

Designing tangible/free-form applications for navigation in audio/visual collections (by content-based similarity)

Christian Frisson

numediart Institute, TCTS lab, University of Mons (UMONS)
Boulevard Dolez 31, B-7000 Mons, Belgium
christian.frisson@umons.ac.be

ABSTRACT

This paper focuses on one aspect of doctoral studies, within the last year of completion, consisting in designing applications for the navigation (by content-based similarity) in audio or video collections: the choice of tangible or free-form interfaces depending on use cases. One goal of this work is to determine which type of gestural interface suits best each chosen use case making use of navigation into media collections composed of audio or video elements, among: classifying sounds for electroacoustic music composition, de-rushing video, improvising instant music through an installation organizing and synchronizing audio loops. Prototype applications have been developed using the modular MediaCycle framework for organization of media content by similarity. We conclude preliminarily that tangible interfaces are better-suited for focused expert tasks and free-form interfaces for multiple-user exploratory tasks, while a combination of both can create emergent practices.

Author Keywords

Interface design, multimedia content organization, free-form interfaces, tangible interfaces.

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation:
Miscellaneous—*Input devices and strategies*

General Terms

Design, Experimentation

INTRODUCTION

In his book [4], particularly chapters 8 “Multisensory and Multimedia” and 10 “People and Prototypes”, the late Mogridge broaches the scope of our work: how to design seamless interaction for applications featuring multimedia content. Our work intends to draw guidelines on how to design interfaces for navigation in audio or video collections, depending on the use case, as classified by the following questions:

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.

TEI 2013, February 10 - 13, 2013, Barcelona, Spain.

Copyright 2013 ACM 978-1-4503-1898-3/13/02...\$15.00.

- How many users are involved? One only? Or multiple?
- What is the setting of the task? Lab, exhibition, home...?
- How deep is the task focused? Exploratory, intended?

METHOD

User-centered and agile

These doctoral studies have been undertaken with fast-paced cycles of trimestrial projects within the numediart research program in digital art technologies¹. It entailed collaboration with artists or industrial partners. This situation has been determining the progress of this work: pushing forward new technologies on a short notice while trying to ensure use cases could be satisfied on a longer term by a close relationship with users.

CONTRIBUTIONS

Modular media content organization with MediaCycle

An actual and accurate summary of the description of the MediaCycle framework² is provided in [2]. In short, it allows to create applications for multimedia navigation and organization (by content-based similarity), making use of plugins dedicated for the support of different media types (audio, video, images, 3D models, text...) and for each sequence of the workflow of organization (file reading, feature extraction, thumbnailing, clustering or classification, computation of positions of media elements in a 2D space).

Gestural prototypes

Several interactive applications have been created using the MediaCycle framework, among them:

- the DeviceCycle toolbox [3] allows to interface many off-the-shelf USB/HID device to control MediaCycle applications through the OpenSoundControl (OSC) protocol by refining interaction cues with the PureData modular environment, particularly for tangible interfaces, as illustrated in Figure 1, including: jog wheels, multitouch trackpads, (force-feedback) 3D mice;
- the LoopJam installation [1] allows multiple visitors to compose instant musical creations, with a free-form interface sensing their center of gravity, corresponding to a position in a 2D map of sounds incidentally triggering the closest loop, as seen in Figure 2.

¹<http://www.numediart.org>

²<http://www.mediacycle.org>



Figure 1. *The Remote Controller* (2003) by People Like Us alias Vicki Bennett in VideoCycle: zooming the inter-segment “browser” (up), scrubbing the intra-media “timeline” (down) with the jog shuttle wheel (down). Yellow dotted lines/arrows are annotations.

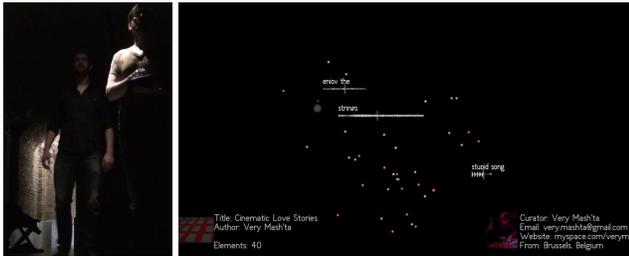


Figure 2. Two participants playing with the LoopJam installation, looking at the displayed map of organized sounds: both trigger one loop each by the position of their center of gravity sensed by the Kinect camera, while one of them triggers one additional sound touching an iPad. Both free-form and tangible interactions feature gestural-to-visual space projection.

A more recent tryout tries to combine both types of gestural interfaces: tangible and free-form. One of the visitors of the LoopJam installation, audio collagist Gauthier Keyaerts, in Figure 2, triggers one sound by its location of the space, but also triggers other sounds by means of a tablet that allows access to the same sound map, with one loop per finger.

EVALUATION

Qualitative: contextual inquiries

A contextual inquiry has been undertaken with 5 electroacoustic music composers asked to create collections of loops for the LoopJam installation. Quoting [5], such a number of users has been sufficient to grasp 80% of the most common usability issues of our application, ranging from generic shortcuts widely used in other multimedia applications to obstacles in the workflow of the application. A similar inquiry is being considered with video artists.

Quantitative: known-item search tasks

We are participating to a challenge consisting in assessing how fast and accurately expert and novice users can find a segment of a video presented on a screen using diverse video browsing systems including ours [2]. For that contest, we proposed a tangible interface as illustrated in Figure 1. We will replicate similar usability tests locally.

We are preparing a quantitative analysis of the LoopJam installation from data obtained by the depth-sensing camera, generic “metrics” featuring the average participation time per user, the agitation of each user, the participation along the time of the day or the day of the week, and so on...

CONCLUSION

We come up with the following preliminary assumptions that still require a complete validation:

- Daily, focused, precise and repetitive tasks, such as multimedia data mining including video de-rushing, work best with tangible interfaces that allow minimal movements in the gestural space covering all the media content space.
- Exploratory, often unintentional tasks, such as playing with public installations, can take advantage from free-form interfaces, for instant playability without a steep virtuous learning curve, especially when collaborative use including concurrent access to the interface is at stake.
- “Mashup” applications that combine both types of gestural interfaces, tangible and free-form, can alleviate the drawbacks of each.

ACKNOWLEDGEMENTS

These doctorate studies are benefiting from supervision by advising professors Thierry Dutoit from University of Mons (UMONS) and Jean Vanderdonck from Université catholique de Louvain (UCL), both from Belgium. This work is being funded by the numediart Institute along a grant (N°716631) from Région Wallonne, Belgium. The MediaCycle framework has been developed in collaboration with (alphabetically): Stéphane Dupont, Alexis Moinet, Cécile Picard-Limpens, Thierry Ravet, Xavier Siebert.

REFERENCES

1. C. Frisson, S. Dupont, J. Leroy, A. Moinet, T. Ravet, X. Siebert, and T. Dutoit. LoopJam: turning the dance floor into a collaborative instrumental map. In *Proceedings of the New Interfaces for Musical Expression (NIME)*, Ann Arbor, Michigan, USA, May 21-23 2012.
2. C. Frisson, S. Dupont, A. Moinet, C. Picard-Limpens, T. Ravet, X. Siebert, and T. Dutoit. VideoCycle: user-friendly navigation by similarity in video databases. In *Proceedings of the 19th International Conference on Multimedia Modeling (MMM2013), Video Browser Showdown*, Huangshan, China, January 7-9 2013.
3. C. Frisson, S. Dupont, X. Siebert, D. Tardieu, T. Dutoit, and B. Macq. 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.
4. B. Moggridge. *Designing Interactions*. MIT Press, 2007.
5. T. Tullis and W. Albert. *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*. Interactive Technologies. Morgan Kaufmann, 2008.