

# Un espace de classification pour l'interaction sur dispositifs mobiles

*Anne Roudaut -Eric Lecolinet*

Ecole Nationale Supérieure des Télécommunications (GET/ENST)  
46 rue Barrault, 75013, Paris, France  
{roudaut, elc} @enst.fr

## RESUME

En quelques années, la multiplication et la diversité des dispositifs mobiles dans nos vies quotidiennes ont entraîné de nouveaux défis pour les concepteurs : la taille de l'écran, la difficulté du pointage et de la saisie de texte, mais aussi la gestion des services en situation de mobilité constituent de nouvelles contraintes. De nombreuses techniques d'interaction ont été proposées ces dernières années pour répondre à ces nouvelles problématiques. Dans cet article, nous décrivons un espace de classification regroupant l'ensemble des techniques d'interaction existantes qui répondent aux problèmes spécifiques des dispositifs mobiles.

**MOTS CLES :** Dispositif mobile, technique d'interaction, modalité d'interaction, Dispositif physique, langage d'interaction.

## ABSTRACT

The multiplication and the diversity of mobile devices in everyday life involve new challenges for the designer: the size of the screen, text input, pointing, and the management of services in situation of mobility constitute new constraints. Many interaction techniques have been proposed these last years to answer these problems. In this article, we describe a space for classifying the existing interaction techniques that have been proposed for mobile devices.

**CATEGORIES AND SUBJECT DESCRIPTORS:** H.5.2 [Information Interfaces And Presentation]: User Interface. I.3.6 [Methodology and Techniques]: Interaction techniques.

**GENERAL TERMS:** Design, Human Factors.

**KEYWORDS:** Mobile device, interaction technique, modality, physical device, interaction language.

---

## INTRODUCTION

Les dispositifs mobiles ne sont pas conçus pour les mêmes usages qu'un ordinateur classique [6] : ils sont en effet utilisés la plupart du temps dans des conditions de mobilité où l'utilisateur ne dispose pas du même temps de réalisation de ses tâches et où son attention n'est pas constamment dirigée vers le dispositif. De plus, les facteurs de forme et les modes d'entrée des dispositifs mobiles sont différents (taille de l'écran, clavier restreint, stylet), ce qui rend l'interaction plus complexe. Par exemple, la petitesse de l'écran et des objets d'intérêt rend difficile le pointage ou la sélection de cibles avec un stylet. De plus, la manipulation d'un stylet nécessite d'utiliser les deux mains, alors qu'il est fréquent que l'une d'entre-elles ne soit pas disponible en situation de mobilité. L'entrée de texte est également un problème en soi sur ce type de dispositifs. Les claviers physiques les plus répandus sur les téléphones se limitent ainsi principalement à l'envoi de messages réduits tels que des SMS. Enfin, l'utilisateur peut être à tout moment interrompu par un appel, ou dérangé par des éléments de son environnement.

Les dispositifs mobiles nécessitent donc de nouvelles techniques d'interaction adaptées aux spécificités de leurs usages, de leurs facteurs de forme et de leurs capacités interactives. Divers travaux de recherche et plusieurs taxonomies ont déjà été réalisés dans ce domaine. Les taxonomies de [47] et [57] font état des modes d'entrée de texte sur dispositifs mobiles. D'autres ([34], [56]) traitent le problème de la visualisation d'information sur des surfaces d'affichages et celle de [1] cible ce problème sur des écrans de petite taille. Enfin [3] examine les techniques d'interaction employant le dispositif mobile lui-même comme des dispositifs d'entrée.

Les taxonomies existantes sont donc relativement nombreuses mais éparpillées, ciblant chacune un des aspects de l'interaction sur dispositifs mobiles. Afin de fournir une vue d'ensemble de l'interaction sur ces dispositifs mobiles et ce dans le but de faciliter la conception de nouvelles techniques, nous proposons un espace de classification regroupant les différentes techniques d'interaction en entrée et en sortie adaptées aux terminaux mobiles. En exposant cet espace nous présentons des exemples de l'état de l'art avant de conclure.

## DEFINITION DE L'ESPACE DE CLASSIFICATION

Selon Norman [45], la réalisation d'une tâche par l'utilisateur se fait selon un cycle action/perception : L'utilisateur percevant l'état du système, l'évalue pour décider de l'action nécessaire à la réalisation de sa tâche. Par la suite l'état du système change, ceci entraînant un nouveau cycle. L'exécution de ce cycle se base sur un ensemble varié de techniques d'interaction (en entrée et en sortie), que nous assimilons au terme de *modalité*.

L'approche des ergonomes et des psychologues est d'associer la modalité aux capacités de perception et d'action du sujet humain : le sens visuel, auditif, et tactile pour la perception (les sens olfactif et gustatif sont encore peu exploités), et le mouvement et la voix pour l'action. Avec une vision plus logicielle, Nigay [44] définit la modalité par un couple <D, L>, où D est un dispositif physique et L un langage d'interaction. Un dispositif physique est un élément du système qui acquiert (dispositif d'entrée) ou fournit (dispositif de sortie) des informations à l'utilisateur. Un langage d'interaction est un ensemble d'expressions bien formées et significatives. Par exemple, l'utilisateur inclinant son PDA pour zoomer sur une carte utilise la modalité < capteur, langage gestuel >. De plus cette définition introduit également la notion de modalité active ou passive. Contrairement à une modalité passive, une modalité active suppose une action explicite de l'utilisateur pour réaliser

une action ou pour percevoir les informations. Les services contextuels, comme la capture GPS pour localiser un utilisateur, constituent un exemple de modalité passive.

Notre espace de classification (Fig 1) regroupe ces définitions du terme de modalité. L'objectif premier est de combiner une approche système et une approche utilisateur. Ensuite, il aborde les modalités passives. Ce point est un atout car le développement de services contextuels sur dispositifs mobiles est aujourd'hui un enjeu. D'autre part il permet d'envisager des techniques d'interaction combinant plusieurs modalités à la fois (multi-modalité [44]). Le second objectif de cet espace de classification est de proposer une vue globale des tendances et failles des systèmes actuels, comme premier pas pour la conception de nouvelles techniques d'interaction. Nous présentons par la suite chaque dimension de cet espace avant de faire une synthèse étudiant l'apport de ces modalités face aux contraintes des dispositifs mobiles.

## MODALITE EN ENTREE PAR MOUVEMENT

L'interaction par mouvement comporte trois catégories classées selon les dispositifs en entrée : l'interaction sur dispositifs continus, l'interaction sur dispositifs discrets et l'interaction embarquée où l'utilisateur agit sur le dispositif mobile lui-même comme un périphérique d'entrée.

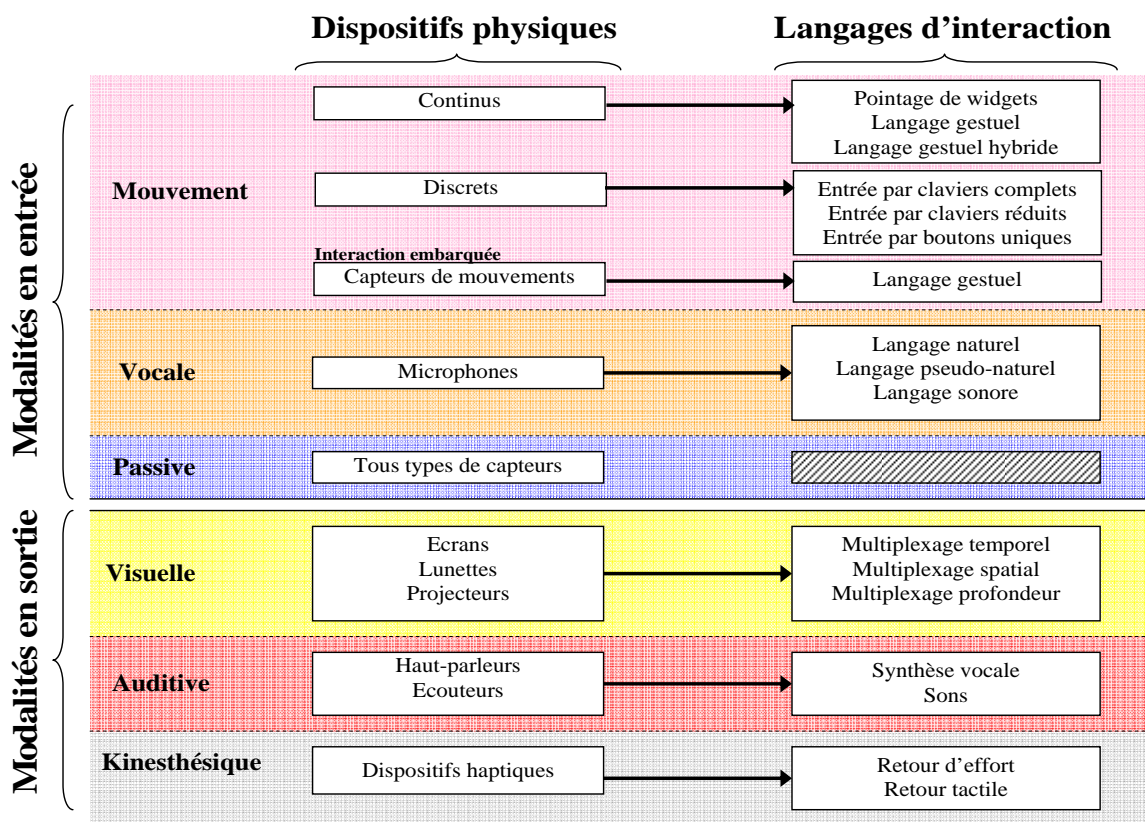


Figure 1. Espace de classification des techniques d'interaction en entrée et en sortie sur dispositifs mobiles

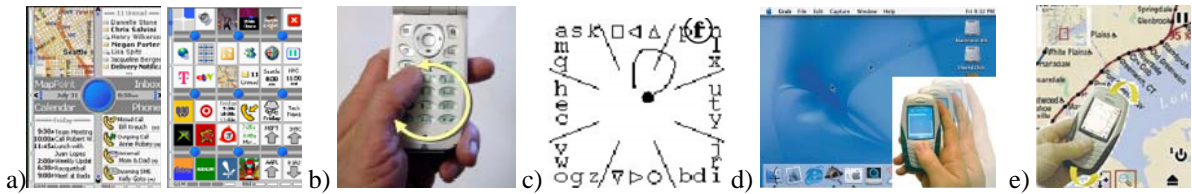


Figure 2. a) LauchTile [31], b) SmartPad [51], c) QuikWriting [46], d) Sweep & Point & shoot [2], e) CamBlaster [63]

### Interaction sur dispositifs continus

Les écrans tactiles présents sur de nombreux terminaux mobiles sont des dispositifs continus. C'est-à-dire qu'il récupère un signal continu représentant la position du pointeur à l'écran (rien ou plusieurs valeurs). Cependant les écrans tactiles sont aussi des dispositifs d'affichage, ce qui nous amène à distinguer trois formes de langage d'interaction : la *pointage de widgets* qui intervient lors de la sélection ou l'activation des widgets et nécessite l'attention de l'utilisateur sur l'écran, la *langage gestuel* qui permet de réduire cette attention visuelle et la *langage gestuel hybride* qui combine les deux approches.

**Pointage de widgets.** Les widgets sont des composants graphiques (menus, icônes, fenêtres ou barres de défilement) utilisés dans la plupart des systèmes graphiques. L'interaction se base sur les principes de la manipulation directe [54], où l'utilisateur agit directement sur les objets d'intérêt grâce à un dispositif de pointage (stylet ou doigt), au lieu d'exprimer les actions qu'il souhaite effectuer. Par exemple, dans LauchTile [31] (Fig 2a) l'utilisateur interagit grâce à un large bouton circulaire lui permettant de naviguer parmi ses applications. Le style d'interaction WIMP, actuellement mis en œuvre sur les dispositifs mobiles grand public, utilise ce principe.

**Langage gestuel.** Le langage gestuel repose sur l'interprétation de la dynamique des actions physiques de l'utilisateur. Par exemple, avec le système de reconnaissance d'écriture EdgeWrite [60], l'utilisateur forme les lettres au stylet sur un écran tactile, selon des patrons de marques prédéfinies (par exemple, pour la lettre I, l'utilisateur fait une marque rectiligne de haut en bas). Dans AppLens [29], des gestes directionnels sont utilisés pour naviguer dans un ensemble d'application. Une fois les gestes connus, l'utilisateur n'est pas forcé de regarder l'écran pour interagir. Cependant cette phase d'apprentissage des gestes correspondant aux actions peut être contraignante pour l'utilisateur. Notons que la notion de dispositif continu est relative au niveau d'abstraction dans lequel on se place. Un clavier (qui techniquement est un dispositif discret) peut être considéré au niveau de l'interaction comme un dispositif continu, comme avec la technique du SmartPad [51] (Fig 2b) où l'utilisateur effectue un geste sur plusieurs touches du clavier mobile.

**Langage gestuel hybride.** Le langage gestuel hybride résout le problème d'apprentissage de l'interaction gestuelle en utilisant des widgets pour accompagner l'utilisateur dans son geste. La technique QuikWriting [46] (Fig 2c) illustre ce concept : à la différence de EdgeWrite, les marques sont associées à un menu.

L'utilisateur fait une marque allant du centre vers la cible souhaitée. Avec l'habitude, l'utilisateur passe implicitement du mode novice (où il apprend la position des lettres dans le menu) au mode expert (le geste moteur étant imprimé en mémoire, il connaît désormais le placement des lettres souhaitées). Les menus circulaires (Pie menus [28], Marking menus [37], Flow menus [23], Control menus [39], Tracking menu [21]) sont des variantes des menus contextuels linéaires utilisant un principe similaire. Grâce à ces techniques, seule la direction du geste suffit à désigner un item contrairement aux menus linéaires qui imposent de regarder l'écran pour pointer un item dont la distance à la base est figée. L'avantage de telles techniques est que l'utilisateur expert peut lancer des commandes sans regarder l'écran.

### Interaction sur dispositifs discrets

A la différence des dispositifs continus, l'information récupérée par un dispositif discret est une valeur booléenne (par exemple, une touche de clavier est appuyée ou non). Nous distinguons les langages d'interaction associés selon les caractéristiques physiques des dispositifs : les *entrées par claviers complets*, les *entrées par claviers réduits* et les *entrées par boutons uniques*.

**Entrée par claviers complets.** Les claviers présents sur terminaux mobiles adoptent généralement une configuration de touche qwerty ou azerty, plus habituelles pour la majorité des utilisateurs. Ce n'est cependant pas la stratégie la plus performante et d'autres configurations ont été proposées pour améliorer la saisie de texte comme le clavier Dvorak où les lettres les plus fréquemment utilisées sont rapprochées, le clavier Metropolis où les voyelles sont organisées au centre, autour de la touche espace, et d'autres encore (Knits, Fitaly, Opti, la typologie de [47] synthétise l'ensemble de ces approches). Certaines de ces configurations sont particulièrement pertinentes pour les dispositifs mobiles, mais elles ont l'inconvénient de nécessiter une phase d'apprentissage plus ou moins longue. Elles peuvent aussi s'avérer utiles pour les claviers virtuels affichés sur l'écran, d'autant plus que ceux-ci peuvent aisément prendre toutes les formes souhaitées. Le cas des claviers virtuels peut par ailleurs être considéré comme un cas particulier de pointage de widgets (décrit dans la section précédente).

**Entrée par claviers réduits.** Sur ce type de claviers, les techniques d'entrée évoluent vers du "Multitap" comme dans les claviers T12 qui imposent d'appuyer plusieurs fois sur la même touche pour atteindre une lettre. D'autres approches combinent plusieurs touches en faisant des «accords» comme dans [61] où l'utilisateur doit former les patrons de lettre EdgeWrite sur quatre boutons.

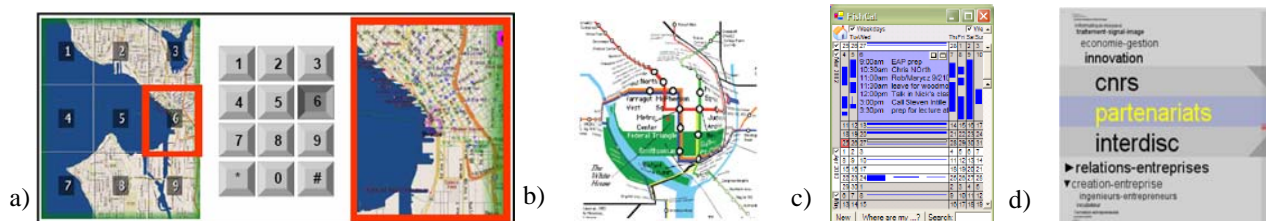


Figure 3. a) ZoneZoom [52], b) Vue FishEye, c) DateLens [8], d) FishTree [38]

**Entrée par boutons uniques.** Les boutons physiques sont souvent utilisés pour lancer des commandes. Avec des boutons évolués (molette ou joystick), l'utilisateur peut se déplacer dans plusieurs directions sur une fenêtre

### Interaction embarquée

Les interfaces embarquées (Embodied user interface [19]) utilisent le dispositif mobile lui-même comme dispositif d'entrée. Ceci est rendu possible par divers types de capteurs : capteurs de déplacement (GPS, détecteurs de mouvements, etc.), capteurs inertiels (accéléromètres, inclinomètres, gyroscopes), capteurs d'identification (tags RFID, cellules biométriques) ou encore les cellules photosensibles (photo/caméra) que l'on trouve sur la plupart des dispositifs mobiles. Ces capteurs permettent à l'utilisateur d'interagir par gestes.

**Langage gestuel.** Dans [2] (Fig 2d), l'utilisateur utilise son téléphone mobile comme dispositif de pointage. Grâce aux différences d'images récupérées par la caméra, le téléphone devient une "souris mobile". La caméra peut être aussi utilisée pour contrôler un zoom ou se déplacer sur un plan, comme dans CamBlaster [63] (Fig 2e). On peut aussi utiliser la caméra pour prendre en photo des tags visuels, et récupérer un service associé [43]. Le passage automatique du mode paysage au mode portrait peut être réalisé en détectant l'orientation du dispositif [27], une idée qui a été reprise dans l'iPhone d'Apple. Enfin, la détection des chocs a été utilisée pour le couplage de dispositif : dans [26] l'utilisateur couple deux tablets PC en les choquant l'un contre l'autre.

### MODALITE EN ENTREE VOCALE

Par reconnaissance vocale, les systèmes peuvent interpréter les commandes d'un utilisateur. Ce processus de reconnaissance en toutes conditions étant encore un problème de recherche non totalement résolu, la majorité des systèmes se basent sur une reconnaissance de mots clefs (langage pseudo-naturel) ou de sons (langage sonore).

**Langage pseudo-naturel.** Le langage pseudo-naturel est défini par une grammaire spécifique. Par exemple dans [33], l'utilisateur peut effectuer des requêtes sur son dispositif mobile grâce à des phrases suivant le schéma : « Je voudrais des informations sur /Canal d'Information/ dans /Arrondissement(s)/ pour /Date(s)/ ». Des sites tels que [64] proposent notamment l'accès à ce type d'interaction pour le grand public.

**Langage sonore.** Ce langage repose sur l'interprétation d'entrées sonores non verbales. Par exemple, dans cer-

tains jeux de Nintendo DS, l'utilisateur souffle sur le microphone pour interagir avec les personnages. Un autre exemple applicable aux cas des dispositifs mobiles est la technique du "Query by humming" [10], qui consiste à siffler, fredonner ou chanter un air de musique pour retrouver un morceau de musique dans une base de données.

### MODALITE EN ENTREE PASSIVE

Les modalités passives permettent aux interfaces utilisateurs de s'adapter au contexte d'interaction sans action explicite de l'utilisateur. Elles se basent également sur l'utilisation de capteurs comme le microphone et les capteurs utilisés dans les interfaces embarquées, mais aussi les capteurs de luminosité, de température, de pression, etc. Les informations récupérées par ces capteurs (informations sur l'utilisateur, sur l'environnement physique et social ou sur les ressources disponibles), permettent de détecter des changements pouvant conduire à de nouvelles situations. Par exemple, dans [53] les alertes du téléphone changent (sonnerie, vibreur, ajustement du niveau de la sonnerie ou coupure de la sonnerie) selon la position du mobile (dans la main, sur une table, dans un sac ou à l'extérieur). Les interfaces "follow me" sont un autre exemple où l'interface s'adapte automatiquement aux ressources disponibles autour de l'utilisateur. De nombreux autres exemples pourront être trouvés dans l'article [15]. L'action de l'utilisateur étant implicite, nous considérons qu'il n'y a pas de langage d'interaction puisque ce sont les variations du contexte d'interaction qui influent sur le système.

### MODALITE EN SORTIE VISUELLE

Ces modalités de sortie sont les plus courantes, mais ce n'est pas surprenant puisque 70% des informations perçues par l'homme proviennent du canal visuel [59]. Sur dispositif mobile, la taille de l'écran impose d'adopter des stratégies pour présenter les informations à l'utilisateur. Nous classons ces stratégies grâce aux travaux de [48], qui utilise la métaphore du multiplexage<sup>1</sup> pour décomposer les modalités visuelles en trois types : multiplexage temporel, spatial ou en profondeur.

**Le multiplexage temporel.** Seule une vue de l'information est affichée à un moment donné, et cette vue change continuellement dans le temps en fonction des interactions de l'utilisateur. Ce contrôle peut être de plu-

<sup>1</sup>Faire passer plusieurs informations à travers un même support de transmission

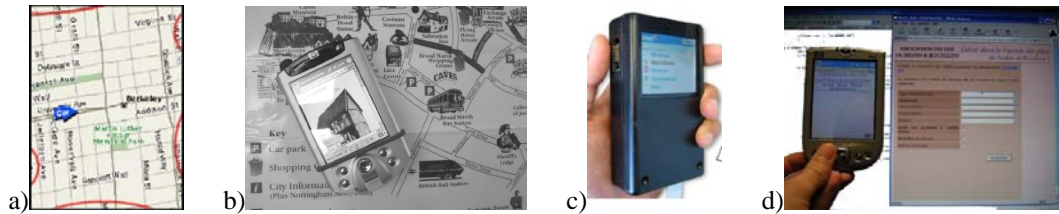


Figure 4. a) Halo [5], b) Just point and click [50], c) Dispositif haptique [42], d) Sedan Bouillon [4]

sieurs types : zoom ou filtrage de l'information. Les interfaces Pan & Zoom en sont un exemple : Les informations sont affichées sur un écran virtuel plus large que l'écran physique et l'utilisateur se déplace ou zoome pour trouver son point d'intérêt. Par exemple dans ZoneZoom [52] (Fig 3a) l'utilisateur appuie sur une touche de clavier pour zoomer sur une zone correspondante à l'écran. Le zoom sémantique [7] est un autre exemple de contrôle par zoom. A la différence des interfaces Pan & Zoom, la représentation des données change au cours du zoom pour faciliter la visualisation et la compréhension comme dans [35]. D'autres techniques de zoom utilisent une représentation hiérarchique de l'information comme par exemple les TreeMap (introduit par [55]) sur PDA [18]. Une autre façon de contrôler l'affichage de l'information est le filtrage. Par exemple, dans les lecteurs de médias, l'utilisateur peut trier ses morceaux de musique en fonction du style, du titre des albums ou des noms d'artistes. Fathumb [32] utilise aussi ce principe pour retrouver des informations hiérarchiques grâce au filtrage selon les attributs des données (facettes).

**Le multiplexage spatial.** Les informations sont affichées simultanément sur une même vue. Elles permettent à l'utilisateur de garder une vue globale mais aux prix de déformations. Les techniques Focus + Context comme les FishEyes (Fig 3b) illustrent ce concept : une déformation optique centrée sur le curseur permet d'afficher plus de détails dans la zone d'intérêt. Le multiplexage spatial est souvent combiné avec du multiplexage temporel afin de détailler une zone d'information sans perdre le contexte global. DateLens [8] (Fig 3c) utilise ce principe : un calendrier global est affiché et le jour sélectionné est agrandi au centre, ceci déformant la représentation des autres jours. Le FishTree [38] (Fig 3d) fonctionne selon le même principe mais pour des représentations hiérarchiques.

**Le multiplexage en profondeur.** Plusieurs vues de l'information sont présentées au même moment grâce à un jeu de superposition et de transparence. Par exemple Macroscopie [40] superpose deux cartes, la carte globale et une autre en transparence qui représente une vue zoomée d'une partie de la carte globale. Dans Halo [5] (Fig 4a) des cercles englobant sont superposés à une carte pour indiquer la distance et la direction de lieux recherchés se trouvant en dehors de la vue courante. D'autres outils comme les See-Through tools [9] utilisent ce concept de transparence (par exemple, une Toolglass est une surface transparente, sur laquelle sont disposés des outils semi-transparents qui agissent sur les objets de l'application située en dessous). Sur dispositifs

mobiles, la mise en œuvre de ce principe amène de nouvelles techniques d'interaction comme dans [50] (Fig 4b) où le dispositif mobile devient une Toolglass : en superposant le mobile à un plan papier, l'utilisateur peut récupérer des informations sur les lieux.

#### MODALITE EN SORTIE AUDITIVE

Les modalités en sortie auditive se basent sur l'utilisation de haut-parleurs ou d'écouteurs. Le langage d'interaction associé peut être de deux types : synthèse vocale ou sons.

**Synthèse vocale.** Les développements des technologies favorisant l'accès des malvoyants aux interfaces font l'objet de nombreuses recherches concernant la synthèse vocale. L'article [17] traite de façon complète ces techniques qui ne seront pas développées ici faute de place.

**Sons.** L'utilisation du son peut être très efficace pour améliorer l'interaction sans avoir à surcharger l'interface graphique. Notamment, les expériences [12] sur PDA ont démontré que l'interaction avec notifications sonores sur des boutons réduits s'avérait aussi efficace que l'interaction avec des boutons de taille normale. A terme, les utilisateurs experts peuvent se passer de la visualisation, évitant ainsi de porter une attention constante sur l'écran. Le son sert aussi d'alerte lors d'un changement de l'état du système, notamment lors d'un appel, ou pour signaler la réception de messages. L'utilisation des écouteurs révèle aussi un autre atout du son qui est l'utilisation de la stéréophonie. Par exemple, dans [25] l'utilisateur navigue dans une collection de morceaux de musique et se repère spatialement grâce à la localisation des musiques dans ses écouteurs.

#### MODALITE EN SORTIE KINESTESIQUE

Les modalités en sortie kinesthésique reposent sur l'utilisation de dispositifs à retour sensoriels. Les sensations perçues par l'utilisateur se décomposent en deux catégories : le retour tactile et le retour d'effort.

**Le retour tactile.** Grâce à ce type de système, l'utilisateur peut percevoir des états de surface (rugosité, texture), des températures, des glissements ou des détections d'arêtes. L'utilisation du vibreur des téléphones mobiles est l'exemple le plus répandu de retour tactile. Utilisé simplement pour remplacer la sonnerie lorsque celle-ci n'est pas souhaitable, il peut aussi servir pour augmenter l'interactivité : Dans des jeux [41] ou comme dans Tactimod [16] où l'utilisateur se dirige sur un parcours grâce à deux vibreurs placés sur chaque côté du dispositif lui indiquant où tourner. Dans [42] (Fig 4c) le dispositif mobile est augmenté d'une technologie plus

complexe : un stimulateur tactile qui donne une impulsion à l'utilisateur pour l'aider à naviguer dans une page Web. Avec le Zoom Haptique [62], l'utilisateur reçoit des impulsions tactiles lorsqu'il zoome pour se repérer et lui éviter de se perdre dans le zoom.

**Le retour d'effort.** Les systèmes à retour d'effort consistent à rendre perceptible à l'utilisateur des forces de contact, de dureté, de poids ou d'inertie d'un objet. Ces techniques, lourdes à mettre en place, ne sont à priori pas appliquées aux cas des dispositifs mobiles. On pourra consulter l'article [36] pour plus de détail sur ce sujet.

## SYNTHESE DES MODALITES

Comme nous l'avons exposé en introduction, l'interaction sur dispositifs mobiles doit prendre en compte de nouvelles contraintes, qui sont liées à trois facteurs : le facteur de forme (petite taille et prise en main), les modes d'entrée (réduits et limités) et l'usage en conditions de mobilité (attention réduite et manipulation à une main). Dans cette section, nous synthétisons chaque modalité en détaillant les apports et les inconvénients liés à ces contraintes.

### Modalité en entrée par mouvement

Sur dispositifs continus, cette modalité se heurte à la contrainte de la petite taille du dispositif et des objets d'intérêt. En particulier, la loi de Fitts [20] nous indique que la vitesse de sélection d'une cible dépend de la distance et de la taille de la cible. Ceci rend les interfaces à widgets mal adaptées à l'interaction sur dispositifs mobiles car les boutons sont souvent petits et donc difficiles à atteindre au stylet. Une autre contrainte majeure est celle de la mobilité : l'utilisateur souhaiterait souvent pouvoir utiliser son dispositif mobile avec une seule main et interagir directement avec les doigts. Mais, du fait de la taille des doigts, ce type d'interaction pose de nouveaux problèmes comme le manque de précision du pointage ou l'occlusion de l'espace écran. Les techniques d'offset [49] permettent de résoudre les problèmes d'occlusion en affichant un curseur décalé par rapport au doigt. Le pointage sémantique qui augmente la taille de la cible dans l'espace moteur sans changer sa représentation visuelle peut permettre d'améliorer la précision. Cependant, cette technique ne peut être utilisée telle quelle sur un dispositif à pointage absolu, comme un écran tactile, car l'espace moteur et l'espace visuel sont confondus. Une autre stratégie, illustrée dans SpiraList [29] pour la sélection d'items dans une liste, repose sur l'utilisation de techniques de filtrage des mouvements non significatifs et de stabilisation du curseur. Le système stabilise ainsi le déplacement du doigt et compense les éventuelles maladresses engendrées dans l'espace moteur.

Le problème majeur concernant l'interaction sur dispositifs discrets concerne la saisie de textes longs. Les configurations de clavier classiques (qwerty/azerty) s'avèrent encore moins bien adaptées aux petits terminaux qu'aux ordinateurs de bureau. D'une part les petits claviers physiques sont utilisés avec un ou deux doigts et non à deux

mains, d'autre part les conditions de mobilité rendent les tâches d'entrées encore plus difficiles. Dans le cas des claviers virtuels, les configurations peuvent changer comme par exemple avec Glyph [58] où l'utilisateur spécifie des lettres par la composition de formes primitives (boucle, barre verticale ou horizontale). Cependant les claviers virtuels ont l'inconvénient d'occuper une place non négligeable à l'écran et de ne pas fournir de retour tactile contrairement aux claviers physiques.

L'interaction embarquée présente de nombreux avantages. La manipulation est naturelle, simple, ludique et se fait souvent à seule une main. Les usages de ce type de manipulation commencent à se répandre, notamment avec l'arrivée sur le marché de l'iPhone d'Apple, ou encore de la console Wii qui utilise un système capable de détecter la position, l'orientation et les mouvements des manettes dans l'espace. De plus, les travaux sur les interfaces tangibles [30] peuvent être appliqués à ce type d'interaction. Un possible inconvénient en conditions de mobilité serait la "peur du ridicule" (par exemple devoir secouer son dispositif en public).

### Modalité en entrée vocale

Nous avons vu que l'entrée textuelle avec un clavier s'avérait difficile. Les modalités vocales peuvent être une solution, le langage oral étant plus naturel et facile. Cependant ce type d'interaction se heurte à des problèmes techniques (fiabilité des systèmes de reconnaissance vocale) ainsi qu'aux contraintes de mobilité. En effet, en situation de mobilité, le contexte de l'utilisateur peut être bruyant, et interférer sur ses commandes vocales. De plus, parler en public à son dispositif mobile peut être gênant.

### Modalité en entrée passive

Ces types de modalité permettent d'enrichir et de multiplier les services sur dispositifs mobiles. L'utilisateur peut alors se décharger de certaines tâches (telles qu'éteindre son téléphone en réunion) ou profiter de nouveaux services (comme récupérer le journal du jour en passant près d'une borne). Cependant les informations provenant du contexte d'utilisation sont nombreuses et nécessitent une infrastructure logicielle complexe pour leur traitement. Sur dispositif mobile il est difficile de déployer une telle architecture (peu de mémoire, gestion de la batterie et puissance limitée des processeurs). Par ailleurs les comportements des utilisateurs sont très variables d'un utilisateur à un autre et il peut être difficile d'inférer ce que souhaite l'utilisateur (par exemple, certains souhaitent être joignable au domicile alors que d'autres coupent leur téléphone ou n'acceptent que les appels familiaux). L'utilisateur doit donc pouvoir finement contrôler ses services contextuels. Enfin ces techniques peuvent poser des problèmes de protection de la vie privée d'utilisateurs qui pourraient se sentir "suivis" dans leurs déplacements.

### Modalité en sortie visuelle

Nous avons vu que la petite taille de l'écran est un problème pour la modalité visuelle. Malgré des techniques

d'organisation spatiales de l'information plus adaptées, les dispositifs mobiles ne peuvent, à priori, pas devenir des ordinateurs de bureau réduits. On peut cependant compter sur des dispositifs de sortie additionnels qui augmentent la surface d'affichage tels que des lunettes opaques ou semi-transparentes [63], ou des mini projecteurs embarqués sur dispositifs mobiles [14]. De plus, on peut tirer avantage des conditions de mobilité, car l'utilisateur est amené à rencontrer des surfaces d'affichage plus larges (ordinateurs conventionnels, tableaux d'affichages numériques ou grands écrans). Le couplage des dispositifs mobiles à ces surfaces est une piste prometteuse, comme dans [4] (Fig. 4d) où l'utilisateur place le menu de navigation d'un site sur son PDA et visualise le reste du site sur un ordinateur conventionnel.

#### **Modalité en sortie auditive**

Nous avons vu que les modalités auditives pouvaient améliorer l'interaction sur dispositif mobile en déchargeant les modalités visuelles. Cependant, de la même façon que les modalités vocales, ces modalités peuvent gêner l'utilisateur ou son entourage en conditions de mobilité. De plus un niveau sonore ambiant élevé risque de dégrader la qualité de l'interaction en masquant les informations auditives. Les modalités auditives restent cependant utiles pour les alertes (prévenir l'utilisateur d'un changement d'état du système). L'utilisation des écouteurs, qui se généralise avec l'émergence des lecteurs mp3, peut s'avérer intéressante, par exemple la stéréophonie pour se repérer dans l'espace [25].

#### **Modalité en sortie kinesthésique**

L'avantage de ces modalités est de décharger les canaux visuels et auditifs de l'utilisateur, sans engendrer de gêne en condition de mobilité. Elles améliorent aussi l'interaction de façon ludique, comme pour des jeux [41]. Cependant, mis à part le vibreur, les dispositifs haptiques sont souvent difficiles à mettre en place sur terminaux mobiles (taille importante, forme non adaptée, consommation électrique). De plus, les informations perçues par l'utilisateur ont une bande passante limitée, même pour des utilisateurs malvoyants ayant développés leurs aptitudes à reconnaître des formes ou des rythmes [22].

#### **Multimodalité**

La multimodalité combine de façon harmonieuse différentes modalités au cours de l'interaction. L'avantage principal est de tirer profit des atouts de chaque modalité. Sur dispositif mobile on trouve diverses applications qui couplent plusieurs modalités : les modalités en entrée vocale et par mouvement dans [2], les modalités en sortie haptiques et visuelle comme dans [41], les modalités en sortie sonore et haptique [25], etc. Le deuxième objectif est de calquer l'interaction sur les actions de la vie quotidienne qui sont en majorité multimodale. Par exemple, dans "Put-That-There" [11] l'utilisateur exprime la commande vocale « met ça là » en pointant l'objet à sélectionner (ça) puis en pointant un endroit (là). Enfin la multimodalité peut spécifiquement répondre à certains

besoins propres aux conditions de mobilité. Les études de [13] ont notamment montré que l'utilisation des modalités changeait selon le contexte d'interaction (au repos l'utilisateur préfère la modalité en entrée par mouvement tactile, alors qu'en marchant, il utilise la modalité en entrée vocale pour la même tâche). C'est le concept d'équivalence des propriétés CARE de [44].

#### **CONCLUSION**

Dans cet article nous avons présenté un espace de classification pour l'interaction sur dispositifs mobiles. Après avoir décrit les dimensions de cet espace, nous avons étudié l'impact des conditions de mobilité sur les techniques d'interaction en entrée et en sortie. En proposant une vue synthétique des techniques existantes, cette étude a pour but d'aider au développement de nouvelles techniques d'interaction spécifiquement adaptées à la mobilité. Nous projetons désormais de développer des techniques d'interaction innovantes afin de résoudre certains des problèmes mis en évidence par cette étude.

#### **BIBLIOGRAPHIE**

1. Baffoun, S., Robert, J.M., Etat de l'art des techniques de présentation d'information sur écran d'assistant numérique personnel, IHM 2006.
2. Ballagas, R., Rohs, M., Sheridan, J.G., Sweep and Point & Shoot: Phonedcam-Based Interactions for Large Public Displays, CHI 2005.
3. Ballagas, R., Rohs, M., Sheridan, J., Borchers, J., The Smart Phone: A Ubiquitous Input Device, Pervasive Computing 2006.
4. Balme, L., Demeure, A., Calvary, G., Coutaz, J., Sedan-Bouillon: A Plastic Web Site, Plastic Services for Mobile Devices (PSMD), Workshop Interact 2005.
5. Baudisch, P., Rosenholtz, R., Halo: a Technique for Visualizing Off-Screen Locations, CHI 2003.
6. Beaudoin-Lafon, M., Ceci n'est pas un ordinateur – Perspectives sur l'Interaction Homme-Machine, TSI An 2000.
7. Bederson B., Hollan J.D., Perlin K., Meyer J., Bacon D. et Furnas G., Pad++: A Zoomable Graphical Sketchpad For Exploring Alternate Interface Physics, Visual Languages and Computing 1996.
8. Bederson, B., Czerwinski, M.P., Robertson, G.G., A Fish-eye Calendar Interface for PDAs: Providing Overviews for Small Displays, CHI 2002.
9. Bier, E.A., Stone, M.C., Fishkin, K., Buxton, W., Baudel, T., A Taxonomy of See-Through Tools, CHI 1994.
10. Birmingham, W., Dannenberg, R., Pardo, B. Query by humming with the vocals search system, communication of the ACM, 2006.
11. Bolt, R. A., Put-that-there: Voice and gesture at the graphics interface, Computer Graphics and Interactive Techniques, 1980.
12. Brewster, A., Overcoming the Lack of Screen Space on Mobile Computers, Ubicomp 2002.
13. Calvet, G., Salembier, S., Kahn, J., Zouinar, M., Etude empirique de l'interaction multimodale en mobilité : approche méthodologique et premiers résultats, IHM 2005.
14. Cao, X., Balakrishnan, R., Interacting with Dynamically Defined Information Spaces using a Handheld Projector and a Pen, UIST 2006.

15. Chen, G., Kotz, D., A Survey of Context-Aware Mobile Computing Research, Technical Report, 381, 2000.
16. Couture, N., Minel, S., TactiMod dirige et oriente un piéton, Ubimob 2006.
17. D'Alessandro, C., 33 ans de synthèse de la parole à partir de texte : une promenade sonore (1968-2001), Traitement automatique de la parole, 2001.
18. Engdahl, B., Köksal, M., Marsden, G., Using Treemaps to Visualize Threaded Discussion Forums on PDAs, CHI 2005.
19. Fishkin, K. P., Moran, T. P. & Harrison, B. L., Embodied User Interfaces: Towards Invisible User Interfaces, EHCI 1998.
20. Fitts, P.M., The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology, 1954.
21. Fitzmaurice, G., Khan, A., Pieké, R., Buxton, B., Kurtenbach, G., Tracking Menus, UIST 2003.
22. Fritz, J.P., Barner, K.E, Design of a Haptic Data Visualization System for People with Visual Impairments, IEEE rehabilitation engineering, 1999
23. Guimbretière, F., Winograd, T., Flow Menu: Combining Command, Text and Parameter Entry, UIST 2000.
24. Harrison, B., Fishkin, K. P., Gujar, A., Mochon, C., Want, R., Squeeze Me, Hold Me, Tilt Me! An Exploration of Manipulative User Interfaces, CHI 1998.
25. Hiipakka J., Lorho, G., a spatial audio user interface for generating music playlists, Auditory Display, 2003.
26. Hinckley, K., Synchronous Gestures for Multiple Persons and Computers, UIST 2003.
27. Hinckley, K., Pierce, J., Sinclair, M., Horvitz, E., Sensing Techniques for Mobile Interaction, UIST 2000.
28. Hopkins, D., The design and implementation of pie menus. Dr. Dobb's Journal, 16(12), p16-26, 1991.
29. Huot, S., Lecolinet, E., Focus+Context Visualization Techniques for Displaying Large Lists with Multiple Points of Interest on Small Tactile Screens, Interact 07.
30. Ishii, H., Ullmer, B., Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms, CHI 1997.
31. Karlson, A K., Bederson, B., SanGiovanni, J., AppLens and LaunchTile: Two Designs for One-Handed Thumb Use on Small Devices, CHI 2004.
32. Karlson, A K., Robertson, G., Robbins, D. C., Czerwinski, M., Smith, G., FaThumb: A Facet-based Interface for Mobile Search, CHI 2006.
33. Karsenty, L., Sire, S., Causse, M., Deherly, N., Quel impact de l'entrée vocale sur la conception graphique d'un service mobile ?, IHM 2005.
34. Keim, D.A., Information Visualization and Visual Data Mining, IEEE Transactions on Visualization and Computer Graphics, 2002.
35. Khella, A., Bederson, B., Pocket PhotoMesa: A Zoomable Image Browser for PDAs, MUM 2004.
36. Kuchenbecker, K.J., Fiene, J., Niemeyer, G., Improving Contact Realism through Event-Based Haptic Feedback. IEEE Trans. Vis. Comput. Graph , 2006.
37. Kurtenbach, G., Buxton, W., Issues in Combining Marking and Direct Manipulation Techniques, UIST 1991.
38. Lecolinet E., Nguyen, D., Représentation focus+contexte de listes hiérarchiques zoomables, IHM 2006.
39. Lecolinet, E., Pook, S., Interfaces zoomables et « Control menus » : Techniques focus+contexte pour la navigation interactive dans les bases de données, Revue les Cahiers du numérique, 2002.
40. Lieberman, H., A Multi-Scale, Multi-Layer, Translucent Virtual Space, IV 1997.
41. Linjama, J., kaaresoja, T., Novel, minimalist haptic gesture interaction for mobile devices, NordiCHI 2004.
42. Luk, J., Pasquero, & Al. , J., A Role for Haptics in Mobile Interaction: Initial Design Using a Handheld Tactile Display Prototype, CHI 2006.
43. McCune, J.M., Perrig, J.M., Reiter, M. K., Seeing-Is-Believing: Using Camera Phones for Human-Verifiable Authentication, Security and privacy, 2004.
44. Nigay, L., Coutaz, J., Espaces conceptuels pour l'interaction multimédia et multimodale, TSI, spéciale Multimédia et collectif, AFCET & HERMES, 1996.
45. Norman, D.A., Draper, S.W., User Centered System Design: New Perspectives on Human-Computer Interaction. Erlbaum Associates, Hillsdale, NJ, 1986.
46. Perlin, K., QuickWriting: continuous stylus-based text entry, User interface software and technology, 1998.
47. Poirier, F., Schadle, I., État de l'art des méthodes de saisie de données sur dispositifs nomades, IHM 2004.
48. Pook, S., Interaction et contexte dans les interfaces zoomables, Thèse, Paris, 2001.
49. Potter, R., Weldon, L., Shneiderman, B., Improving the accuracy of touch screens: an experimental evaluation of three strategies, CHI 88.
50. Reilly, D., Welsman-Dinelle, M., Bate, C., Inkpen, K., Just Point and Click? Using Handhelds to Interact with Paper Maps, MobileHCI 2005.
51. Rekimoto, J., Ishizawa, T., Oba, H., SmartPad: A Finger-Sensing Keypad for Mobile Interaction, CHI 2003.
52. Robbins, D. C., Cutrell, E., Sarin, R., Horvitz, E., Zone-Zoom: Map Navigation for Smartphones with Recursive View Segmentation, AVI 04.
53. Schmidt A., Aidoo K.A., Takaluoma A., Tuomela U., Van Laerhoven K., Van de Velde W., Advanced interaction in context, HUC 1999.
54. Shneiderman, B., Direct Manipulation: A Step Beyond Programming Languages, IEEE Computer, 1983.
55. Shneiderman, N., Tree visualization with tree-maps: 2-d space-filling approach, ACM Transactions on Graphics, 1992.
56. Shneiderman, B., The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations, IEEE Symposium on Visual Languages, 1996.
57. Sirisena, A., Mobile Text Entry, Rapport en informatique, Université de Canterbury (NZ), 2002.
58. Uguen G., Poirier F., Saisie de données pour interfaces réduites avec Glyph : principes, niveaux de saisie et évaluations théoriques. IHM 2005.
59. Ware, C., Multimedia output devices and techniques, Journal ACM Computing Surveys, 1996.
60. Wobbrock, J. O., Myers, B. A., Kembel, J. A., EdgeWrite: A Stylus-Based Text Entry Method Designed for High Accuracy and Stability of Motion, UIST 2003.
61. Wobbrock, J., Myers, B.A., Rothrock, B., Few-key Text Entry Revisited: Mnemonic Gestures on Four Keys, CHI06.
62. Ziat, M., Conception et implémentation d'une fonction zoom haptique sur PDAs, thèse, Université de technologie de Compiègne, 2006.
63. <http://www.realeyes3d.com/pages/home/home.php>
64. <http://www.tellme.com>
65. <http://www.myvu.com>