

# Conception centrée utilisateur de prototypes interactifs pour la gestion de contenu multimedia par similarité

Christian Frisson

Université de Mons, laboratoire TCTS  
Boulevard Dolez 31  
7000 Mons, Belgique  
[christian.frisson@gmail.com](mailto:christian.frisson@gmail.com)

## RESUME

Cet article présente les travaux en cours d'une recherche doctorale visant à proposer une méthodologie de conception centrée utilisateur et de prototypage rapide afin de concevoir des applications interactives destinées à la navigation par similarité dans des bases de données multimedia, adaptées à des cas d'utilisation divers et profils d'utilisateurs variés. Les modalités d'interaction sont volontairement restreintes à la visualisation d'information et l'interaction manuelle. Une méthode de développement rapide, réutilisable et durable est proposée, exemplifiée par quelques prototypes à évaluer par des tests utilisateur.

**MOTS CLES :** Prototypage rapide, visualisation d'information, interaction gestuelle, navigation hypermedia.

## ABSTRACT

This paper presents a work-in-progress of doctoral research aiming at defining a user-centered methodology combined with rapid prototyping dedicated to the design of interactive applications for the navigation by similarity in multimedia databases, adapted to different use cases and multiple user profiles. The interaction modalities are voluntarily restricted to information visualization and gestural interaction. A rapid, reusable and sustainable development method is proposed, applied in several prototypes that are to be validated by usability tests.

**CATEGORIES AND SUBJECT DESCRIPTORS:** H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

**GENERAL TERMS:** Design, Experimentation

**KEYWORDS:** Rapid prototyping, information visualization, gestural interaction, hypermedia navigation.

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, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IHM 2010, 20-23 Septembre 2010, Luxembourg, Luxembourg

## INTRODUCTION

### Contexte

**Media et similarité** L'essor démesuré des technologies des télécommunications et de l'information nous place dans une ère de l'explosion numérique, repoussant les limites de l'accès et de la mise à disposition de divers contenus d'information. Les actualisations successives des termes empruntés pour décrire les systèmes informatisés gérant ces multiples contenus marquent ces évolutions: "multimedia" pour plusieurs types simultanés, "hypermedia" pour la connectivité entre ces divers types, "crossmedia" vers leur corrélation. Face à la variété conséquente, l'utilisateur a de plus en plus recours à des systèmes de recommandation pour être assisté dans ses choix, sur base de son expérience et ses préférences. Mais comment informatiser des propositions de contenu similaire?

### L'interface utilisateur au service des cas d'utilisation

La navigation et manipulation de documents multimedia touche de nombreux praticiens et amateurs:

- sur leurs appareils, mobiles ou fixes: une grande partie de la population mondiale;
- en médiathèque: les abonnés et les documentalistes;
- dans leurs bureaux et studios: les compositeurs de musique électro-acoustique, de bande-son de films, les designers sonores, les monteurs vidéo;
- sur scène: les musiciens (DJ) et vidéastes (VJ) qui utilisent des échantillons audio-visuels;
- dans des musées: les visiteurs d'installations et bornes interactives, leurs concepteurs (artistes, commissaires...).

Mais alors comment bien adapter l'interface utilisateur aux divers cas et domaines d'application, aux multiples profils d'utilisateurs?

Plusieurs familles d'interfaces utilisateur, plus ou moins inspirées par la métaphore des documents sur un bureau, ont déjà permis de mettre en oeuvre ces systèmes: graphiques (GUI), zoomables (ZUI), naturelles (NUI), etc... [7]. Choisir des techniques de visualisation d'information en considérant leurs effets cognitifs [9] permet d'améliorer la fouille ou le "data mining" visuel dans d'autres domaines que la gestion de bases de données multimedia.

## Problèmes

**Dans la méthode** L'interaction homme-machine, interdisciplinaire, rassemble plusieurs domaines d'expertise: le design d'interaction, l'ingénierie logicielle, l'étude de la cognition, de l'ergonomie... La navigation par similarité dans des bases de données multimedia est une pratique naissante issue de technologies récentes qui a peu d'analogies pré-existantes. Une équipe qui tente de proposer des systèmes informatisés le permettant est confrontée à devoir supposer les attentes et besoins de l'utilisateur pour pouvoir concevoir ces systèmes; à moins d'adopter, en parallèle du développement logiciel et matériel, d'une approche centrée-utilisateur. Mais comment bien concilier ces deux voies d'expertise? Comment bien dégager les généralités et spécificités communes à ou distinguant les divers profils et domaines? Comment soutenir un développement logiciel et matériel qui soit rapide, réutilisable, modulaire et durable; en connaissance de ces généralités et spécificités?

**Dans les domaines d'application** Les dispositifs interactifs disponibles permettent de plus en plus de supporter plusieurs types de media, mais sans transversalité dans le classement des medias, actuellement par type de media.

Les questionnaires de bibliothèques multimedia sont souvent peu ergonomiques et pas forcément bien adaptés à la tâche: des listes interminables de documents proposées que l'on peut classer selon leurs méta-données et parfois annoter avec des grades de préférences. Bien que des techniques de visualisations mieux adaptées apparaissent (flux de couvertures d'album, similarité de sous-segments du media), elle restent encore isolées. Certains lecteurs audio agrémentés de visualisations apparemment basées sur le signal affichent une volonté de proposer une interface utilisateur agréable et engageante, pourquoi ne pas rediriger ces efforts pour visualiser correctement une information utile, en plus d'être plaisante?

La démocratisation des appareils et logiciels d'enregistrement, de montage et de traitement multimedia rend la frontière entre l'amateur et l'expert plus floue: l'interface utilisateur peut-elle être adaptée à ces deux profils extrêmes?

## OBJECTIFS

### Choix de restrictions de l'étendue de la recherche

L'implémentation et l'optimisation des techniques d'analyse et de classification basées signal sont confiés aux collaborateurs du projet permettant cette thèse. Toutefois, il serait illusoire de mener à bien ces travaux sans une bonne compréhension de l'influence de ces algorithmes.

Par multimedia nous retiendrons surtout les medias à temporalité (audio et vidéo) qui ont des spécificités en tant que séries temporelles pour la visualisation d'information. Les données textuelles, pouvant faire office de documents à part entière, seront en premier lieu considérées uniquement pour l'annotation sémantique.

Comme modalité d'entrée, nous nous concentrerons essentiellement sur le contrôle manuel, tangible, instrumental; par opposition aux modalités d'entrée sans contact, selon la taxonomie proposée par Saffer [8]. Nous pensons que les experts potentiellement intéressés par la gestion par similarité de contenu multimedia ont une pratique souvent de longue durée et qui nécessite par conséquent une moindre fatigue corporelle avec des mouvements réduits; et que la manipulation d'outils manuels fait partie de leur expérience cognitive (le contrôle rotatif pour la navigation au sein d'un média est présent chez les monteurs vidéo aussi bien que chez les DJs).

En modalité de sortie, nous considérerons particulièrement le retour visuel par le choix de techniques de visualisation d'information adaptées en utilisant pertinemment les variables visuelles selon la perception humaine [9].

## Résultats attendus

Nous comptons proposer plusieurs prototypes adaptés à la tâche et à l'utilisateur et pratiquer des tests utilisateur sur ces prototypes afin trouver une solution à des problèmes tels qu'augmenter la rapidité à trouver un media que l'utilisateur recherche, ou assurer que l'utilisateur aura une expérience attrayante avec de tels systèmes.

## METHODOLOGIE

### Approche centrée utilisateur

Nous avons adopté une approche centrée utilisateur largement inspirée de celle établie par Bernsen et Dybkjaer [1]:

- mener des enquêtes contextuelles auprès d'utilisateurs experts à l'aide de questionnaires comme support de discussion,
- proposer des solutions à travers des brainstormings en présence de tous les collaborateurs du projet;
- produire des maquettes, storyboards... avant d'entreprendre une implémentation logicielle et matérielle qui pourrait s'avérer inutile;
- assurer la continuité entre les divers cas d'utilisation, domaines d'application, profils utilisateurs, par la conception d'un environnement de prototypage rapide, modulaire et réutilisable;
- valider les prototypes produits par des tests utilisateurs en situation réelle.

### Développement durable

Comme de nombreux produits redondants saturent le marché sans toujours offrir de solutions originales et uniques et que la contrainte d'être à jour décharge les produits en fin de vie vers les pays et populations moins favorisés, nous voulons réduire l'impact de nos travaux de recherche sur l'environnement: en pré-évaluant des maquettes, par prototypage rapide et réutilisable en réadaptant des dispositifs interactifs pré-existants. Des implémentations multi-plateformes et sous licences libres devraient permettre d'assurer aux solutions proposées une meilleure pérennité.

## CONTRIBUTIONS

### Conceptualisation

**Vocabulaire et métaphores** Un classement de la terminologie utilisée pour décrire des métaphores issues de cas analogues à la gestion de contenu multimedia mais non informatisés permet de dégager et identifier des modes d'interaction et de suggérer des modalités d'interaction:

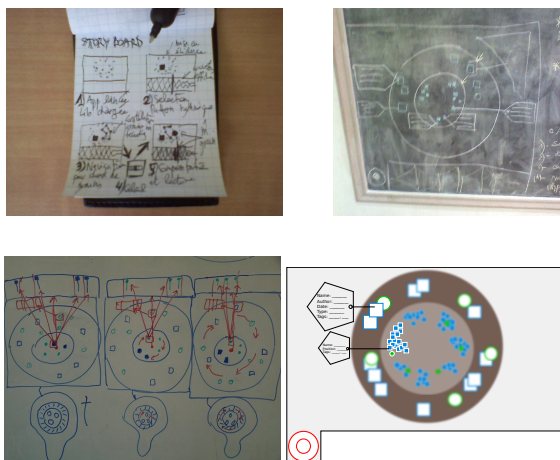
- à l'échelle de la base de données: "classer", "empiler" (classer en groupes), "compiler" (retenir l'essentiel)...
- au sein d'un élément: "skim" et "scrub" (parcourir cursivement), "cue" (mettre des marqueurs temporels), "jist" (produire un résumé)...
- conjoints: "naviguer", "manipuler", "annoter"...

**Modes d'interaction et échelles** Nous distinguons deux modes d'interaction: la navigation dans la base de données, entre éléments; et la navigation au sein d'un élément.

**Modes de navigation dans la base de données** Nous discernons deux modes de navigation: exploratoire par "clustering" avec visibilité intégrale sur la bibliothèque multimedia organisée par similarité autour d'un élément choisi comme référent; et de proches en proches par voisinages, partant également d'un référent.

### Prototypes

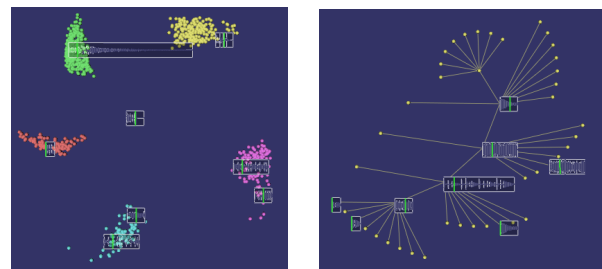
**Maquettes, storyboards** Les travaux de recherches de ce projet sur la gestion de contenu multimedia par similarité ont été démarrés avant cette thèse avec une méthode mue par l'ingénierie logicielle, avec un accent sur l'implémentation et l'amélioration des algorithmes d'extraction de caractéristiques du signal et de classification sur base de ces caractéristiques. Nous nous sommes cependant astreints à précéder les implémentations de l'interface utilisateur par des maquettes papier et storyboards, illustrés en Figure 1, suivant l'approche centrée utilisateur décrite précédemment. Cela nous a permis de gagner du temps de développement et d'améliorer notre compréhension mutuelle des objectifs.



**Figure 1 :** Exemples de *mockups* et *storyboards* sur divers supports: papier, tableaux blanc et noir, dessin vectoriel

**Framework MediaCycle, applications AudioCycle** Pour implémenter nos prototypes, nous avons consolidé au cours de court projets applicatifs le *framework* MediaCycle, dont les dernières avancées sont présentées dans [3]. MediaCycle a pour vocation de permettre le développement rapide d'applications en ligne ou autonomes nécessitant le classement par similarité de contenu multimedia (actuellement: audio, images et vidéos, séparément). Les premiers travaux dont il est question ici ont été essentiellement basés sur l'application AudioCycle pour la navigation dans des bibliothèques de boucles et échantillons audio. Outre la variation possible sur la nature du media, MediaCycle adopte une architecture modulaire, le noyau interconnectant des composants d'extraction de caractéristiques, de classification, de visualisation, d'interaction et de communication client/serveur.

**Visualisation d'information modulaire** Pour réaliser les deux types de vues (exploratoires par clusters, illustrée en Figure 2 en haut; et de proches en proches, en bas), nous avons commencé par emprunter des techniques existantes de visualisation d'information, notamment les diagrammes noeud-lien pour la vue par voisinages. Les algorithmes de classification (par clusters ou voisins), de calcul de position des noeuds représentant chaque media, et de rendu graphique sont tous dissociables, ce qui permet de s'adapter rapidement aux contraintes d'implémentation logicielles (par exemple: choix du langage). Nous avons porté certaines implémentations d'algorithmes de visualisation en arbre de la librairie Prefuse<sup>1</sup>. Nous avons également introduit des transitions animées [5] lors des changements de vues et de noeuds référents car nous pensons que cela améliore cognitivement pour l'utilisateur sa perception progressive de la nature de la base de donnée multimedia et de son historique de navigation.

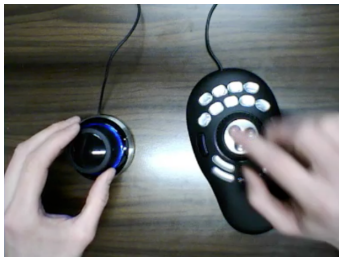


**Figure 2 :** Exemples de visualisations en modes exploratoire (haut) et de proches en proches (bas) d'une base de données de boucles audio avec AudioCycle.

**Prototypage rapide et interaction gestuelle** Pour le contrôle manuel de la navigation dans la base de données et la manipulation d'éléments de cette base, nous avons commencé par tester des dispositifs du marché ("off-the shelf") qui proposaient un contrôle rotatif et/ou dans un espace 2D/3D. Bien que de plus en plus de boîtes à outils dédiées au prototypage rapide d'interaction multimodale

<sup>1</sup><http://www.prefuse.org>

sont disponibles [2], nous avons choisi l’environnement de programmation visuelle de traitement de signal Pure-Data<sup>2</sup>, comme expliqué dans [4], car cette boîte à outils multi-plateforme et sous licence libre permet: d’explorer les dispositifs HID (Human Interface Device), d’en filtrer les événements, et de communiquer par réseau avec d’autres applications. Nous avons par conséquent implémenté un support client / serveur du protocole Open-SoundControl<sup>3</sup> dans le framework MediaCycle et défini un *namespace* classifiant les différentes actions de l’utilisateur selon les modes d’interaction, augurant de futures applications multi-utilisateur et multi-dispositifs. Nous avons pour l’instant conçu trois prototypes: (1) la combinaison traditionnelle souris et clavier “WIMP”, améliorée par une surface de contact multi-doigts; (2) un contrôle bimanuel de la navigation de la base par une souris 3D et dans un élément par une molette rotative, illustré en Figure 3; (3) une variante à retour d’effort, avec une souris 3D à 3 degrés de liberté.



**Figure 3** : Interaction bimanuelle et assignation: navigation à la main gauche (souris 3D) et manipulation à la main droite (contrôleur rotatif).

## PERSPECTIVES

### Renforcer l’expression de la similarité par l’annotation

Pour consolider la représentation de la base de données, nous comptons allier les algorithmes de traitement de signal à des techniques d’annotation sémantique, vers une classification semi-automatique, plus proche de l’utilisateur.

### Affiner et diversifier l’interaction

Nous comptons étudier des méthodes d’interaction collaboratives, pour proposer par exemple une application de recherche multi-utilisateur dans une médiathèque. Pour mieux adapter l’interaction aux souhaits de l’utilisateur et sa morphologie, nous avons commencé à ré-adapter les dispositifs d’interaction testés.

### Rendre le prototypage infovis plus rapide

Outre la possibilité de changer de vue et de technique de visualisation au sein d’une vue, nous évaluerons comment modéliser les scénarios de fouille visuelle et les programmer visuellement, similairement à VisTrails<sup>4</sup>.

<sup>2</sup><http://www.puredata.info>

<sup>3</sup><http://www.opensoundcontrol.org>

<sup>4</sup><http://vistrails.org>

## Valider les prototypes par des tests utilisateurs

Il est nécessaire de valider et améliorer nos prototypes par des tests utilisateurs, qualitatifs à l’aide de questionnaires de satisfaction [6] pour évaluer des facteurs comme la “*pleasurability*”; quantitatifs, notamment la rapidité à trouver un media cible. Nous pensons que rétablir le retour d’effort dans l’interaction gestuelle informatisée devrait permettre d’accroître ces deux critères.

## Enquête contextuelles

Des enquêtes contextuelles visant le design sonore et les performances DJ sont en cours pour préciser les besoins ou attentes d’utilisateurs potentiels.

## REMERCIEMENTS

Ce travail a été financé par la Région Wallonne, Belgique, à travers numediart<sup>5</sup>, programme de recherche autour des technologies des arts numériques (financement 716631).

## PROMOTEURS

Thierry Dutoit, Université de Mons.

Jean Vanderdonckt, Université catholique de Louvain.

## BIBLIOGRAPHIE

1. Bernsen, N. O., and Dybkjaer, L. *Multimodal Usability*. Springer, 2009.
2. Dumas, B., Lalanne, D., and Oviatt, S. *Human Machine Interaction*, volume 5440 of *LNCIS*, chapter Multimodal Interfaces: A Survey of Principles, Models and Frameworks, pages 3–26. Springer, 2009.
3. Dupont, S., Frisson, C., Siebert, X., and Tardieu, D. Browsing sound and music libraries by similarity. In *128 AES Convention*, London, UK, May 22-25 2010.
4. Frisson, C., Dupont, S., Siebert, X., Tardieu, D., Dutoit, T., and Macq, B. DeviceCycle: rapid and reusable prototyping of gestural interfaces, applied to audio browsing by similarity. In *Proceedings of the New Interfaces for Musical Expression++ (NIME++)*, Sydney, Australia, June 15-18 2010.
5. Heer, J., and Robertson, G. Animated transitions in statistical data graphics. In *IEEE Information Visualization (InfoVis)*, 2007.
6. Lewis, J. R. IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 7, 57-78., 7:57–78, 1995.
7. Moggridge, B. *Designing Interactions*. The MIT Press, 2007.
8. Saffer, D. *Designing Gestural Interfaces*. O’Reilly Media, Inc., 2009.
9. Ware, C. *Visual Thinking: for Design*. Interactive Technologies. Morgan Kaufmann, 2008.

<sup>5</sup><http://www.numediart.org>