

# Un espace de caractérisation de la télécommande dans le contexte de la télévision interactive

Dong-Bach Vo<sup>1</sup>, Gilles Bailly<sup>2</sup>, Eric Lecolinet<sup>1</sup>, Yves Guiard<sup>1</sup>

<sup>1</sup>Telecom Paristech  
46, rue Barrault  
75013 Paris  
{prénom.nom}@telecom-paristech.fr

<sup>2</sup>Deutsche Telekom Laboratories, TU Berlin  
Ernst-Reuter-Platz 7,  
D-10587 Berlin  
gilles.bailly@telekom.de

## RESUME

Vestige du simple instrument de zapping qu'elle a été dans les années 1970, la télécommande traditionnelle de téléviseur est désormais inadaptée aux exigences multiformes de la télévision interactive. La conception de nouvelles télécommandes commence à mobiliser la recherche en IHM. L'article indique les diverses directions qui sont aujourd'hui explorées. Après nous être employés à caractériser le contexte spécifique de la télévision interactive, nous proposons un tableau de l'espace de conception et d'évaluation auquel sont confrontés les chercheurs. Malgré les améliorations et les augmentations dont elle peut certainement encore bénéficier, la télécommande traditionnelle ne survivra probablement pas longtemps à la crise qu'elle connaît depuis l'avènement du numérique, face à la concurrence de divers dispositifs procédant de logiques d'interfaçage entièrement renouvelées, dont l'article fait état.

## Mots clés

Télévision interactive, télécommande, technique d'interaction.

## ABSTRACT

Initially designed in the nineteen seventies as a mere zapping tool, obviously the traditional device known as the TV remote control no longer responds to the multifarious needs of today's interactive television. Designing new remote-control devices is a challenge that the HCI community has started to face. The paper indicates the various directions that are currently being investigated by researchers. The paper starts with an attempt to characterize the specific context of interactive TV. We then offer a tentative account of the design and evaluation space of interest. There is little doubt that the traditional remote control may still be improved and augmented. It is unlikely, however, that it will survive for long the crisis it has been undergoing since the beginning of the digital era, given the emerging plethora of alternative interaction possibilities based on a variety of new interfacing logics, which the paper reviews.

## Categories and Subject Descriptors

H5.2. [Information Interfaces And Presentation]: User Interface.  
I.3.6. [Methodology and Techniques]: Interaction techniques.

## General Terms

Design, Human Factors.

## Keywords

Interactive television, remote control, interaction technique.

## 1. INTRODUCTION

Depuis plus de 50 ans, la télévision (TV) n'a cessé d'évoluer vers toujours plus d'interactivité. La TV numérique passe aujourd'hui par un boîtier décodeur offrant l'accès à de multiples contenus multimédia et fonctionnalités (e.g. vidéo à la demande, guides de programmes (EPG), possibilité de réagir à certaines émissions ou encore d'accéder à des jeux en lignes ou à des informations issues de l'Internet via des widgets [29]). De son côté, la télécommande, dispositif central pour le contrôle de ces nombreux contenus et fonctionnalités, a relativement peu évolué [15]. Fidèle au paradigme « un bouton, une fonctionnalité », la télécommande est jugée malcommode [7].

Les enjeux sociétaux de la TV sont considérables : par exemple, 98% des ménages américains étaient équipés d'une TV en 2009<sup>1</sup>. Alors même qu'on assiste à un développement soutenu de la TV interactive, la question des techniques d'interaction que soulève inmanquablement ce nouveau type de TV a récemment suscité un intérêt relativement important dans la recherche. Nous proposons dans ce qui suit un espace de caractérisation visant à guider, dans le contexte spécifique de la TV interactive, la conception et l'évaluation des techniques d'interaction liées à la télécommande, ou visant à la remplacer. Nous considérerons les enjeux et les problématiques spécifiques de la TV interactive, le cœur de notre sujet, ensuite nous décrirons et comparerons les solutions proposées, tout en suggérant quelques directions de recherche.

## 2. LA TELEVISION INTERACTIVE

La TV interactive associe un ensemble de dispositifs pour l'information et le divertissement, et un ensemble de services interactifs [6], [15]. Elle intervient dans divers contextes. Tout

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IHM'11, October 24-27, 2011, Sophia Antipolis, France

Copyright © 2011 ACM 978-1-4503-0822-9/11/10 ...\$10.00.

<sup>1</sup> Nielsen Company, Television Audience 2009 (2010).

d'abord, on peut regarder la TV seul ou à plusieurs [12]. Ensuite, l'attention accordée à la TV, souvent partagée entre plusieurs activités concurrentes, est fort variable [7], [12]. Enfin, les activités liées à la TV sont associées au divertissement, un contexte sensiblement différent d'un contexte de travail. Nous reviendrons sur ces spécificités.

## 2.1 Des services interactifs

De nombreux genres, formats et contenus enrichissent l'interaction entre l'utilisateur et la TV [29] :

- les *guides de programmes électroniques* (EPG) [5], [29] permettent d'explorer la liste des programmes disponibles. Un EPG peut comprendre un moteur de recherche interactif, des fonctionnalités de rappel de diffusion ou d'enregistrement de programmes, ou encore des outils de personnalisation.
- les *programmes augmentés* permettent d'accéder à des informations supplémentaires lors du visionnage, comme par exemple l'affichage de statistiques durant la diffusion d'un événement sportif.
- la *vidéo à la demande*, présentée sous forme de catalogue, permet de choisir du contenu multimédia et de le visionner au moment de son choix.
- la *TV personnalisée* fournit un contrôle plus précis sur le contenu diffusé, comme le contrôle parental ou la possibilité de visionnage différé, de rejouer le contenu ou de s'y déplacer temporellement.
- les services de *e-commerce* permettent d'acheter directement un produit vu à la TV.
- des *jeux, quizz*, services de *vote interactif*, *chats* et *forums* peuvent compléter l'émission.
- enfin, divers services Internet peuvent être rendus accessibles directement sur le téléviseur, notamment à l'aide de *widgets* (météo, informations, alertes, messagerie, vidéos...).

Cette prolifération et cette complexification des services interactifs bouleversent le vieux concept de TV. Le rôle du téléspectateur devient moins passif, pouvant choisir parmi un large éventail de contenus et interagir avec les programmes ou d'autres utilisateurs. De nouvelles formes d'interaction [9] sont alors nécessaires comme la saisie de texte [27], ou encore le pointage d'éléments à l'écran [28], [37], [48].

## 2.2 Une activité à différents degrés d'engagement

Le téléspectateur peut naturellement mener d'autres activités en parallèle. Berglund et al. [5] identifient quatre types d'usage de la TV en fonction du niveau d'engagement : activité d'arrière plan, d'accompagnement, partielle ou principale. Gueddana et Roussel [22] définissent le niveau d'engagement comme « la limite dans laquelle une personne est prête à s'exposer et à être sollicitée par d'autres ». Lorsque certaines modalités sensorielles sont requises dans une autre tâche, elles sont plus ou moins indisponibles pour la TV. Ce facteur de niveau d'engagement est un facteur critique pour la conception des applications pour la TV [13]. Par exemple, dans [13] une application de diffusion de clips musicaux, le flux de la diffusion vidéo est continu, sauf si le bouton « chanson suivante » est pressé. Cette application transforme de fait le téléviseur en radio, permettant à l'utilisateur de se consacrer à une autre activité.

## 2.3 Une activité sociale

L'activité télévisuelle est souvent partagée entre amis ou en famille. L'utilisation de services interactifs peut alors s'avérer problématique si l'accès à un service accapare une partie ou la totalité de la TV, ceci au détriment des autres utilisateurs. La prise de contrôle du dispositif peut également être source de tension, Meerbek et al. [41] suggérant que cette question du contrôle pèse sur la satisfaction des utilisateurs. De plus, des problèmes de vie privée peuvent surgir si l'échange d'informations entre un utilisateur et le service interactif est visible par les autres personnes [11]. L'utilisation d'un dispositif auxiliaire [11], [16], [50] comme un téléphone portable ou une tablette pourrait constituer une solution en offrant un espace d'interaction privé opaque aux autres utilisateurs et non susceptible de les gêner [16], [50].

Certains travaux de recherche prennent en considération le degré d'engagement des utilisateurs et l'activité sociale dans la conception des systèmes. Pêle-Mêle [22] est un système de communication vidéo multi-sites qui utilise la détection de présence, de mouvement et de visage pour classifier en permanence le niveau d'activité des utilisateurs selon trois niveaux, absent, disponible, engagé. Ces niveaux impactent le contenu : les images des lieux de niveau "engagé" apparaissent au centre de l'écran, les autres étant relégués vers la périphérie. Ils impactent aussi la fréquence de mise à jour des images : si le niveau d'engagement est "absent", l'image est figée et dégradée. La transition et combinaison avec d'autres services est aussi influencée : des animations permettent de faciliter la perception et la compréhension des transitions entre les différents niveaux d'engagement.

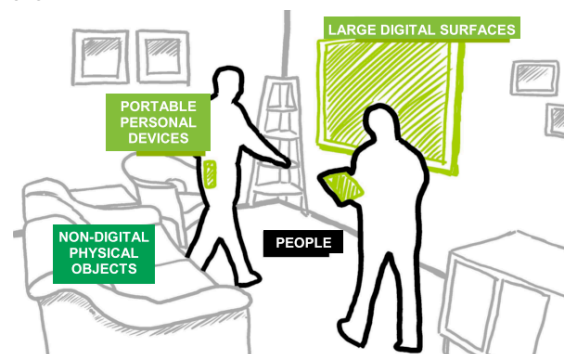


Figure 1 : Un environnement proxémique [3]

Ballendat et al. [3] ont introduit la notion d'*interaction proxémique* : selon le lieu, la distance, l'orientation, le mouvement des utilisateurs par rapport au dispositif, et selon l'identité de ce dispositif, le système offre différentes options (Figure 1). Par exemple, si l'utilisateur s'approche du téléviseur, le système adapte la visualisation afin de faciliter l'exploration des vidéos disponibles. S'il s'éloigne pour s'asseoir dans son canapé, la vidéo sélectionnée passe en plein écran et sa diffusion commence. Inversement, la vidéo s'arrête si l'utilisateur reçoit un appel téléphonique ou ouvre un magazine. Si un autre participant s'approche de l'écran, des informations sont affichées à son attention mais l'écran est partagé de manière à ne pas interrompre le flux intéressant le participant assis. Le système met en pause la vidéo si les participants rentrent en discussion et se met en veille s'ils quittent la pièce. D'autres fonctionnalités permettent aussi de détecter si l'utilisateur pointe l'écran avec un objet (pour interagir avec), s'il s'approche un téléphone mobile près de l'écran (pour

transférer des contenus multimédia), etc. D'autres exemples de tels systèmes sont consultables dans [23].

## 2.4 Une activité de divertissement

L'objet de l'activité télévisuelle est de se divertir, de s'informer et de se détendre [7], [14], affalé dans un canapé plutôt qu'assis devant un écran d'ordinateur [7], [51]. Bernhaupt et al. [7] rapportent que les téléspectateurs ne sont pas disposés à utiliser clavier et souris, et qu'ils sont même réticents à l'idée d'une interface graphique exploitant la métaphore du bureau. Les modèles qui ont fait le succès de l'ordinateur personnel classique semblent donc peu adaptés au contexte de la TV interactive.

Il y a lieu dès lors de s'interroger sur la pertinence, par rapport à notre sujet, des méthodes d'évaluation traditionnelles en IHM qui reposent sur des mesures de performance—e.g., un temps d'exécution de tâche que l'on demande de minimiser [14], [17]. Les critères d'évaluation des techniques d'interaction pour les applications productives ne suffisent pas pour celles en TV [14]. Druker et al. [17] rapportent que bien que les performances des utilisateurs avec leur système de navigation vidéo ne soient pas les meilleures, les utilisateurs jugent leur système plus amusant, moins frustrant et aussi plus facile à apprendre et à utiliser que les deux autres systèmes comparés. Les auteurs suggèrent que les préférences des utilisateurs et leurs performances dépendent du niveau d'engagement face à la TV. Chorianopoulos et Spinellis [14] proposent une méthodologie d'évaluation basée sur les émotions et prenant en compte les caractéristiques particulières de la TV, des téléspectateurs et du contexte d'usage. Outre l'examen traditionnel de l'efficacité des interfaces, ils proposent de réutiliser des méthodes d'évaluation déjà éprouvées dans d'autres domaines, comme l'étude des média et de la communication de masse.

## 3. LA TELECOMMANDE

La télécommande est un dispositif d'entrée qui permet de contrôler les divers services proposés par la TV interactive et de manipuler les données multimédia mises à disposition de l'utilisateur. Comme précédemment mentionné, la télécommande a beaucoup moins évolué que les services offerts à l'utilisateur [15].



Figure 2 : Télécommande et interface usuelles

La littérature fait état de nombreuses tentatives pour améliorer la télécommande, en partant de sa configuration actuelle. En s'appuyant sur les restrictions de la télécommande traditionnelle pour contrôler le contenu multimédia, Cesar et al. [10] classent les techniques d'interaction TV en trois catégories : l'augmentation de la télécommande traditionnelle, l'utilisation opportuniste des mobiles personnels (smartphones, tablettes tactiles) et la réutilisation des objets du quotidien. Ces auteurs indiquent des pistes notamment pour faciliter l'utilisation des télécommandes

pour sélectionner, naviguer et manipuler les contenus multimédia, pour enrichir l'interaction avec la télévision et prendre en compte l'environnement dans lequel se déroule l'activité. Contrairement à la réutilisation des objets du quotidien qui tiennent compte de la fonction principale de l'objet, des travaux offrent de nouvelles télécommandes qui reposent sur la capacité des utilisateurs à manipuler directement des objets physiques abstraits et sur la représentation des données associées qui pourrait changer la façon d'interagir avec la télévision. Enfin, l'indispensabilité de commandes à distance pour contrôler la télévision est de mise. Certains travaux s'intéressent aussi au cas où, la télécommande étant inaccessible ou indisponible, des dispositifs alternatifs permettent d'interagir quand même avec la télévision. La prise en compte de ces deux nouvelles voies de conception nous impose d'ajouter deux catégories : la dématérialisation de la télécommande et la télécommande tangible.

### 3.1 La télécommande traditionnelle

La télécommande traditionnelle (Figure 2) est composée de touches dédiées à la navigation, aux fonctionnalités courantes, ou encore à la saisie de données. Elle tend à être surchargée de boutons [15], [45] à cause de l'augmentation et de la richesse des services interactifs. En plus d'accaparer l'attention visuelle normalement dédiée à la TV, sa difficulté d'utilisation [7] est un frein à l'adoption des nouvelles fonctionnalités par les utilisateurs [45]. Une alternative offerte par les télécommandes classiques consiste à multiplier les modes (un par appareil physique) et les menus dans lesquels il faut naviguer au moyen de nombreux appuis successifs sur des touches directionnelles [6], [14]. Contrairement à l'interaction classique avec une souris et un clavier, la navigation dans des menus hiérarchiques est plus difficile avec une télécommande [14]. Mirlacher et al. [40] montrent que six touches sur la télécommande (4 touches directionnelles, plus 1 pour valider et 1 pour annuler) est un bon compromis entre simplicité et efficacité pour naviguer dans les menus hiérarchiques.

Certains attributs des boutons, comme leur couleur ou leur forme, peuvent être mis à profit pour réaliser des actions soit globales soit relatives au mode ou à la sélection courante [15]. Une telle ressource est toutefois d'une portée limitée et en pratique n'excédant pas quatre couleurs (Figure 2), à plus forte raison si l'on garde à l'esprit que la TV se regarde généralement à faible niveau de luminosité ambiante. Les distinctions de forme évitent ce problème et ont l'avantage d'être reconnaissables par un simple toucher, ceci permettant d'activer des fonctionnalités sans avoir à déplacer le regard de la TV vers la télécommande. Enfin, les touches numériques permettent d'accéder directement aux chaînes (si tant est que l'utilisateur se remémore le nombre qui leur est associé) ou parfois à naviguer dans les menus hiérarchiques (chaque item étant associé à un chiffre).

L'entrée de texte constitue un autre problème [39] car le facteur de forme interdit généralement de disposer d'un clavier classique. De même que pour les téléphones mobiles, le Multitap [15] peut permettre d'entrer du texte si la télécommande dispose d'un clavier numérique. Une alternative consiste à afficher un clavier virtuel sur l'écran de la TV. Plusieurs techniques et configuration de claviers sont envisageables. Par exemple avec *The Numeric-keypad Typer* [27], l'alphabet présenté sur l'écran est divisé en 9 zones disposées matriciellement et identiquement aux touches numériques sur la télécommande (Figure 3). De la même manière, chaque zone contient neuf lettres ou symboles. Pour sélectionner un caractère, il suffit d'appuyer deux fois sur une touche

|   |   |   |
|---|---|---|
| 1 | 2 | 3 |
| 4 | 5 | 6 |
| 7 | 8 | 9 |

Marshall et al. [39] comparent six autres techniques et configurations de saisie de texte avec la télécommande. Notons que les auteurs affirment que la meilleure solution pour saisir des données alphanumériques reste le clavier physique classique. C'est sans doute pourquoi certaines télécommandes incluent un clavier complet que l'on accède en le dépliant ou en retournant la télécommande comme dans [6] ou chez Boxee<sup>2</sup>.

Il faut enfin noter que la transmission infrarouge (IR) utilisée par la plupart des télécommandes traditionnelles implique de fortes limitations. Tout d'abord la fréquence d'émission des signaux IR est incompatible avec du contrôle continu. De plus, afin d'économiser l'énergie, l'appui sur une touche ne génère le plus souvent l'émission que d'un seul signal IR (à l'appui, mais pas au relâchement ni pendant que la touche est appuyée). Enfin, l'infrarouge implique de viser le récepteur avec la télécommande et qu'aucun objet n'empêche la lumière de l'atteindre. Ceci engendre nombre d'erreurs de réception et nécessite parfois de tenir la télécommande d'une manière peu commode.

### 3.2 Augmentation de la télécommande traditionnelle

### 3.2.1 Gestes 2D

<sup>2</sup> <http://www.boxee.tv/buy>

|   |   |    |   |   |   |   |   |    |
|---|---|----|---|---|---|---|---|----|
| a | b | c  | j | k | l | s | t | u  |
| d | e | f  | m | n | o | v | w | x  |
| g | h | i  | p | q | r | y | z | ä  |
| ä | ö | ½  | / | ( | ) | 1 | 2 | 3  |
| ! | ° | #  |   |   | ? | 4 | 5 | 6  |
| π | % | &  | < | > | , | 7 | 8 | 9  |
| = | - | \$ |   | \ | ~ | : | 0 |    |
| @ | £ | §  | ^ | * |   |   |   |    |
|   |   |    | µ | : |   |   |   | SH |

### 3.2.2 Gestes 3D

Gurrin et al. [24] utilisent des gestes 3D pour naviguer dans une collection de souvenirs numériques. Ces gestes permettent d'aller d'évènement en évènement ou de jour en jour. GesText [30] permet d'entrer du texte en utilisant les accéléromètres d'une Wiimote tandis que Aoki et al. [1] combinent l'appui de touches directionnelles avec des rotations du poignet. Afin de guider l'utilisateur, la forme de la représentation graphique du clavier suggère les gestes possibles pour saisir un caractère dans cette étude. Par ailleurs, le pointage dans l'air via un dispositif gyroscopique est comparé à d'autres techniques dans [28] (touches directionnelles ou trackball) et [37] (souris ou d'un joystick). Le pointage en l'air est la technique la plus rapide dans les deux cas mais pas toujours la plus précise.

L'utilisation du langage naturel a été proposée pour la navigation parmi les chaînes de TV [2], [26] et l'utilisation d'un guide électronique des programmes [6]. La modalité vocale, et plus spécifiquement la reconnaissance du langage naturel, permet en un seul énoncé d'exécuter une commande et de donner les paramètres qui y sont associés, comme enregistrer un programme sur une chaîne particulière à un horaire spécifique [6]. Une des ambitions des commandes vocales est de faciliter l'utilisation des systèmes par des utilisateurs novices. Elles permettent aussi de créer des raccourcis afin de réduire le nombre d'actions pour réaliser les tâches fréquentes. Par exemple dans [26] les utilisateurs énoncent les noms des chaînes au lieu d'avoir à naviguer dans des listes.

Des moteurs de vibration de petite taille peuvent être employés pour fournir une forme de retour haptique compatible avec les contraintes d'une télécommande (poids, taille et consommation en énergie limités). Le retour tactile est par exemple utilisé dans [43], [44] pour fournir des informations à l'utilisateur ainsi que dans [49] où une cellule Braille permet de faciliter la sélection des chaînes favorites.

Le dispositifs mobiles comme les "smartphones" et les tablettes tactiles, dont le nombre ne cesse de croître, ont le double avantage de permettre un accès facile aux infrastructures réseau et de disposer d'une surface d'affichage suffisamment grande pour véhiculer une quantité d'information importante. Ils sont de plus généralement équipés d'un écran tactile et de divers capteurs comme des accéléromètres.

<sup>4</sup> <http://www.gyration.com/index.php/fr/home.html>

### 3.3.1 Dispositif d'interaction

Le dispositif mobile peut jouer le même rôle qu'une télécommande augmentée dans la mesure où il peut être équipé de capteurs. Par exemple le système multimodal de Turunen et al. [53] permet de naviguer ou zoomer en effectuant des gestes avec le téléphone ou d'activer des commandes par la voix ou par contact du téléphone sur des objets comportant des étiquettes RFID. Le retour tactile est utilisé pour avertir l'utilisateur du bon déroulement, ou non, des opérations (sous forme d'icônes haptiques associées à des comportements vibratoires différents).

### 3.3.2 "Companion device"

De par la taille de leur écran ces dispositifs peuvent aussi servir d'écran auxiliaire [11], [16], [46] ceci permettant d'éviter l'apparition d'informations sur l'écran de la TV propres à gêner les autres utilisateurs. Cette solution permet de contrôler, de partager, de commenter ou d'enrichir le contenu diffusé, particulièrement dans le cas des tablettes dont la taille facilite l'entrée de texte. Cette solution permet aussi de personnaliser l'interaction avec le système [50] en adaptant l'interface en fonction de l'utilisateur (par exemple pour contrôler l'accès aux contenus [50] ou pour personnaliser les contenus favoris [16]). Une architecture pour réaliser un tel système est détaillée dans [11].

Deux cas de figure peuvent être considérés suivant que le dispositif mobile est personnel (typiquement un téléphone) ou partagé par plusieurs personnes (typiquement une tablette se trouvant habituellement à proximité de la télévision). Le premier cas correspond plutôt à un usage en complément d'une télécommande dédiée, surtout s'il y a des enfants dans le foyer ou souvent des personnes de passages susceptible d'utiliser la TV (amis, famille, baby sitter...). On notera à ce propos qu'il peut être gênant de prêter un dispositif personnel à une autre personne du fait des questions de privauté. Par ailleurs, le facteur de forme des smartphones n'est pas très approprié : étant petits et plats car conçus pour tenir dans une poche ils risquent de tomber entre les coussins d'un canapé. Les tablettes tactiles à vocation domotique semblent donc mieux adaptées à ce type d'usage.

## 3.4 Réutilisation des objets du quotidien

La réutilisation des objets du quotidien permet de tirer avantage de la familiarité avec ces objets pour faciliter l'apprentissage de nouvelles techniques ou technologies interactives. Basé sur une étude des routines quotidiennes au domicile, Paper Remote [5] combine un stylo numérique [38] et un guide papier des programmes télévisés pour spécifier les programmes à visionner ou à enregistrer. Ce système introduit donc une nouvelle technologie par le biais d'un artefact lié à la TV. L'interaction avec le papier a été aussi investiguée dans [25] pour personnaliser la télécommande pour contrôler un environnement multimédia et dans [36] dans le cadre de l'éducation électronique. Via un lecteur et des codes barres imprimés sur le cours, l'étudiant peut jouer des vidéos liées aux documents papier sur la TV.

Certains travaux s'intéressent davantage aux relations avec les objets et à l'expérience utilisateur [34], [42], [47], [54]. Dans [34], lancer un ballon en mousse contre la TV permet de modifier le contenu diffusé par celle-ci en fonction de l'état émotionnel. Nack et al. [42] envisagent d'utiliser un coussin pour contrôler la TV, en particulier lorsque la télécommande est hors d'atteinte. De plus, la relation intime entre le téléspectateur et le coussin permet de surveiller ses données biométriques. L'analyse de ces données peut permettre de suggérer des activités à l'utilisateur tout en

personnalisant l'interface sur la surface du coussin pour cette activité. Dans le même esprit ludique, Schmidt et al. [47] proposent d'utiliser un pouf comme télécommande. Enfin, la table basse du salon peut aussi être utilisée comme surface de contrôle [54]. Celle-ci a de plus le mérite d'être partageable par tous les utilisateurs situés autour de la table, lesquels interagissent au moyen de gestes.

## 3.5 Dématérialisation de la télécommande

La télécommande n'est pas toujours à portée de main [20], l'utilisateur pouvant être plus ou moins éloigné du canapé ou engagé dans une autre tâche. Elle peut de plus avoir été égarée [20], [50]. Enfin certains utilisateurs peuvent avoir du mal à utiliser les télécommandes traditionnelles, notamment s'ils sont mal voyants [6]. Ces raisons motivent la conception de techniques alternatives [12].

### 3.5.1 Langage naturel

Comme pour le cas des télécommandes augmentées, la reconnaissance du langage naturel constitue une première possibilité [52], [21]. Dans [21] l'utilisateur dialogue avec un agent conversationnel, représenté par une peluche, afin de contrôler les fonctionnalités basiques de la TV et d'obtenir des informations sur les programmes. L'agent répond à chaque requête et suggère des opérations en fonction du contexte, comme enregistrer un programme si la diffusion est prévue ultérieurement. Les auteurs soulignent l'imperfection de la reconnaissance vocale. De plus, le remplacement de la télécommande traditionnelle par la reconnaissance vocale semble difficile [6], [26] lorsque le bruit ambiant est trop important, lorsque la modalité vocale est engagée dans une autre activité, ou encore si la requête est difficile à exprimer oralement. Tseng et al. [52] rapportent la difficulté des utilisateurs à apprendre à parler naturellement avec TRIBA, un agent en forme de sphère.

### 3.5.2 Gestes 3D

Une autre alternative consiste à utiliser les gestes dans l'air, généralement en utilisant des techniques basées sur la vision [20], [35], [56]. Bien que des produits commerciaux récents comme Microsoft Kinect<sup>5</sup> rencontrent aujourd'hui un grand succès, cette idée n'est pas nouvelle [4], notamment dans le domaine de la TV où Freeman et al. [20] ont proposé un système de contrôle basé sur l'analyse des gestes de la main. Filmés par une caméra, les gestes sont interprétés par analyse vidéo en temps réel. La main agit comme une souris, avec un geste de fermeture de la main en guise de clic. Les auteurs rapportent l'excitation des utilisateurs mais aussi, la fatigue physique liée à l'utilisation intensive du système.

Tenant de restaurer le potentiel intégral de la main pour l'interaction gestuelle, g-stalt [56] permet de manipuler un environnement 3D et d'interagir avec les médias vidéos qui s'y trouvent grâce à un vocabulaire de geste complexe capturé par un système Vicon. Actual Remote Control [35], porté sur le poignet comme une montre, est capable de reconnaître les directions cardinales des gestes de la main et les rotations horaires et anti-horaires du poignet grâce à un accéléromètre et un gyroscope 3 axes. Le système permet de naviguer dans un menu et de sélectionner des commandes de nombreux appareils domestiques dont la TV via des gestes simples.

<sup>5</sup><http://www.xbox.com/en-US/Xbox360/Accessories/Kinect/Home>

### 3.6 Les télécommandes tangibles

Certains travaux de recherche s'intéressent aussi à la forme des télécommandes [8] [19], [32], [49], [52], [55]. L'idée de concevoir des télécommandes de formes élémentaires permet de tirer partie de la faculté de la main à saisir et manipuler les objets [19]. Par ailleurs, l'adéquation entre la forme de la télécommande et l'interface graphique sur la TV facilite la prise en main des techniques d'interaction.

#### 3.6.1 Formes

L'utilisation de formes simples a été explorée dans [8], [19], [31], [52]. Un cube peut par exemple être posé sur n'importe quelle surface horizontale et être facilement manipulé : les opérations pour ramasser, poser, pivoter, translater un cube sont intuitives [8], [19]. Basée sur ce principe, [8] permet de sélectionner et visualiser un ensemble de chaînes à l'écran tandis que [19] permet aussi de contrôler le volume et de naviguer à travers un menu. Le vocabulaire gestuel décrit par les auteurs inclut les rotations du cube, secouer le cube, ou poser ce dernier dans une certaine position pour mettre le système télévisuel en veille. La combinaison d'un appui sur un bouton situé sur le cube et des différentes rotations permet d'augmenter sensiblement les possibilités. De forme cylindrique la télécommande de Kimman et al. [31] permet d'atteindre facilement avec le pouce les quatre anneaux disposés verticalement sur sa surface (Figure 5). L'apparence sphérique de TRIBA [52] suggère qu'on la manipule au moyen de rotations. La forme plane et large de Rebo [32] invite l'utilisateur à interagir avec des gestes linéaires sur sa surface. Cependant, la taille de Rebo nécessite d'être posé sur une surface stable et ne permet pas la manipulation avec une seule main.

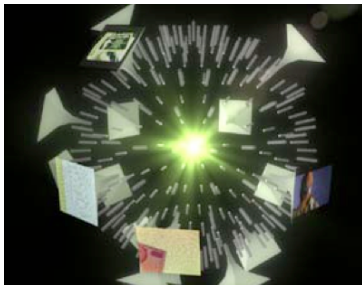


Figure 4 : Interface de TRIBA [52]

Des télécommandes de forme plus complexe comme *Peppermill* [55] entraînent diverses affordances. La partie plane de *Peppermill* et la présence de boutons invitent l'utilisateur à tenir la télécommande alors que sa partie cylindrique suggère de visser ou dévisser pour sélectionner un élément dans une liste tout en rechargeant la télécommande en énergie. Enfin, la forme de *Tactimote* [49], proche d'un pistolet, permet de contrôler un joystick avec le pouce (à la place du chien) pour naviguer dans une représentation matricielle des chaînes TV. La gâchette permet de valider avec le majeur tandis que l'index repose sur une cellule Braille permettant de véhiculer des informations tactiles destinées à faciliter la sélection des favoris.

#### 3.6.2 Couplage

L'association entre la forme de la télécommande et la représentation logicielle à l'écran peut créer une connexion évidente entre le dispositif physique et l'interface numérique [52]. TRIBA [52] permet de naviguer à travers une collection de contenus multimédias représentée sous forme d'une composition sphérique (Figure 4). Chaque chaîne est représentée par une

épingle pouvant se dilater pour afficher un aperçu du contenu. Cette solution permet la visibilité de toutes les options offertes, même si l'intérêt est seulement porté sur un groupe particulier de chaîne. De même, [8] est associée à une représentation cubique de telle sorte que lorsque le cube physique est soumis à une rotation, le cube virtuel restitue cette rotation sur la TV et permette de visionner l'aperçu de trois chaînes différentes simultanément sur les faces visibles. Si le cube est posé et immobile, la chaîne sur la face la plus visible du cube virtuel s'élargit pour couvrir tout l'écran. Visser ou dévisser *Perpermill* [55] permet d'actionner un carrousel virtuel présentant les différents contenus multimédia. Enfin, dans [31] la représentation visuelle se présente comme quatre menus, chacun étant associé à chaque anneau de la télécommande et permettant d'interagir sans avoir à détourner le regard de l'écran de TV (Figure 5).

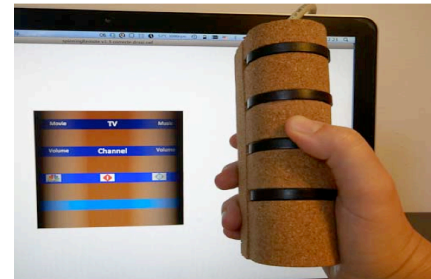


Figure 5 : Spinning remote [31]

L'association peut aussi relier les gestes et la représentation graphique. Les gestes réalisés sur Rebo [32] permettent d'engager une commande et de la valider. Par exemple, pour changer de chaîne, l'utilisateur effectue un trait horizontal sur la télécommande. Au fur et à mesure qu'il réalise son geste, un aperçu de la chaîne suivante apparaît et pousse la chaîne actuelle vers l'extérieur de l'écran. Si le geste s'arrête, la commande avorte et la chaîne n'est pas modifiée. Ce système d'exécution partielle, appelé *action slopping* par ses auteurs, permet de prendre conscience des fonctionnalités existantes et introduit une certaine tolérance à l'erreur, indisponible avec une télécommande traditionnelle. D'autres exemples sont présentés dans [32].

## 4. DISCUSSION

Les travaux évoqués ci-dessus visent tous à faciliter, d'une manière ou d'une autre, la sélection, la navigation et/ou la manipulation des contenus multimédias. Nous avons vu que la télécommande tangible introduit de nouvelles formes d'exploration et de représentation des données qui pourraient changer le paradigme actuel d'interaction pour la télévision.

Il reste qu'au stade où nous en sommes du développement de la TV interactive, les solutions qui font appel à l'augmentation de la télécommande traditionnelle paraissent jouir d'un avantage non négligeable sur leurs concurrentes : augmentée de manière adéquate, la télécommande devrait en principe permettre de répondre aux besoins nouveaux de la télévision interactive sans pour autant sacrifier un instrument hautement familier, et dont la conception intègre malgré tout plusieurs décennies d'expérience industrielle à grande échelle. Certes, l'idée presque évidente que la télécommande traditionnelle est dépassée motive de nombreux travaux plus ou moins radicalement innovants. Mais ce sont divers problèmes particuliers—comme l'accès aux commandes les plus courantes, la navigation dans les listes, la saisie de texte ou l'utilisation d'une application particulière—que l'on s'est pour l'essentiel efforcé de résoudre jusqu'à ce jour. Il est clair, à en

juger par l'état de l'art, qu'on ne dispose pas actuellement une réponse d'ensemble, coordonnée et viable, à la problématique de la commande à distance telle qu'elle se pose désormais dans le contexte de la télévision interactive multimédia. Avant de pouvoir concurrencer véritablement la télécommande augmentée, la plupart des techniques d'interaction qui ont été décrites dans la littérature IHM vont sans doute devoir progresser davantage vers l'intégration.

Une partie des travaux référencés s'intéressent à l'enrichissement de l'expérience télévisuelle. La réutilisation des dispositifs mobiles permet notamment d'échanger des données avec le décodeur (par exemple, apporter de l'information complémentaire sur les programmes ou encore transférer des médias pour une utilisation ultérieure [11]). L'interaction entre les deux médias ouvre de nouvelles pistes de recherche tant en terme d'usage que d'interaction. Cela complique un peu plus la problématique faisant l'objet de cet article, notamment par les différentes situations d'usage envisageables.

Il est important de rappeler que, dans leur majorité, les travaux de la littérature laissent de côté la variabilité du contexte d'usage de la télévision, de nouvelles techniques d'interaction étant typiquement proposées pour des situations à un seul téléspectateur. Or, des contraintes spéciales apparaissent lorsque la télévision est regardée entre amis ou en famille—par exemple, la voix peut devenir tout simplement inapproprié comme moyen de contrôle de la télévision, tout comme le geste dans l'espace 3D.

## 5. CONCLUSION

Nous avons proposé un espace de caractérisation des télécommandes pour la TV interactive utile pour guider la conception et l'évaluation de nouvelles techniques d'interaction dans ce contexte d'usage. Nous avons d'abord présenté les spécificités de la TV interactive, puis nous avons organisé les différents travaux sur les techniques d'interaction avec des télécommandes. Nous prévoyons de nous appuyer sur le concept d'interaction proxémique [3] pour proposer de nouvelles techniques d'interaction réactives au contexte d'usage en adéquation avec les besoins de l'utilisateur. Nous envisageons également d'approfondir des critères qui ont reçu peu d'attention comme le niveau d'engagement de l'utilisateur dans l'activité [17] et la réponse émotionnelle aux techniques d'interaction [14], qui semblent essentiels pour l'évaluation et l'acceptation des techniques d'interaction dans le contexte de la TV interactive.

Enfin, toutes les télécommandes référencées dans cet article ainsi que des prototypes issus de l'industrie sont présentés sur le site internet <http://www.gillesbailly.fr/remote>.

## 6. REMERCIEMENTS

Le travail présenté dans cet article a été financé par le programme de recherche européen QUAERO.

## 7. REFERENCES

- [1] Aoki, R., Maeda, A., Watanabe, T., Kobayashi, M., and Abe, M. Twist&tap: text entry for television remotes using easy-to-learn wrist motion and key operation. *IEEE Trans. on Consumer Electronics*, 56, 1 (2010).
- [2] Balchandran, R., Epstein, M.E., Potamianos, G., and Seredi, L. A multi-modal spoken dialog system for interactive television. *Proc. of IMCI '08* (2008).
- [3] Ballendat, T., Marquardt, N., and Greenberg, S. Proxemic interaction. *Proc. of ITS '10* (2010).
- [4] Baudel, T. and Beaudouin-Lafon, M. Charade: remote control of objects using free-hand gestures. *ACM CACM*, 36, 7 (1993).
- [5] Berglund, A., Berglund, E., Larsson, A., and Bang, M. Paper remote: an augmented television guide and remote control. *Univ. Access in the Info. Society* 4, 4 (2005).
- [6] Berglund, A. and Johansson, P. Using speech and dialogue for interactive television navigation. *Univ. Access in the Info. Society*, 3, 3-4 (2004).
- [7] Bernhaupt, R., Obrist, M., Weiss, A., Beck, E., and Tscheligi, M. Trends in the living room and beyond: results from ethnographic studies using creative and playful probing. *ACM CIE*, 6, 1 (2008).
- [8] Block, F., Schmidt, A., Villar, N., and Gellersen, H.W. Towards a playful user interface for home entertainment systems. *Ambient Intelligence*, Springer (2004).
- [9] Cesar, P. and Chorianopoulos, K. The Evolution of TV Systems, Content, and Users Toward Interactivity. *Foundations and Trends in Human-Computer Interaction*, 2, 4 (2007).
- [10] Cesar, P., Chorianopoulos, K., and Jensen, J.F. Social television and user interaction. *ACM CIE*, 6, 1 (2008).
- [11] Cesar, P., Bulterman, D.C. a, and Jansen, J. Leveraging user impact: an architecture for secondary screens usage in interactive television. *Multimedia Systems*, 15, 3 (2009).
- [12] Chorianopoulos, K. Interactive television design that blends seamlessly with everyday life. *Universal Access in Ambient Intelligence Environments*, (2007).
- [13] Chorianopoulos, K. User Interface Design Principles for Interactive Television Applications. *Int. Journal of Human-Computer Interaction*, 24, 6 (2008).
- [14] Chorianopoulos, K. and Spinellis, D. User interface evaluation of interactive television: a media studies perspective. *Universal Access in the Information Society*, 5, 2 (2006).
- [15] Cooper, W. The interactive television user experience so far. *Proc. of UXTV '08* (2008).
- [16] Cruickshank, L., Tsekleves, E., Whitham, R., Hill, A., and Kondo, K. Making Interactive Television Easier to Use: Interface Design for a Second Screen Approach. *The Design Journal*, 10, 3 (2007).
- [17] Drucker, S.M., Glatzer, A., Mar, S.D., and Wong, C. SmartSkip: consumer level browsing and skipping of digital video content. *Proc. of CHI '02* (2002).
- [18] Enns, N. and MacKenzie, I.S. Touchpad-based remote control devices. *CHI '98* (1998).
- [19] Ferscha, A., Vogl, S., Emsenhuber, B., and Wally, B. Physical shortcuts for media remote controls. *Proc. of INTETAIN '08* (2008).
- [20] Freeman, W.T. and Weissman, C. Television control by hand gestures. *Proc. Intl. Workshop on Automatic Face and Gesture Recognition*, (1994).
- [21] Goto, J., Kim, Y.-B., Miyazaki, M., Komine, K., and Uratani, N. A spoken dialogue interface for television

- operations based on data collected by using WOZ method. *Proc. of Assoc for Computational Linguistics '03* (2003).
- [22] Gueddana, S. and Roussel, N. Pêle-Mêle, une étude de la communication multi-échelles. *Proc. of IHM'09* (2009).
- [23] Greenberg, S., Marquardt, N., Ballendat, T., Diaz-Marino, R., and Wang, M. Proxemic interactions: the new ubicomp? *Interactions* 18, 2011 (2011).
- [24] Gurrin, C., Lee, H., Caprani, N., Zhang, Z.X., O'Connor, N., and Carthy, D. Browsing Large Personal Multimedia Archives in a Lean-Back Environment. *Advances in Multimedia Modeling*, (2010).
- [25] Hess, J., Küstermann, G., and Pipek, V. Premote. *Ext. Abs. CHI '08*, (2008).
- [26] Ibrahim, A. and Lundberg, J. Speech enhanced remote control for media terminal. *European Conference on Speech*, (2001).
- [27] Ingmarsson, M., Dinka, D., and Zhai, S. TNT: a numeric keypad based text input method. *Proc of CHI '04* (2004).
- [28] Ishiyama, K. and Yano, S. A study of characteristics of pointing devices for television operation. *Proc. of IEEE Int. Conf. on Systems, Man and Cybernetics*. (2000).
- [29] Jensen, J.F. Interactive television: new genres, new format, new content. *Proc. of the second Australasian conference on Interactive entertainment* (2005).
- [30] Jones, E., Alexander, J., Andreou, A., Irani, P., and Subramanian, S. GesText: accelerometer-based gestural text-entry systems. *Proc. of CHI '10* (2010).
- [31] Kimman, F., Weda, H., Hoven, E. van den, Zeeuw, T. de, and Luitjens, S. Spinning in control: design exploration of a cross-device remote. *Proc. of TEI'11* (2011).
- [32] Kobayashi, K., Nakagawa, Y., Yamada, S., Nakagawa, S., and Saito, Y. Rebo: A remote control with strokes. *Proc. of The 18th IEEE International Symposium on Robot and Human Interactive Communication* (2009).
- [33] Kurtenbach, G. and Buxton, W. Issues in combining marking and direct manipulation techniques. *Proc. of UIST '91* (1991).
- [34] Lee, C.-H.J., Chang, C., Chung, H., Dickie, C., and Selker, T. Emotionally reactive television. *International Conference on IUI'07*, (2007).
- [35] Lee, D.-W., Lim, J.-M., Sunwoo, J., Cho, I.-Y., and Lee, C.-H. Actual remote control: a universal remote control using hand motions on a virtual menu. *IEEE Transactions on Consumer Electronics* (2009).
- [36] Lima, J.V.D. and Nevado, R.A.D. Common Paper as an Interface for Digital television. *Ninth IEEE International Symposium on Multimedia Workshops* (2007).
- [37] MacKenzie, I. and Jusoh, S. An Evaluation of Two Input Devices for Remote Pointing. *Engineering for Human-Computer Interaction* (2001).
- [38] Malacria, S. and Lecolinet, E. Espace de caractérisation du stylo numérique. *Proc. of IHM '08* (2008).
- [39] Marshall, D., Foster J.C., and Jack, M.A. User performance and attitude towards schemes for alphanumeric data entry using restricted input devices. *Behaviour & information technology* 20, 3 (2001).
- [40] Mirlacher, T., Pirker, M., Bernhaupt, R., Fischer, T., Schwaiger, D., Wilfinger, D. and Tscheligi, M. Interactive simplicity for iTV: minimizing keys for navigating content. *Proc. of EuroITV '10* (2010).
- [41] Meerbeek, B., Bingley, P., Rijnen, W., and Hoven, E. van den. Pipet: a design concept supporting photo sharing. *Proc. of NordiCHI '10* (2010).
- [42] Nack, F., Schiphorst, T., Obrenovic, Z., Kauwatjoe, M., De Bakker, S., Rosillo, A. P., and Aroyo, L.. Pillows as adaptive interfaces in ambient environments. *Proc of the Int. Workshop on Human-centered multimedia* (2007).
- [43] O'Modhrain, S. and Oakley, I. Touch TV: Adding feeling to broadcast media. *Proc of EuroITV '03* (2003).
- [44] O'Modhrain, S. and Oakley, I. Adding interactivity: active touch in broadcast media. *Proc. of HAPTICS '04*. (2004)
- [45] Pirker, M., Bernhaupt, R., and Mirlacher, T. Investigating usability and user experience as possible entry barriers for touch interaction in the living room. *Proc. of EuroITV '10* (2010).
- [46] Robertson, S., Wharton, C., Ashworth, C., and Franzke, M. Dual device user interface design: PDAs and interactive television. *Proc. of CHI '96*. (1996)
- [47] Schmidt, A., Holleis, P., and Kranz, M. Sensor virrig-a balance cushion as controller. *Workshop, Adjunct Proc. of UbiComp '04*. ACM Press (2004).
- [48] Sohn, M. and Lee, G. SonarPen : an ultrasonic pointing device for an interactive television. *IEEE Trans. on Consumer Electronics*, 50, 2 (2004).
- [49] Tahir, M., Bailly, G., Lecolinet, E., and Mouret, G. TactiMote: a tactile remote control for navigating in long lists. *Proc of ICMI '08*, (2008).
- [50] Tsekleves, E., Cruickshank, L., Hill, A., Kondo, K., and Whitham, R. Interacting with Digital Media at Home via a Second Screen. *Ninth IEEE International Symposium on Multimedia Workshops* (2007).
- [51] Tsekleves, E., Whitham, R., Kondo, K., & Hill, A. Investigating media use and the television user experience in the home. *Entertainment Computing*, (2011).
- [52] Tseng, M. and Kolko, J. TRIBA: a cable television retrieval & awareness system. *Ext. Abs. of CHI'05* (2005).
- [53] Turunen, M., Mäkinen, E., Valkama, P., et al. Multimodal interaction with speech and physical touch interface in a media center application. *Proc. of the Int. Conf. on Advances in Computer Entertainment Technology* (2009).
- [54] Vatavu, R.-D. and Pentiu, S.-G. Interactive Coffee Tables: Interfacing television within an Intuitive, Fun and Shared Experience. *Proc. of EuroITV '08* (2008).
- [55] Villar, N. and Hodges, S. The peppermill: a human-powered user interface device. *Proc. of TEI '10* (2010).
- [56] Zigelbaum, J., Browning, A., Leithinger, D., Bau, O., and Ishii, H. g-stalt: a Chirocentric, Spatiotemporal, and Telekinetic Gestural Interface. *Proc. of TEI '10* (2010).