangibles (Ishii & Ullmer 97), réalité virtuelle, mixte (Milgram & Kishino 94) ou augmentée, environnement et objet augmentés (Mackay 96) vidéo augmentée, bit/atomes, etc. Un cadre de définition consensuel est donc nécessaire pour comparer les systèmes existants et capitaliser les expériences de conception, en vue de l'établissement de méthodes et principes généraux de conception, en d'autres termes pour appréhender de façon systématique la fusion des deux mondes, réel et virtuel (informatique).

5.2 Contribution à cet axe de recherche

Nos travaux sont motivés par le fait que l'intégration de l'environnement physique de l'utilisateur dans l'espace d'interaction ouvre des voies nouvelles pour la conception de modalités en entrée comme en sortie. Mais les limites sont mal cernées et notre première contribution est un espace taxinomique, identifiant les caractéristiques intrinsèques d'un système de RA (paragraphe 5.2.2). Afin de mieux cerner la conception d'un système de RA, la notation ASUR permet de modéliser un système en termes d'entités et de relations (paragraphe 5.2.3.1). Une modélisation ASUR permet de raisonner sur les deux types d'interaction coexistant dans un système de RA : l'interaction avec l'environnement réel et l'interaction avec le monde informatique. Nous caractérisons la fusion harmonieuse des deux mondes, réel et virtuel, par les propriétés de compatibilité et de continuité (paragraphe 5.2.3.2). A partir d'une modélisation ASUR, une démarche d'analyse de ces propriétés est expliquée. Nous proposons enfin deux concepts génériques de RA qui visent la fluidité entre les deux mondes. Comme annoncé au Chapitre 1, notre contribution

concerne outre la conception ergonomique, la conception logicielle de systèmes de RA. Au paragraphe 5.2.4, nous présentons deux solutions architecturales générales résultant de notre expérience dans le développement de CASPER (Dubois et al. 01c) et MAGIC (Renevier & Nigay 01) (Nigay et al. 01). Ces deux systèmes sont introduits dans le paragraphe suivant.

5.2.1 Deux systèmes de Réalité Augmentée : CASPER et MAGIC

CASPER (Computer ASSisted PERicardial puncture) est un système de chirurgie assistée par ordinateur ou chirurgie augmentée, développé en collaboration avec le service de chirurgie cardiaque du CHU de Grenoble. Le chirurgien réalise une ponction péricardique au moyen d'une aiguille de ponction. Cette aiguille, ainsi que le patient sont localisés dans l'espace par des caméras. Dans une première version de CASPER-V1, qui a été validée cliniquement, la trajectoire prédéfinie avant l'opération et la position courante de l'aiguille sont affichées sur un écran d'ordinateur en temps réel pendant l'intervention. Dans une deuxième version de CASPER-V2, la trajectoire est affichée en superposition du champ opératoire et donc de l'aiguille de ponction, le chirurgien portant un casque semi-transparent.

Les deux versions de CASPER reposent sur la connaissance de la position en 3D (1) des images échographiques acquises avant l'opération pour calculer la trajectoire, (2) de l'outil au cours de l'intervention et (3) du patient, tous les calculs étant effectués par un ordinateur PC. Le principe de localisation retenu consiste en un système optique. Il se compose de deux caméras qui peuvent localiser en temps réel un nombre important de marqueurs : avec le système Polaris, ces marqueurs sont des diodes émettrices d'infrarouges ou des pastilles réfléchissant les infrarouges. Chaque élément à localiser (sonde échographique, aiguille, etc.) est accolé fermement à un groupe de marqueurs (entre 3 et 6 en général), qui forme un corps-rigide. Dans CASPER-V2, le système doit aussi connaître la position du regard du chirurgien, afin d'afficher la trajectoire dans le casque semi-transparent en superposition du champ opératoire. Pour cela, le centre des deux écrans semi-transparents composant le casque constitue une approximation de la position du regard. Un corps-rigide supplémentaire est conséquemment fixé sur l'armature du casque.

Le domaine d'application du projet MAGIC (Mobile, Augmented reality, Group Interaction, in Context) est l'archéologie. La plate-forme MAGIC est destinée à des fouilles archéologiques d'urgence, qui durent moins de trois mois, suite à la découverte d'un site lors de travaux de construction. Ce projet repose sur une collaboration avec le Centre d'Etudes Alexandrines. L'objectif est d'aider un groupe d'archéologues à effectuer les fouilles le plus rapidement possible en leur fournissant des outils informatiques pour le relevé de terrain mais aussi pour la collaboration à distance sur le site en partageant leurs trouvailles.

La plate-forme matérielle conçue est un assemblage de dispositifs commercialisés. Nous utilisons une tablette Fujitsu Stylistic 3400. Véritable PC (processeur Pentium III 400, mémoire vive portée à 192 Mo, carte vidéo 3D), cette tablette tactile est de la dimension d'une feuille A4 et d'un poids de 1,5 kg. En outre, elle dispose d'une sortie vidéo permettant l'affichage double, à laquelle nous connectons un casque semi-transparent, un SONY LDI D100 BE, d'une résolution de 800x600 pixels. Nous exploitons l'option de Windows 98 qui permet d'obtenir deux "écrans" distincts : l'écran de la tablette et ceux du casque. Une caméra est fixée entre les deux écrans du

casque (entre les deux yeux) : il s'agit d'une caméra USB Philips Ucam Pro d'une résolution 600x480 pixels. La plate-forme matérielle contient aussi un magnétomètre (HMR 3000 de Honeywell) qui détermine l'orientation du regard de l'utilisateur ainsi qu'un GPS qui localise l'utilisateur mobile dans le champ de fouille. Pour le partage des données entre les utilisateurs, nous utilisons un réseau sans fils Wavelan à travers des cartes PCMCIA.

La plate-forme logicielle comporte les fonctions de base, plusieurs fonctions de communication (forum et messages électroniques), de coordination (positionnement de chaque archéologue sur un plan du site) et de production (édition d'objets, base de données).

5.2.2 *Espace taxinomique*

5.2.2.1 *Propriétés intrinsèques à la Réalité Augmentée*

Nous retenons deux principes constituant les fondements de la réalité augmentée. Elle permet notamment de sortir les capacités de l'ordinateur de sa "boîte" habituelle, dont les modalités d'entrée et de sortie reposent essentiellement sur un écran, un clavier et une souris. L'objectif est de parvenir à rompre la frontière entre l'ordinateur et le monde réel, qui empêche l'entrelacement simple et transparent des informations issues du monde réel et celles issues du monde informatique. Un autre principe de base de la RA est qu'elle permet d'ajouter en temps réel des données ou des services informatiques dans l'environnement où évolue l'utilisateur. Il s'agit donc d'accroître la capacité de l'utilisateur à percevoir des informations, ou à exécuter des tâches, sans le couper de son environnement. Partant de ces deux principes, les approches de la littérature expliquent principalement comment le monde réel est augmenté : dans (Dubois et al 01a), nous proposons une revue des espaces taxinomiques existants. Notre espace taxinomique vise à intégrer et organiser les facettes variées des systèmes, que nous qualifions de systèmes mixtes. Une caractéristique fondamentale réside dans l'objet de la tâche réalisée par l'utilisateur. Nous qualifions un système mixte de :

- réalité augmentée (RA), lorsque l'objet de la tâche est une entité de l'environnement physique de l'utilisateur, c'est-à-dire du monde réel,
- virtualité augmentée (VA) lorsque l'objet de la tâche est une entité virtuelle dont la manipulation ou la perception s'appuient sur des objets du monde réel.

Un système mixte permettant la réalisation de plusieurs tâches, il peut certes appartenir aux deux classes RA et VA pour des tâches distinctes. Afin de distinguer les deux classes de systèmes, nous définissons les deux continua de la Figure 7 :

- Un continuum noté **Réalité Augmentée** : il permet de caractériser des systèmes dont l'objet de la tâche est dans le monde réel. Nos deux systèmes CASPER et MAGIC appartiennent à cette classe.
- Un continuum noté **Virtualité Augmentée** : il est dédié aux systèmes dont l'objet de la tâche réside dans le monde informatique. Les systèmes considérés visent à rendre l'interaction plus "réaliste". A la Figure 7, nous avons cité les différentes approches, comme les interfaces tangibles et les interfaces incorporées, que nous qualifions de Virtualité Augmentée. Ces approches sont issues d'une taxonomie présentée dans (Fishkin et al. 98). L'ordre de citation est un ordre chronologique, et il ne signifie pas que nous considérons les interfaces tangibles plus "réelles" que les interfaces gestuelles. Notre système présenté à la Figure 2 (paragraphe

3.1) est un exemple d'interface incorporée donc de virtualité augmentée, l'objet de la tâche étant le courrier électronique. Un autre exemple est notre application de retouche d'images basée sur les pixels miroirs (Vernier et al. 99). Dans cette application, la caméra observe, comme dans le cas du Bureau Digital (Wellner 91) une feuille de papier sur laquelle l'utilisateur dessine avec les instruments de son métier (pinceau, crayon, gomme, etc.). Face à lui, est présentée sur un écran, l'image de sa main et de son crayon superposée à l'image à retoucher, comme le montre la Figure 8. L'apport majeur est la souplesse offerte dans le choix des modalités d'entrée : pour une tâche donnée, l'utilisateur a le choix entre les instruments du métier et les outils informatiques. De plus, ce choix est décidable à tout instant : l'utilisateur peut dessiner avec son feutre favori et changer ensuite la couleur du trait en utilisant les services du logiciel. Dans cette application, bien que l'utilisateur dessine avec un feutre sur une feuille de papier, l'objet de la tâche est bien virtuel, il s'agit de retoucher une image contenue dans un fichier informatique.

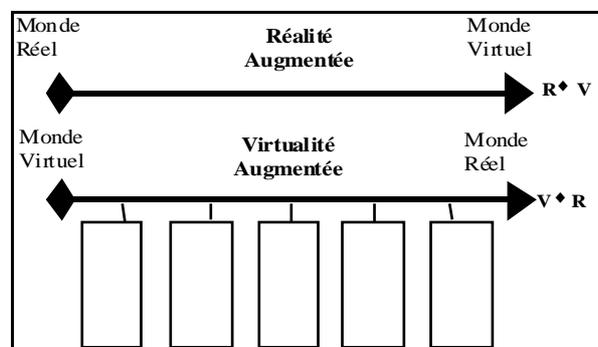


Figure 7 : Deux continua pour distinguer "Réalité Augmentée" et "Virtualité Augmentée". Figure étendue de la Figure 5 dans (Dubois et al. 01a).



Figure 8 : Application de retouche d'images basée sur les pixels miroirs.

Nous n'avons présenté ici qu'un seul trait de la classification, la classification complète étant exposée dans (Dubois & Nigay 00). Une classification n'est utile que si elle permet de comparer les systèmes existants. Dans la thèse de E. Dubois (Dubois 01), quatorze systèmes de la littérature en IHM et une vingtaine de systèmes de Chirurgie Augmentée ont été classés dans notre espace. Outre les caractéristiques intrinsèques d'un système mixte, considérées dans notre espace taxinomique, il convient aussi d'étudier la mobilité de l'utilisateur et le caractère mono-/multi-

utilisateur du système. Pour cela, nous cernons notre étude aux systèmes mixtes que nous qualifions de RA.

5.2.2.2 *Réalité Augmentée mobile et collaborative*

Dans le livre "Mixed Reality: Merging Real and Virtual Worlds" (Azuma 99), Azuma souligne la mobilité inhérente de la réalité augmentée. Concevoir des systèmes de RA fonctionnant en extérieur par exemple est une étape incontournable du développement de la RA puisque l'utilisateur se déplace naturellement dans le monde réel. Ainsi la RA s'est orientée très tôt vers la mobilité et de nombreux systèmes sont déjà mobiles ou le sont dans l'esprit. Ces systèmes s'appuient en général sur un casque semi-transparent, sur la technologie sans fil et sur des créations sur mesure, comme pour le système NETMAN (Bauer et al. 96), où un PC a été intégré à une veste faite sur-mesure. L'objet de la tâche dans NETMAN est le réseau informatique d'une université. Le technicien mobile peut par exemple obtenir des informations superposées à un câble, information indiquant ce qui est connecté au câble. Il est néanmoins important de noter que tous les systèmes de RA ne sont pas mobiles. Aussi la mobilité est un trait de classification des systèmes de RA :

- Les systèmes non mobiles : Ce sont les systèmes de RA où l'objet de la tâche est statique, de taille raisonnable et localisé en un seul lieu. Par exemple les systèmes de chirurgie augmentée comme CASPER sont conçus pour fonctionner dans le bloc opératoire.
- Les systèmes mobiles : Ce sont les systèmes qui visent à augmenter l'environnement dans lequel évolue l'utilisateur. Les objets de la tâche ne sont plus co-localisés. C'est par exemple le cas du système NETMAN que nous venons de citer ou notre plate-forme MAGIC, les objets de la tâche étant alors les trouvailles disséminées sur le site archéologique. Mobilité n'est pas synonyme d'extérieur, citons à titre illustratif le musée augmenté (Rekimoto 95) où les œuvres d'art sont accompagnées d'information textuelle, affichée dans le casque semi-transparent.

Enfin un dernier trait de classification est le caractère mono-/multi-utilisateur du système de RA. Dans (Renevier&Nigay 01), nous faisons la distinction entre un système de RA qui fournit aussi des possibilités de collaboration comme des outils de communication, des systèmes de RA collaborative, où les actions d'un utilisateur influencent l'augmentation de l'environnement d'un autre utilisateur. Pour les systèmes de RA collaborative, nous identifions trois classes selon la situation de collaboration (Renevier&Nigay 01), en considérant d'une part la co-localisation ou non des utilisateurs, d'autre part la localisation de l'objet de la tâche par rapport à celles des utilisateurs.

5.2.3 *Éléments de conception*

5.2.3.1 *Notation OPAS*

Dans (Dubois et al 99), nous présentons une notation pour modéliser un système mixte : ASUR (Adaptor, System, User, Real entity). La notation ASUR repose sur un ensemble de composants matériels et informatiques constitutifs d'un système mixte et considère les relations entre ces composants. Une relation entre deux composants ASUR traduit un échange de données. Parmi les composants ASUR, le composant

"Real entity" est affiné en plusieurs types selon son rôle dans l'interaction : outil ou objet de la tâche. Ces rôles sont repris dans notre espace caractérisant les relations entre un système informatique et le monde réel (Graham et al 00). Le composant ASUR "Adaptor" établit une passerelle entre le monde réel et le monde informatique, comme un localisateur mécanique ou optique (CASPER) ou encore un casque semi-transparent (MAGIC). La modélisation ASUR d'un système interactif, par composants et relations, définit une description de l'interaction induite par le système. Dans cette modélisation, l'interaction utilisateur-système se manifeste sous la forme de relations connectées à l'utilisateur :

Composant-Srce \supset "User" et "User" \supset Composant-Dest

Pour chacune de ces relations faisant intervenir l'utilisateur, il est possible de dresser la liste des modalités utilisées en identifiant les couples (composant-Srce ou composant-Dest, relation), instances du couple (dispositif, langage).

La modélisation ASUR de l'interaction mise en œuvre par un système existant constitue un outil de classification complémentaire à notre espace taxinomique. Typiquement la modélisation ASUR d'un système de Virtualité Augmentée ne contient pas de composant "Real entity" qui est objet de la tâche. De plus ASUR permet d'affiner notre espace taxinomique en distinguant des systèmes appartenant à la même classe dans notre espace. Dans (Dubois 01), plusieurs systèmes mixtes existants sont modélisés selon ASUR afin de montrer l'apport taxinomique de la notation.

Outre cet aspect taxinomique, une évaluation ergonomique prédictive de l'interaction d'un système en cours de conception peut être menée sur la base d'une modélisation ASUR. ASUR devient alors un outil pour la conception. Pour cela nous caractérisons plus finement les composants en relation avec l'utilisateur en définissant l'environnement de perception/d'action associé. Cet environnement est défini par le couple (sens de perception/d'action, localisation physique de perception/d'action). Nous retrouvons ici les caractéristiques du dispositif d'une modalité : capacités sensorielles mises en jeu, et l'extension exposée dans le paragraphe 3.3 liée à l'espace. Nous caractérisons aussi les relations qui partent ou aboutissent à l'utilisateur en considérant la forme des données véhiculées par la relation. Nous retrouvons ici la notion de système représentationnel d'une modalité. Au paragraphe suivant, nous définissons deux propriétés ergonomiques, qui caractérisent la fusion des deux mondes, réel et virtuel. Ces propriétés sont étudiées de façon prédictive au sein d'une modélisation ASUR en considérant les modalités mises en jeu, c'est-à-dire les environnements de perception/d'action et systèmes représentationnels.

5.2.3.2 *Propriétés ergonomiques*

Nous définissons deux propriétés, la compatibilité et la continuité, déclinées selon deux niveaux, le niveau perceptuel et le niveau cognitif (Nigay et al. 01a).

- La compatibilité perceptuelle (extension de la propriété d'observabilité (Gram & Cockton 96) paragraphe 2) traduit la capacité du système à garantir que l'utilisateur peut percevoir les données relatives à N concepts à un instant donné.
- La compatibilité cognitive (extension de la propriété d'honnêteté (Gram & Cockton 96)) traduit la capacité du système à garantir que l'utilisateur peut

interpréter correctement les données perçues, relatives à N concepts à un instant donné.

Similairement, nous définissons la continuité perceptuelle et cognitive pour le cas où un seul concept est représenté physiquement ou informatiquement plusieurs fois. Nous étudions alors la perception et l'interprétation de ces représentations multiples du même concept. Dans le paragraphe 4.2.1.2, nous considérons la représentation multiple d'un espace informationnel.

Dans le système de retouche d'images, nous constatons une compatibilité perceptuelle et cognitive par la mise en œuvre d'un environnement perceptuel unique (vision, écran) et d'un même système représentationnel (le dessin). Au contraire, la phase de calibrage du casque de la plate-forme MAGIC illustre un cas d'incompatibilité perceptuelle, l'utilisateur devant simultanément regarder ce qu'il voit au travers du casque et regarder la tablette. La phase de calibrage est nécessaire pour que la caméra fixée sur le casque (entre les yeux) capture ce que perçoit l'utilisateur à travers le casque. Certes la visibilité de l'utilisateur n'est pas réduite par l'ajout de la caméra. Toutefois le casque focalise la vue à la manière d'une paire de jumelles. Ainsi, la caméra capture plus d'information que le regard de l'utilisateur à travers le casque. Il est donc nécessaire de calibrer la caméra, c'est-à-dire de définir quelle est la partie de l'image prise par la caméra qui correspond à la vue de l'utilisateur. Cette opération se fait lors de la première utilisation de la caméra, par le positionnement d'une lentille magique (Bier et al. 93) dans une fenêtre où le retour caméra est affiché. Le taux de rafraîchissement serait trop lent si ce retour caméra était affiché dans le casque, aussi la fenêtre est affichée sur la tablette. La procédure de calibrage actuelle est donc complexe, l'utilisateur devant regarder alternativement la tablette et à travers le casque. Le problème d'incompatibilité perceptuelle est identifiée et nous sommes en phase de recherche d'une solution.

Enfin, l'étude prédictive de CASPER V-1 a conclu à une discontinuité perceptuelle et cognitive pour le concept d'aiguille de ponction. Deux environnements perceptuels distincts sont mis en jeu, le champ opératoire pour percevoir l'aiguille de ponction et l'écran sur le côté du patient pour percevoir l'aiguille représentée sur l'écran. De plus le système représentationnel 2D à base de croix pour représenter l'aiguille à l'écran rend très difficile la correspondance entre la position courante de l'aiguille réelle perçue en 3D dans l'espace et celle de l'aiguille représentée à l'écran. Les deux systèmes représentationnels sont trop distants créant une discontinuité cognitive. Dans (Dubois et al. 01d), nous décrivons la discontinuité dans CASPER-V1 et le rationnel de conception de CASPER-V2. Une étude empirique basée sur CASPER-V1 et CASPER-V2, démontre l'impact des propriétés de compatibilité et de continuité sur la réalisation de la tâche (Dubois et al. 01b).

Pour étudier la fusion des deux mondes, réel et virtuel, nous considérons les deux propriétés dans le cas d'une représentation physique et d'une représentation virtuelle informatique. Ces deux propriétés concernent alors la perception et l'interprétation d'entités virtuelles et réelles. Pour cela nous avons considéré les relations ASUR qui aboutissent à l'utilisateur. Cet étude doit être prolongée en considérant les relations ASUR partant de l'utilisateur. En particulier la continuité dans la forme des actions (systèmes représentationnels) à effectuer sur la même entité réelle et virtuelle constitue un point intéressant à étudier en vue de la fluidité entre les deux mondes.

Une méthode de conception itérative reposant sur l'analyse prédictive des deux propriétés de compatibilité et de continuité au sein d'une modélisation ASUR est expliquée dans (Dubois et al 01a).

Plus généralement, par rapport à notre espace de composition des modalités (paragraphe 3.2), ces deux propriétés focalisent sur les aspects articulatoire (sens de perception), spatial (lieu de perception) et syntaxique (système représentationnel) de la composition de deux modalités de sortie. Dans le cas de la compatibilité, l'aspect temporel est simultanément et l'aspect sémantique est complémentaire. Dans le cas de la continuité, l'aspect temporel est indéfini et l'aspect sémantique est partiellement ou totalement redondant. Ces deux propriétés sont donc générales à la multimodalité. En effet, dans le cas de la composition de modalités graphiques (paragraphe 4), nous avons présenté deux règles de conception (Nigay & Vernier 98a) détaillées dans (Vernier 01 page 122) qui sont liées à la continuité visuelle : la continuité spatiale et la continuité temporelle.

- La continuité spatiale (observabilité du focus dans son contexte) caractérise la composition de deux modalités dont l'aspect temporel est simultanément, l'aspect sémantique est complémentaire. Ce cas se rattache donc à la compatibilité. Un seul environnement de perception est ici mis en jeu : (vision, écran). La continuité spatiale est donc une instantiation de la compatibilité cognitive pour le cas de la composition de deux modalités graphiques ayant un même environnement de perception (vision, écran).
- La continuité temporelle caractérise la composition de deux modalités dont l'aspect temporel est indéfini, l'aspect sémantique étant partiellement ou totalement redondant. Ce cas se rattache donc à la continuité. Il s'agit d'une instantiation de la continuité cognitive, puisque là encore un seul environnement perceptif est défini (vision, écran). Lorsque cette continuité cognitive ne peut être vérifiée, une règle consiste à aider l'utilisateur à percevoir la transition entre les deux systèmes représentationnels en montrant des étapes intermédiaires de façon animée (Mackinlay et al. 91). Cette solution permet de réduire la charge cognitive en la déléguant au niveau perceptuel.

La localisation des deux propriétés dans notre espace de composition de modalités, nous permet d'entrevoir une approche systématique pour étudier quand les propriétés sont vérifiées. Il convient d'étudier les trois aspects articulatoire, spatial et syntaxique de la composition au regard des différents schémas. Les schémas définissent un cadre systématique pour étudier ces propriétés selon des théories de psychologie comme ICS (Barnard & May 93).

5.2.4 Deux techniques génériques de RA

Au sein de la plate-forme MAGIC de RA mobile et collaborative, nous avons conçu deux techniques génériques qui participent à la fluidité entre les deux mondes, réel et virtuel. Ces deux techniques sont nommées Réalité Cliquable et Ballade Augmentée.

5.2.4.1 Réalité Cliquable

La réalité cliquable consiste en la capture d'images du monde réel en cliquant avec le stylet. Pour ce faire, la caméra fixée sur le casque (entre les yeux) doit capturer ce que perçoit l'utilisateur à travers le casque. Une phase de calibrage difficile à réaliser est nécessaire, comme nous l'avons expliquée au paragraphe précédent. Une fois

cette opération de calibrage réalisée, la capture d'images est par contre aisée. Elle se fait par le positionnement d'une lentille magique superposée à la zone du monde réel à "sauvegarder". La visée se fait à travers le casque, mais la manipulation de la lentille se fait sur la tablette, dans la fenêtre "gateway". Une fois la lentille positionnée et dimensionnée, l'utilisateur appuie sur le bouton "take" (prendre une photo). Il est alors possible d'enregistrer cette image sur le disque dur, et de l'intégrer au formulaire de description de l'objet dans la base de données.

Lors de la sélection d'une zone du monde réel, nous constatons une compatibilité perceptuelle et cognitive par la mise en œuvre d'un environnement perceptuel unique (vision, casque) et de systèmes représentationnels conciliables (la réalité et la lentille magique superposée). De plus la réalité cliquable vérifie la continuité dans la forme des actions (systèmes représentationnels) effectuées par l'utilisateur. En effet dans le transfert d'un objet d'un monde à l'autre, que l'objet manipulé soit matériel ou virtuel, les actions s'effectuent de la même manière grâce à la fenêtre "gateway". La fenêtre "gateway" " définit une passerelle entre le monde virtuel (informatique) et le monde réel. La fenêtre affichée sur la tablette est une copie mise à l'échelle de la fenêtre dans le casque. L'utilisateur dispose alors d'un lieu d'interaction sur la tablette pour manipuler avec le stylet, ce qui est présent dans le casque. Ainsi le stylet peut servir à désigner des objets matériels perçus au travers du casque. La fenêtre "gateway" évite donc l'introduction d'une modalité supplémentaire qui serait dédiée à la désignation des objets matériels. Tous les événements souris produits dans la "gateway" sont dupliqués dans le casque. En particulier, les déplacements de la souris dans la fenêtre "gateway" sont reproduits dans le casque. Pour illustrer l'utilisation de la "gateway" du monde virtuel vers le monde réel, le cas inverse de la réalité cliquable, considérons le cas d'un utilisateur qui souhaite comparer une image de la base de données avec un objet réel. Afin de faciliter la comparaison, un glisser-déposer, avec le stylet, de l'image dans la "gateway" va permettre à l'utilisateur de voir l'image affichée dans le casque, avec l'objet matériel à côté perçu au travers du casque semi-transparent.

La réalité cliquable est à opposer à notre système de retouche d'images (Figure 8). Tandis que lors de la retouche d'une image, l'utilisateur effectue des actions dans le monde réel (papier et crayon) sur un objet de la tâche virtuel (images à l'écran), l'utilisateur de la plate-forme MAGIC effectue des opérations dans le monde informatique sur la tablette pour manipuler un objet réel. Ces deux exemples soulignent la différence entre Réalité Augmentée et Virtualité Augmentée (Figure 7).

5.2.4.2 *Ballade Augmentée*

La ballade augmentée consiste à restituer un objet matériel dans son contexte d'origine, l'objet ayant été déplacé. La technique consiste à superposer une image de l'objet à son contexte d'origine (dans le monde réel), grâce au casque semi-transparent. La ballade augmentée est une technique de RA mobile et collaborative. D'une part, il s'agit de RA mobile car l'augmentation du monde réel est basée sur la localisation de l'utilisateur. Dans la plate-forme MAGIC, la localisation de l'utilisateur est connue grâce à un GPS, l'orientation de son regard par un magnétomètre. Lors de la capture d'un objet matériel (réalité cliquable), l'objet est stocké avec sa position. Ainsi un objet déplacé peut être restitué dans son contexte

d'origine. D'autre part, il s'agit d'une collaboration asynchrone : un utilisateur capture d'abord un objet dans son contexte pour ensuite le déplacer ; un autre utilisateur peut ultérieurement percevoir l'objet déplacé dans son contexte d'origine.

5.2.5 Architecture logicielle

Reposant sur le modèle PAC-Amodeus (Nigay 94), nous proposons deux solutions générales à la conception logicielle de systèmes mixtes. La première concerne la réalisation logicielle de la capture d'information du monde réel et la deuxième est dédiée à la réalisation de passerelles entre les deux mondes.

5.2.5.1 Capture d'information du monde réel

Nous considérons ici la réalisation logicielle de la capture d'information en provenance du monde réel, comme la position de l'utilisateur, l'orientation de son regard ou encore la position d'un outil matériel. Nous nous intéressons donc à la réalisation logicielle de techniques dites "passives" car elles ne requièrent pas de commandes explicites de la part de l'utilisateur ("perceptual user interfaces" (Turk & Robertson 00)). Comme pour une technique "active" (modalité d'interaction), nous localisons le processus d'acquisition et d'interprétation dans les deux composants CBNI (Composant Bas Niveau d'Interaction) et CTP (Composant Technique de Présentation) du modèle PAC-Amodeus. En appliquant le mécanisme de branches hérité du modèle ARCH (UIMS 92), un CBNI et un CTP sont associés à une technique. En cela, nous sommes en accord avec l'approche boîte à outils pour le contexte proposée dans (Salber et al 99), la fusion de plusieurs sources d'information se faisant dans le CIBN et/ou le CTP. L'aspect nouveau des techniques passives par rapport aux techniques actives est le fait que certaines informations ne seront pas transmises au Contrôleur de Dialogue. Nous modélisons alors un flot d'information horizontal d'un CTP à un autre et la boucle d'interaction suivante :

Utilisateur \rightarrow CIBN1 \rightarrow CTP1 \rightarrow CTP2 \rightarrow CIBN2 \rightarrow Utilisateur.

Cette boucle d'interaction est présente dans le logiciel de CASPER, lorsque le chirurgien bouge la tête. L'information affichée dans le casque est conséquemment modifiée sans faire intervenir le Contrôleur de Dialogue. Le dialogue n'est pas modifié par la position du regard du chirurgien. Au contraire dans le cas de MAGIC, l'orientation de la tête de l'utilisateur, capturée au niveau du CIBN grâce au magnétomètre fixé sur le casque, est transmise au Contrôleur de Dialogue. En effet, l'orientation couplée à la localisation de l'utilisateur est pertinente au Contrôleur de Dialogue pour déterminer si des objets sont visibles ou non lors de la ballade augmentée.

Notons que ce flot d'information horizontal entre PTCs et cette boucle d'interaction ont déjà été identifiés dans le cas de techniques "actives" (Nigay 94) pour réaliser des retours d'information qualifiés de niveau intermédiaire. Par exemple, l'affichage de la phrase reconnue par un système de reconnaissance de la parole localisé dans un CTP ne fait pas intervenir le Contrôleur de Dialogue. La différence de réalisation entre techniques actives et celles passives réside donc dans le fait que certaines informations issues de techniques passives ne sont pas exploitées par le Contrôleur de Dialogue.

5.2.5.2 *Passerelle entre les deux mondes*

La réalisation d'une passerelle entre les mondes, réel et virtuel, implique son incarnation dans les deux mondes. Un exemple de passerelle est la fenêtre "gateway" affichée sur l'écran de la tablette et dans le casque de la plate-forme MAGIC. Pour la réalisation logicielle, nous appliquons notre motif architectural "Vue multiple" (paragraphe 2). Au sein du Contrôleur de Dialogue, un agent est dédié à chaque représentation, l'une dans le monde réel, l'autre dans le monde virtuel. Ces agents sont indépendants des modalités d'interaction (Nigay & Coutaz 95), mais correspondent à des lieux d'interaction distincts. Un agent ciment, père de ces agents, gère la fluidité entre les deux mondes. Cette solution a été appliquée pour la réalisation de la fenêtre "gateway" de la plate-forme MAGIC, l'architecture logicielle PAC-Amodeus étant détaillée dans (Renevier & Nigay 01). Le même motif architectural aurait pu être appliqué à la réalisation logicielle du système de retouche d'images. Un agent aurait été dédié au lieu d'interaction défini par la feuille de papier, un autre à la fenêtre à l'écran où est affiché l'image.

5.3 Perspectives de recherche

Nos perspectives de recherche liées à la réalité augmentée s'organisent selon deux axes complémentaires : le premier axe de recherche concerne la conception et réalisation de nouvelles techniques pour combiner les deux mondes, réel et virtuel tandis que le deuxième concerne les aspects méthodologiques de conception dans le but de concevoir des techniques utiles et utilisables.

5.3.1 *Conception et réalisation de techniques de RA et de liens dynamiques entre les deux mondes*

Pour montrer la généralité de la plate-forme MAGIC, nous envisageons de développer un jeu collaboratif dans le campus en utilisant la plate-forme qui est matérielle et logicielle. Dans (Renevier & Nigay 01), nous avons montré pourquoi les fouilles archéologiques constituaient un domaine d'application adéquat à la RA mobile et collaborative. Nous pensons qu'un jeu collaboratif comme une chasse au trésor ou un jeu de rôles constitue un exemple tout aussi adéquat. A la différence des jeux sur ordinateur où un monde est entièrement reconstitué, le jeu envisagé repose sur le monde réel qui est augmenté, par exemple par des zones magiques. Un jeu de ce type est décrit dans (Starner et al. 00) mais les utilisateurs ne sont pas mobiles. Outre le fait de montrer la généralité de la plate-forme, ce jeu sera aussi un terrain d'investigation pour étudier des augmentations autres que visuelles. En particulier nous envisageons l'usage de sons spatialisés qui pourraient par exemple indiquer la position approximative des autres joueurs.

Au-delà de la conception de techniques, l'établissement de liens entre les deux mondes, réel et virtuel, constitue un de nos sujets d'étude. Tandis que dans CASPER, l'augmentation (par l'affichage de la trajectoire) est figée par le système et donc décidée par le concepteur, nous constatons que dans MAGIC-archéologie, l'augmentation (lors de la ballade augmentée) est dynamique, définie par un utilisateur. Le lien entre l'image, objet virtuel, et sa localisation, point d'ancrage dans le monde réel est néanmoins spécifié implicitement : l'utilisateur n'a pas à explicitement spécifier qu'à cet endroit du terrain de fouille, une information supplémentaire est accessible. Le lien entre les deux mondes est donc dynamique

mais spécifié implicitement. Nous souhaitons étudier des liaisons dynamiques explicites, où l'utilisateur par une commande explicite, définit le lien entre d'une part un lieu ou un objet matériel et d'autre part un objet virtuel. Par exemple, dans le cadre de notre jeu collaboratif, nous pouvons envisager un scénario comme un utilisateur spécifiant que dans cette boîte est cachée une arme ou que le franchissement de cette porte donne un pouvoir magique d'invisibilité sur la carte affichée sur la tablette. L'établissement de liens entre les deux mondes, effectué explicitement par l'utilisateur est évoqué dans (Mackay 00b) qui souligne le risque potentiel de laisser la décision de ces liens aux concepteurs : "another strategy is to explicitly give the control to the users, allowing them to define and more importantly, continue to evolve, the relationship between the physical and virtual documents". Pour spécifier ces liens entre les deux mondes, la multimodalité, comme la parole et un geste de désignation, est une solution qui semble prometteuse (McGee et al. 00) comme la phrase "cette porte donne le pouvoir d'invisibilité" accompagnée d'un geste de désignation de la porte.

5.3.2 Méthode de conception : des systèmes de RA utiles et utilisables

Certes le RA offre de nombreuses possibilités enthousiasmantes et vise à rendre le monde réel "magique" ("Magic by Metaphors" (Madsen 00)). Ceci en fait un axe de recherche très attrayant aux démonstrateurs époustouflants. Mais très peu de systèmes de RA ne sont effectivement utilisés. Citons néanmoins des systèmes comme CASPER-V1 de Chirurgie Augmentée qui sont cliniquement validés. Nous pouvons avancer comme explication à l'état de démonstrateurs des systèmes de RA les limites actuelles de la technologie. Par exemple dans CASPER-V2, la précision de l'affichage de la trajectoire est une barrière à son usage clinique. De même, le déploiement du système MAGIC-Archéologie à Alexandrie soulève des problèmes très pratiques et concrets comme l'alimentation en énergie ou la luminosité ambiante qui empêche de voir l'écran de la tablette. Il convient de constater que cet axe de recherche est cependant fortement aiguillonné par les technologies. Répondant à cet approche centrée technologies, W. MacKay (MacKay 00) propose d'étudier les différences entre les deux mondes, réel et virtuel, et rappelle que la RA n'est pas la panacée à tous les problèmes difficiles d'interaction. Dans (Dubois et al. 99), nous avons aussi constaté cette approche guidée par les technologies en étudiant les taxonomies existantes de Chirurgie Augmentée. Ce constat nous a incités à adopter une approche centrée sur l'utilisateur et à étudier les étapes de conception afin de concevoir des systèmes adéquats à la pratique de l'utilisateur. Dans les paragraphes précédents, nous avons avancé des éléments de conception centrée sur l'utilisateur. Nous allons prolonger cet effort, en particulier dans le cadre du projet RNTL MMM (Médecin, Malade, Machine) qui a démarré en octobre 2001. Pour cela nous avons distingué deux parties au sein du projet : le Protocole Chirurgical Informatisé (PCI) décrira les besoins interactionnels d'une manière indépendante des paradigmes technologiques logiciels et matériels actuels. Il constituera le socle conceptuel d'un Ancillaire Chirurgical Informatisé (ACI), "boîte à outils" matériels et logiciels améliorant l'interaction homme-machine, en innovant dans la combinaison de capteurs de données multimédia et d'effecteurs spécifiquement conçus. La Chirurgie Augmentée est un domaine exigeant, nécessitant la prise en compte fine de ses particularités. De plus les objectifs sont définis précisément pour chaque spécialité chirurgicale. Dans les domaines d'application aux objectifs et contraintes moins

cernés, nous avons commencé une étude complémentaire sur les processus de conception basée sur les scénarios. Les scénarios constituent à la fois un support de communication permettant le partage d'information entre les participants à la conception, mais également à l'occasion de leurs constructions un outil favorisant la communication entre les participants. Ils sont flexibles et permettent de rendre des activités de conception plus accessibles aux différents acteurs et en particulier les utilisateurs finaux (Carroll 00). Dans la conception de MAGIC-Archéologie, nous avons suivi une méthode fondée sur des scénarios et enrichie des apports méthodologiques de l'ergonomie avec une analyse de la tâche et de l'activité des archéologues. En particulier les scénarios développés servent de support à l'analyse de la tâche et à l'analyse de l'activité présente et future (activité avec les techniques de RA à concevoir) (Nigay et al 01b). L'expérience de ce projet a aussi montré l'intérêt et l'efficacité des scénarios comme langage commun lors de la conception au sein d'une équipe pluridisciplinaire composée de trois psycho-ergonomes et deux informaticiens.

6 Conclusion

Les travaux exposés dans ce mémoire sont résolument ancrés dans l'ingénierie de l'interaction homme-machine. Nous avons pour chaque axe de recherche souligné l'approche pluridisciplinaire et les points d'ancrage entre les sciences cognitives et le génie logiciel (chapitre 2). Tout au long de l'exposé, les deux facettes de nos travaux, conception ergonomique et conception logicielle, sont apparentes dans nos contributions mais également dans nos perspectives. Nos travaux s'articulent autour des termes de modalité d'interaction et de multimodalité. Nous nous sommes appuyés sur des résultats généraux incluant la définition d'une modalité et d'un espace de composition de modalités (chapitre 3), et nous les avons infléchis à deux axes de recherche, la visualisation et la réalité augmentée (chapitre 1). Le choix de ces deux axes a été motivé par les constats suivants :

- L'évolution exponentielle de la quantité d'information disponible, due à l'explosion des technologies communicantes, rend la maîtrise des informations par les utilisateurs de plus en plus difficile. L'explosion du World Wide Web illustre ce phénomène. Dans ce contexte, le problème n'est plus tant l'accès à l'information mais la capacité de sélection, à partir des représentations perçues, d'une information qui fasse sens et qui réponde au besoin.
- Les progrès technologiques élargissent l'espace d'interaction : ce dernier inclut l'environnement physique et n'est plus confiné uniquement à un ordinateur sur un bureau.

Le premier constat nous a incités à étudier la visualisation d'information (chapitre 4) que nous avons abordé en jetant un regard nouveau, celui des modalités graphiques de sortie et de la multimodalité en sortie. Nos contributions à cet axe concernent l'ensemble du processus de développement, de la conception à la réalisation logicielle, sont à la fois générales et dépendantes du type de données visualisées, les données temporelles, et concernent des situations d'interaction mono-utilisateur et multi-utilisateur.

Le deuxième constat a suscité notre étude de la Réalité Augmentée visant la fusion harmonieuse des mondes réels et virtuels (Chapitre 5). L'environnement physique dans l'interaction ouvre de nombreuses perspectives que nous avons étudiées selon le point de vue des modalités d'interaction et de la multimodalité. Nos travaux reposent sur la distinction entre augmenter l'interaction avec le monde réel grâce aux capacités de traitements informatiques et augmenter l'interaction avec un ordinateur grâce à des objets matériels. Comme pour la visualisation d'information, nos résultats sont dédiés aux différentes phases du processus de développement, concernent des situations d'interaction mono-utilisateur et multi-utilisateur.

Soulignons enfin que nous n'avons pas explicitement fait apparaître comme dimension à nos travaux le caractère mobile de l'utilisateur. Or, nous l'avons simplement évoqué au paragraphe 3, la conception de modalités et de combinaisons de modalités adaptées à la situation de mobilité se justifie doublement :

1. La situation de mobilité de l'utilisateur impose de concevoir des modalités d'interaction et des combinaisons de modalités qui sortent du cadre classique d'un grand écran, d'un clavier et d'une souris. Les ordinateurs de poche de même que les téléphones mobiles actuels reposent sur des dispositifs physiques d'entrée/sortie aux capacités restreintes, comme la taille de l'écran. Il convient donc de concevoir de nouvelles formes d'interaction : par exemple (Walker et al. 01) proposent l'usage de sons spatialisés pour parcourir les rendez-vous de la journée en situation mobile.
2. La situation de mobilité de l'utilisateur implique des contextes d'interaction par définition variables (dans le train, dans la rue, etc.). Il convient alors de laisser le choix à l'utilisateur de l'usage des modalités selon le contexte d'utilisation. Par exemple le mode vibratoire qui existe sur la plupart des téléphones mobiles actuels constitue une modalité de sortie que l'utilisateur peut choisir à la place de la modalité sonore. Cet aspect d'adaptation est particulièrement important : la liberté laissée à l'utilisateur de choisir une modalité ou une combinaison de modalités la plus efficace et la plus adaptée à un contexte d'utilisation évolutif (propriété inhérente aux situations de mobilité) constitue une condition déterminante de l'acceptation et de l'utilisation effective de l'artéfact. Nous avons mené une première étude empirique de l'usage de la multimodalité avec un ordinateur de poche. Quatre modalités fonctionnellement équivalentes ont été retenues pour la manipulation de courriers électroniques : tactile (stylet), verbale, gestuelle et "interface incorporée" présentée à la Figure 3. Les premiers résultats ont montré une grande variabilité inter-individuelle et ont confirmé l'importance de laisser le choix, puisque toutes les modalités ont été utilisées. Pour réaliser cette expérimentation, nous avons développé une plate-forme magicien d'oz (Calvet et al 01). Tandis que l'approche magicien d'oz est très séduisante (Aublet et al. 96) pour tester des modalités difficiles à réaliser et coûteuses en temps de développement comme les "interfaces incorporées", elle soulève néanmoins le problème d'expérimentations en laboratoire qui va à l'encontre d'une situation de mobilité.

Nos travaux s'inscrivent, selon une perspective générale, dans la tendance actuelle à "l'accès et la manipulation de l'information par tous, n'importe où, n'importe quand, pour des besoins personnalisés", grâce à la conception et à la réalisation de modalités

d'interaction (simples ou composées). L'information est soit ancrée dans le monde réel, comme une concentration de tessons dans un champ de fouille archéologique, soit numérique. Nous rejoignons ici le point de vue exposé dans (Chalmers 00), qui tend à ne pas considérer deux mondes, mais un seul où les types d'information cohabitent : l'information numérique n'est pas plus virtuelle que celle contenue dans un livre.

7 Bibliographie

- (Alberganti 98) Alberganti, M., "La boîte grise est condamnée à disparaître", interview de Jacques Clay, directeur général de la division ordinateurs de bureau de Hewlett Packard, *Le Monde*, 30 avril 1998.
- (Allen 93) Allen, J., Maintaining Knowledge about Temporal Intervals. *Communications of the ACM*, Vol. 26, N. 11, 1983, p. 832-843
- (André et al. 93) André, E., Finkler, W., Graf, W., Rist, T., Schauder, A., Wahlster, W. WIP: The Automatic Synthesis of Multimodal Presentations, Workshop ERCIM on Multimodal Human-Computer Interaction, Working Material, Nov. 1993.
- (Aublet et al. 96) Aublet-Cuvelier, L., Carraux, E., Coutaz, J., Nigay, L., Portolan, N., Salber, D., Zanello, M.-L., NEIMO, un laboratoire d'utilisabilité numérique : Leçons de l'expérience. Actes de la conférence ERGO-IA'96, AFCET SELF IDLS Publ., 1996, p. 149-160.
- (Azuma 93) Azuma, R., Tracking Requirements for Augmented Reality. *Communication of the ACM*, Vol. 36, N. 7, 1993, p. 50-51.
- (Azuma 99) Azuma, R. Mixed Reality: Merging Real and Virtual Worlds. Yuichi Ohta & Hideyuki Tamura (ed.), Springer-Verlag Publ., 1999, Chp. 21, p. 379-390.
- (Baker et al. 01) Baker, K., Greenberg, S., Gutwin, C., Heuristic Evaluation of Groupware Based on the Mechanics of Collaboration. *Engineering for Human-Computer Interaction*, Little R. & Nigay L. Eds, Springer-Verlag Publ., LNCS 2254, p. 131-148.
- (Barnard & May 93) Barnard, P., May, J., Cognitive modelling for user requirements. *Computers, Communication and Usability: Design issues, research and methods for integrated service*, Byerley, P., Barnard, P., May, J. Eds., Amsterdam: Elsevier, North Holland Series in Telecommunication, 1993, p.101-145.
- (Bass & John 01a) Bass, L., John, B., Evaluating software architectures for usability. *Engineering for Human-Computer Interaction*, Little R. & Nigay L. Eds, Springer-Verlag Publ., LNCS 2254, p. 39-40.
- (Bass & John 01b) Bass, L., John, B., Achieving Usability Through Software Architecture. *Proceedings of ISCE 2001*, IEEE Computer Society Publ., p. 684. CMU/SEI-2001-TR-005.
- (Bauer et al. 96) Bauer, Heibe, Kortuem, Segall A Collaborative Wearable System with Remote Sensing. *Proceedings of the Second International Symposium on Wearable Computers, ISWC'98*, 1998, p. 10-17.

- (Bayle 97) Bayle, E., Bellamy, R., Casaday, G., Erickson, T., Fincher, S., Grinter, B., Gross, B., Lehder, D., Marmolin, H., Potts, C., Skousen, G., Thomas, J., Putting It All Together: Towards a Pattern Language for Interaction Design. CHI'97 Workshop, SIGCHI Bulletin, 1998.
http://www.pliant.org/personal/Tom_Erickson/Patterns.WrkShpRep.html
- (Bellotti et al. 96) Bellotti, V., Blandford, A., Duke, D., MacLean, A., May, J., Nigay, L., Interpersonal Access Control in Computer Mediated Communications: A Systematic Analysis of the Design Space. Human Computer Interaction, Lawrence Erlbaum Publ., Vol. 11, N. 4, 1996, p. 357-432.
- (Bérard 99) Bérard, F., The Perceptual Window: Head Motion as a new Input Stream. Proceedings of INTERACT'99, Sasse A. & Johnson C. Eds, IFIP IOS Press Publ., 1999, p. 238-244.
- (Bertin 81) Bertin, J., Graphics and Graphic Information Processing. Berlin, New York: Walter de Gruyter Publ., ISBN 3-11-006901-6, 1981, 273 pages.
- (Bier et al. 93) Bier, E., Stone, M., Pier, K., Buxton, W., DeRose, T., Toolglass and Magic Lenses: The See-Through Interface. Proceedings of SIGGRAPH'93, ACM Press: New York Publ., 1993, p. 73-80.
- (Blandford & Stelmaszewska 01) Blandford, A., Stelmaszewska, H., Shooting the information rapids. Proceedings of IHM-HCI2001, Vanderdonck J., Blandford A. Derycke A. Eds., CEPADUES Publ., 2001, p. 51-54.
- (Bourguet & Caelen 92) Bourguet, M-L., Caelen, J., Interfaces Homme-Machine Multimodales: Gestion des Evénements et Représentation des Informations. Actes de la conférence ERGO-IA'92, AFCET SELF IDLS Publ., 1992, p. 124-134.
- (Calvary et al. 97) Calvary, G., Coutaz, J., Nigay, L., From Single-User Architectural Design to PAC*: a Generic Software Architecture Model for CSCW. Proceedings of CHI'97, Pemberton S. Ed., ACM Press: New York Publ., 1997, p. 242-249.
- (Calvet et al. 01) Calvet, G., Kahn, J., Zouinar, M., Salembier, P., Briois, J.-C., Nigay, L., Rey, G., Pasqualetti, L., Etude empirique de l'usage de la multimodalité avec un ordinateur de poche. Proceedings of IHM-HCI2001, Vanderdonck J., Blandford A. Derycke A. Eds., CEPADUES Publ., 2001, p. 5-8.
- (Card et al. 99) Card, S., Mackinlay, J., Shneiderman, B., Readings in Information Visualization: Using Vision to Think. The Morgan Kaufmann Series in Interactive Technologies, Card, S., Grudin, J., Nielsen, J., Skelly, T., Series Editors, ISBN 1-55860-533-9, 1999, 712 pages.
- (Carlis & Konston 98) Carlis J., Konston, J. Interactive visualization of serial periodic data. Proceedings of UIST'98, , ACM Press: New York Publ., 1998, p. 29-38.
- (Carroll 00) Carroll, J.M., Making use. Scenario-based design of computer interactions. MIT Press, Cambridge, Massachusetts, 2000.
- (Chalmers 00) Chalmers, M., Interweaving Objects, Types and People. Proceedings of DARE 2000, Designing Augmented Reality Environments, MacKay, W. Ed., ACM, 2000, p. 159-160.
- (Chatty 94) Chatty, S., Extending a Graphical Toolkit for Two-Handed Interaction. Proceedings of UIST'94, , ACM Press: New York Publ., 1994, p. 195 - 204.

- (Coutaz et al. 95) Coutaz, J., Nigay, L., Salber, D., Blandford, A., May, J., Young, R., Four Easy Pieces for Assessing the Usability of Multimodal Interaction: The CARE properties. Proceedings of INTERACT'95, Arnesen S.A. & Gilmore D. Eds., Chapman&Hall Publ., 1995, p. 115-120.
- (Coutaz et al. 96) Coutaz, J., Nigay, L., Salber, D., The MSM framework: A Design Space for Multi-Sensori-Motor Systems. Proceedings of EWHCI'93, Bass, L., Gornostaev, J., Unger C. Eds. Springer-Verlag Publ., LNCS 753, 1993, p. 231-241.
- (Coutaz & Nigay 01) Coutaz, J., Nigay, L., Architecture logicielle conceptuelle des systèmes interactifs. Analyse et Conception de l'Interaction Homme-Machine, Kolski C. Ed., Hermes Publ., ISBN 2-7462-0239-5, 2001, p. 207-246.
- (Daassi et al. 01) Daassi, C., Dumas, M., Fauvet, M-C., Nigay, L., Scholl, P-C., Visual Exploration of Temporal Object Databases. Proceedings of BDA'00, p. 159-178.
- (Dieberger 97) Dieberger, A., Supporting social navigation on the World Wide Web. International Journal Human-Computer Studies, Academic Press Publ., Vol. 46, N. 6, p. 805-825.
- (Dourish & Chalmers 94) Dourish, P, Chalmers M., Running out of space: models of information navigation. Adjunct proceedings of HCI '94 (Short Papers), 1994.
- (Dubois et al. 99) Dubois, E., Nigay, L., Troccaz, J., Chavanon, O., Carrat, L., Design Space for Augmented Surgery, an Augmented Reality Case Study. Proceedings of INTERACT'99, Sasse A. & Johnson C. Eds, IFIP IOS Press Publ., 1999, p. 353-359.
- (Dubois & Nigay 00) Dubois, E., Nigay, L., Augmented Reality: Which Augmentation for Which Reality? Proceedings of DARE 2000, Designing Augmented Reality Environments, MacKay, W. Ed., ACM, 2000, p. 165-166.
- (Dubois 01) Dubois, E., Chirurgie Augmentée, Un Cas de Réalité Augmentée ; Conception et Réalisation Centrées sur l'Utilisateur. Thèse de l'Université Joseph Fourier, 2001, 275 pages.
- (Dubois et al. 01a) Dubois, E., Nigay, L., Troccaz, J., Un Regard Unificateur sur la RA : Eléments de Classification et de Conception. RIHM (Revue Interaction Homme-Machine), Europa Productions Publ., 2001, à paraître.
- (Dubois et al. 01b) Dubois, E., Nigay, L., Ribeiro, D., Troccaz, J., Etude empirique de la Continuité dans la Réalité Augmentée. Proceedings of IHM-HCI 2001, Vanderdonck J., Blandford A. Derycke A. Eds., CEPADUES Publ., 2001, p. 9-12.
- (Dubois et al. 01c) Dubois, E., Nigay, L., Troccaz, J., Carrat, L., Chavanon, O., A methodological tool for computer-assisted surgery interface design: its application to computer-assisted pericardial puncture. Proceeding of the 9th annual Medicine Meets Virtual Reality Conference MMVR'2001, J.D. Westwood Ed., IOS Press, 2001, p. 136-139.
- (Dubois et al. 01d) Dubois, E., Nigay, L., Troccaz, J., Consistency in Augmented Reality Systems. Engineering for Human-Computer Interaction, Little R. & Nigay L. Eds, Springer-Verlag Publ., LNCS 2254, p. 117-130.
- (Dumas et al. 00) Dumas, M., Daassi, C., Fauvet, M-C., Nigay L., Pointwise temporal object database browsing. Proceedings of the ECOOP Symposium on Objects and Databases, 2000, 15 pages.

- (Feiner & McKeown 91) Feiner, S., McKeown, K., COMET: generating coordinated multimedia explanations. Proceedings of CHI'91, Robertson S., Olson G., Olson J. Eds., ACM Press: New York Publ., 1991, p. 449-450.
- (Fishkin et al. 98) Fishkin, K., Moran, T., Harrison, B., Embodied User Interfaces: Towards Invisible User Interfaces. Proceedings of the IFIP TC2/WG2.7 Working Conference on Engineering for Human-Computer Interaction, Chatty S. & Prasun D. Eds., Kluwer Academic Publ., 1998, p. 1-18.
- (Gershon et al. 97) Gershon, N., Card, S., Eick, S. Information Visualization. CHI'97 Tutorial notes, 1997.
- (Gershon & Page 01) Gershon, N., Page, W., What Storytelling Can Do for Information Visualization. Communications of the ACM, Vol. 44, N. 8, 2001, p. 31-37.
- (Gourdol et al. 92) Gourdol, A., Nigay, L., Salber, D., Coutaz, J., Two case Studies of Software Architecture for Multimodal Interactive Systems: VoicePaint and a Voice-enabled Graphical Notebook. Proceedings of the IFIP TC2/WG2.7 Working Conference on Engineering for Human-Computer Interaction, Larson J. & Unger C. Eds, IFIP Transactions Series Amsterdam, The Netherlands: North-Holland Publ., 1992, p. 271-283.
- (Graham et al. 00) Graham, N., Watts, L., Calvary, G., Coutaz, J., Dubois, E., Nigay, L. A Dimension Space for the Design of Interactive Systems within their Physical Environments. Proceedings of DIS'2000, ACM Press: New York Publ., p. 406-416.
- (Gram & Cockton 96) Gram, C., Cockton, G., Eds., Design Principles for Interactive Software. Chapman et Hall Publ., 1996, 248 pages.
- (Gray & Salber 01) Gray, P., Salber, Modelling and Using Sensed Context Information in the Design of Interactive Applications. Engineering for Human-Computer Interaction, Little R. & Nigay L. Eds, Springer-Verlag Publ., LNCS 2254, p. 335-354.
- (Harrison et al. 98) Harrison, B., Fiskin, K., Gujar, A., Mochon, C., Want, R., Squeeze me, Hold me, Tilt Me ! An exploration of Manipulative User Interface. Proceedings of CHI'98, Karat C-M. & Lund A. Eds., ACM Press: New York Publ., 1998, p. 17-24..
- (Hascoët 98) Hascoët, M., Analytical Versus Empirical Evaluation of Spatial Displays. Proceedings of CHI'98 Extended Abstracts, Karat C-M. & Lund A. Eds., ACM Press: New York Publ., 1998, p. 257-259.
- (Hascoët & Beaudouin-Lafon 01) Hascoët, M., Beaudouin-Lafon, M., Visualisation Interactive d'Information, Revue Information-Interaction-Intelligence, Revue en Sciences du Traitement de l'Information, Cépaduès Publ., Vol. 1, N. 1.
- (Huitema & Van Liere 00) Huitema, H., Van Liere R., Interactive Visualization of Protein Dynamics. Proceedings of VIS 2000, IEEE Visualization, Ertl T., Hamann B., Varshney A., Eds., IEEE Computer Society Publ., 2000, p. 465-468.
- (Ignatius et al. 96) Ignatius, E., Senay, H., Heller R. Perceptual Guidance for Visualization Design. Thomson Multimedia Ressource, 1996 Edition, Thomson Technology Labs, ISBN 0-534-50719-0, Chapter 2.
- (Ioannidis 96) Ioannidis Y. The FADIVA Benchmark, Version 0*.
<ftp://ftp.dis.uniroma1.it/pub/santucci/in/FADIVA/Ioannidis.ps>

- (Ishii & Ullmer 97) Ishii, H., Ullmer, B., *Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms*. Proceedings of CHI'97, Pemberton S. Ed., ACM Press: New York Publ., 1997, p. 234-241.
- (Jacko et al. 00) Jacko, J., Sears, A., Beaudouin-Lafon, M., Jacob, R. Eds. *Proceedings of CHI'2001*, ACM Press: New York Publ., 2000, 559 pages.
- (Keim 01) Keim, D., *Visual Exploration of large data Sets*. Communications of the ACM, Vol. 44., N. 8, 2001, p. 39-44.
- (Koleva et al. 00) Koleva, B., Schnädelbach, H., Benford, S., Greenhalgh, C., *Traversable Interfaces Between Real and Virtual Worlds*. Proceedings of CHI'2000, ACM Press: New York Publ., 2000, p. 233-240.
- (Laurillau 99) Laurillau, Y., *Synchronous Collaborative Navigation on the WWW*. Proceedings of CHI'99 Extended Abstracts, Altom M. & Williams M. Eds., ACM Press: New York Publ., 1999, p. 308-309.
- (Laurillau & Nigay 00) Laurillau, Y., Nigay, L., *Modèle de Navigation Collaborative Synchrone pour L'Exploration des Grands Espaces d'Information*. Conférence ERGO-IHM2000, CRT ILS&ESTIA Publ., p. 121-128.
- (Leung & Apperley 94) Leung, Y., Apperley, M., *A Review and Taxonomy of Distortion-Oriented Presentation Techniques*. ACM Transaction on Computer-Human Interaction, ACM Press: New York Publ., Vol. 1, N. 2, 1994, p. 126-160.
- (Mackay 96) MacKay, W., *Réalité augmentée : le meilleur des deux mondes*. La Recherche n° 285, 1996, p. 80-84.
- (MacKay 00a) MacKay, W., Ed., *Proceedings of DARE 2000, Designing Augmented Reality Environments*, ACM, 2000, 171 pages.
- (MacKay 00b) MacKay, W., *Augmented Reality: Dangerous Liaisons or the Best of Both Worlds?* Proceedings of DARE 2000, Designing Augmented Reality Environments, ACM, 2000, p. 170-171.
- (Mackinlay et al. 90) Mackinlay, J., Card, S., Robertson, G., *A Semantic Analysis of the Design Space of Input Devices*. Human Computer Interaction, Lawrence Erlbaum Publ., Vol. 5, N. 2 & 3, 1990, p. 145-190.
- (Mackinlay et al. 91) Mackinlay, J., Robertson, G., Card, S., *The Perspective Wall: Detail and context smoothly integrated*. Proceedings of CHI'91, Robertson S., Olson G., Olson J. Eds., ACM Press: New York Publ., 1991, p. 173- 179.
- (Martin 94) Martin, J.C., *Cadre d'étude de la multimodalité fondé sur les types et buts de coopération entre modalités*. Actes de la conférence InforMatique'94, L'interface des mondes réels & virtuels : Montpellier, EC2 Publ., 1994, p. 97-106.
- (Martin 99) Martin, J.C., *TYCOON: Six primitive types of cooperation for observing, evaluating and specifying cooperations*. Working notes of the AAAI Fall 1999 Symposium on Psychological Models of Communication in Collaborative Systems, 1999.
- (Maybury 93) Maybury M., Ed., *Intelligent Multimedia Interfaces*. AAAI/MIT Press Publ., ISBN 0-262-63150-4, 1993, 405 pages.
- (McGee et al. 00) McGee, D., Cohen, P., Wu, L., *Something from nothing: Augmenting a paper-based work practice via multimodal interaction*. Proceedings of DARE 2000, Designing Augmented Reality Environments, ACM, 2000, p. 71-80.

- (Mignot et al. 93) Mignot, C., Valot, C., Carbonell, N., An Experimental Study of Future 'Natural' Multimodal Human-Computer Interaction. Adjunct Proceedings of INTERCHI'93, Ashlung S., Mullet K., Henderson A., Hollnagel E., White T. Eds., ACM Press: New York Publ., 1993, p. 67-68.
- (Milgram & Kishino 94) Milgram, P., Kishino, F., A Taxonomy of Mixed Reality Visual Displays. IEICE Transactions on Information Systems, E77-D(12), 1994, p.1321-1329.
- (Mulhem & Nigay 96) Mulhem, P., Nigay, L., Interactive Information Retrieval Systems: From User Centred Interface Design to Software Design. Proceedings of SIGIR'96, ACM Press: New York Publ., 1996, p. 326-334.
- (Nigay 94) Nigay, L., Conception et modélisation logicielles des systèmes interactifs : application aux interfaces multimodales. Thèse de l'Université Joseph Fourier, 1994, 341 pages.
- (Nigay 95) Nigay, L. MATIS : un système multimodal d'information sur les transports aériens. Actes de la Conférence IHM'95, Vidéo, Cépaduès Publ., p.131-132.
- (Nigay & Coutaz 93) Nigay, L., Coutaz, J., A design space for multimodal interfaces: concurrent processing and data fusion. Proceedings of INTERCHI'93, Ashlung S., Mullet K., Henderson A., Hollnagel E., White T. Eds., ACM Press: New York Publ., 1993, p. 172-178.
- (Nigay & Coutaz 95a) Nigay, L., Coutaz, J., A Generic Platform for Addressing the Multimodal Challenge. Proceedings of CHI'95, Katz I., Mack R., Marks L. Eds., ACM Press: New York Publ., 1995, p. 98-105.
- (Nigay & Coutaz 95b) Nigay, L., Coutaz, J., Multifeature Systems: From HCI properties to Software Design. Intelligence and Multimodality in Multimedia Interfaces: Research and Applications, Lee, J. Ed., AAAI Press Publ., 1995.
- (Nigay & Coutaz 96) Nigay, L., Coutaz, J., Espaces de conception des interfaces multimédia et multimodales. Revue TSI, numéro spécial Multimédia et Collecticiel, Beaudouin-Lafon M. Ed., Vol. 15, N. 9, AFCET-Hermès Publ., 1996, p. 1195-1225.
- (Nigay & Coutaz 97) Nigay, L., Coutaz, J., Software architecture modelling: Bridging Two Worlds using Ergonomics and Software Properties. Formal Methods in Human-Computer Interaction, Palanque P. & Paterno F. Eds., Springer-Verlag: London Publ., ISBN 3-540-76158-6, 1997, p. 49-73.
- (Nigay & Vernier 98a) Nigay, L., Vernier, F., Design method of interaction techniques for large information spaces . Proceedings of AVI'98, Catarci, T., Costabile, M., Santucci, G., Tarantino, L. Eds, ACM Press Publ., 1998, p. 37-46.
- (Nigay & Vernier 98b) Nigay, L., Vernier, F., Navigational Interaction Techniques in the Search Results Space. CHI'98 workshop on Innovation and Evaluation in Information Exploration Interfaces, 1998.
- (Nigay et al. 01a) Nigay, L., Dubois, E., Troccaz, J., Compatibility and Continuity in Augmented Reality Systems. I3 Spring Days Workshop, Continuity in Future Computing Systems, 2001.

- (Nigay et al. 01b) Nigay, L., Renevier, P., Marchand, T., Salembier, P., Pasqualetti, P., La réalité cliquable : instrumentation d'une activité de coopération en situation de mobilité. Proceedings of IHM-HCI 2001, Vanderdonck J., Blandford A. Derycke A. Eds., CEPADUES Publ., 2001, p. 147-150.
- (Oviatt 99) Oviatt, S., Ten myths of multimodal interaction. Communications of the ACM, Vol. 42, N. 11, 1999, p. 74-81.
- (Oviatt et al. 00) Oviatt, S., Cohen, P., Wu, L., Vergo, J., Duncan, L., Suhm, B., Bers, J., Holzman, T., Winograd, T., Landay, J., Larson, J. & Ferro, D. Designing the user interface for multimodal speech and gesture applications: State-of-the-art systems and research directions. Human Computer Interaction, Lawrence Erlbaum Publ., Vol. 15, N. 4, 2000, p. 263-322.
- (Pfaff 85) Pfaff, G., Ed., User Interface Management Systems. Springer-Verlag Publ., New York, 1985.
- (Pirolli & Card 95) Pirolli, P., Card, S., Information Foraging in Information Access Environments. Proceedings of CHI'95, Katz I., Mack R., Marks L. Eds., ACM Press: New York Publ., 1995, p. 51-58.
- (Plaisant et al. 96) Plaisant, C., Milash, B., Rose, A., Widoff, S., Schneiderman, B., LifeLines: Visualizing Personal Histories. Proceedings of CHI'96, Tauber M., Bellotti V. Jeffries R., Mackinlay J., Nielsen J. Eds., ACM Press: New York Publ., 1996, p. 221-227.
- (Plénacoste et al. 01) Plénacoste, P., Lecolinet, E., Pook, S., Dumas, C., Fekete, J-D., Bibliothèques: comparaisons entre le réel et le virtuel en 3D, 2D zoomable et 2D arborescent. Proceedings of IHM-HCI 2001, Vanderdonck J., Blandford A. Derycke A. Eds., CEPADUES Publ., 2001, p. 55-59.
- (Rekimoto 95) Rekimoto, J., Katashi N., The World through the Computer: Computer Augmented Interaction with Real World Environments. Proceedings of UIST'95, ACM Press: New York Publ., 1995, p.29-36.
- (Renevier & Nigay 01) Renevier, P., Nigay, L., Mobile Collaborative Augmented Reality: the Augmented Stroll. Engineering for Human-Computer Interaction, Little R. & Nigay L. Eds, Springer-Verlag Publ., LNCS 2254, p. 315-334.
- (Roberston et al 93) Robertson, G., Card. S., Mackinlay, J., Information visualization using 3D interactive animation. Communications of the ACM, Vol. 36, N. 4, 93, p.57-71.
- (Salber et al 99) Salber, D., Dey, A., Abowd, G., The Context Toolkit:Aiding the Development of Context-enabled Applications. Proceedings of CHI'99, Altom M. & Williams M. Eds., ACM Press: New York Publ., 1999, p. 434-441.
- (Shneiderman 96) Shneiderman, B., The eyes have it : A task by data-type taxonomy for information visualizations. Proceedings of Visual Languages'96, IEEE Computer Science Press Publ., 1996, p. 336-343.
- (Starner et al. 00) Starner, T., Leibe, B., Singletary, B., Pair, J., MIND-WARPING: Towards Creating a Compelling Collaborative Augmented Reality Game. Proceedings of IUI 2000, International Conference on Intelligent User Interfaces, ACM Press: New York Publ., 2000, p. 256-259.
- (Stasko et al. 00) Stasko, J., Catrambone, R., Guzdial, M., McDonald, K., An Evaluation of Space-Filling Information Visualizations for Depicting Hierarchical Structures. Technical Report GIT-GVU-00-03.

- (Rouillard & Caelen 99) Rouillard, J., Caelen, J., A Multimodal and conversational application in natural language for information seeking on the World Wide Web : The HALPIN system. Proceedings of the 4th International Conference on Applications of Natural Language to Information Systems, 1999, p. 239-244.
- (Troccaz et al. 97) Troccaz, J., Peshkin, M., Davies, B., The use of localizers, robots and synergistic devices in CAS. Proceedings of MRCAS'97, 1997, p.727-736.
- (Turk & Robertson 00) Turk, M., Robertson, G. Eds, Perceptual user Interfaces. Special issue. Communications of the ACM, Vol. 43, N. 3, 2000, p. 32-70.
- (Turo & Johnson 92) Turo, D., Johnson, B., Improving the visualization of hierarchies with Treemaps: Design issues and experimentation. Proceedings of the IEEE Visualization'92, IEEE Computer Society Publ.,1992, p.124–131.
- (Tweedie 97) Tweedie, L., Characterizing Interactive Externalization. Proceedings of CHI'97, Pemberton S. Ed., ACM Press: New York Publ., 1997, p. 375-382.
- (Twidale et al. 97) Twidale, M., Nichols, D., Paice, C., Browsing is a Collaborative Process. Information Processing & Management, Elsevier Science Publ., Vol. 33, N. 6, 1997, p. 761-83.
- (UIMS 92) The UIMS Workshop Tool Developers: A Metamodel for the Runtime Architecture of an Interactive System. SIGCHI Bulletin, Vol. 24, N. 1, 1992, p. 32-37.
- (Ullmer et al. 98) Ullmer, B., Ishii, H., Glas, D., mediaBlocks: Physical containers, transports, and controls for online media. Proceedings of SIGGRAPH'98, ACM Press: New York Publ., 1998, p. 379-386.
- (Vernier & Nigay 97) Vernier, F., Nigay, L., Représentations Multiples d'une Grande Quantité d'Information. Actes de la conférence IHM'97, Cépaduès Publ., 1997, p. 183-190.
- (Vernier et al. 99) Vernier, F., Lachenal, Nigay, L., Coutaz, J., Coutaz Interface Augmentée par Effet Miroir. Actes de la conférence IHM'99, Cépaduès Publ., 1999, p. 158-165.
- (Vernier & Nigay 00a) Vernier, F., Nigay, L., A Framework for the Combination and Characterization of Output Modalities. Proceedings of DSV-IS2000, Palanque P. & Paterno F. Eds., Springer-Verlag Publ., LNCS 1946, 2000, p. 32-48.
- (Vernier & Nigay 00b) Vernier, F., Nigay, L., Modifiable Treemaps Containing Variable-Shaped Units. Work in Progress, Proceedings of VIS 2000, IEEE Visualization, IEEE Computer Society Publ., 2000, 2 pages.
- (Vernier 01) Vernier, F., La multimodalité en sortie et son application à la visualisation de grandes quantités d'information. Thèse de l'Université Joseph Fourier, 2001, 240 pages.
- (Walker et al. 01) Walker, A., Brewster, S., McGookin, D., Ng, A., Diary in the Sky: A Spatial Audio Display for a Mobile Calendar. Human Computer Interaction 20001, Joint Proceedings of HCI 2001 and IHM 2001, Blandford A., Vanderdonck J., Gray P. Eds., BCS Conference Series, Springer-Verlag Publ., 2001, p. 531-539.
- (Ware & Lewis 95) Ware, C., Lewis, M., The DragMag Image Magnifier. Proceedings of CHI'95, Katz I., Mack R., Marks L. Eds., ACM Press: New York Publ., 1995, p. 407-408.

- (Waterworth & Chignell 91) Waterworth, J., Chignell, M. A Model of Information Exploration. *Hypermedia*, London: Taylor Graham Publ., Vol. 3, N. 1, 1991, p. 35-38.
- (Wellner 91) Wellner, P., The DigitalDesk calculator. *Proceedings of UIST'91*, ACM Press: New York Publ., 1991, p. 27-34.
- (Wellner et al. 93) Wellner, P., MacKay, W., Gold, R. Eds. Special Issue of *Communications of the ACM*, Vol. 36, N. 7, 1993.
- (Whitaker et al. 95) Whitaker, R., Crampton, C., Breen, D., Tuceryan, M., Rose, E., Object Calibration for Augmented Reality. *Proceedings of Eurographics'95*, Vol. 14, N. 3, 1995, p. 15-28.
- (Young et al. 89) Young, R., Green, T., Simon, T. Programmable user models for predictive evaluation of interface designs. *Proceedings of CHI '89*, Bice K. & Lewis C. Eds, ACM Press: New York Publ., 1989, p. 15-19.