



UvA-DARE (Digital Academic Repository)

Enriching the user experience with multimodal interfaces

Salah, A.A.; Dutoit, T.

DOI

[10.1007/s12193-011-0061-9](https://doi.org/10.1007/s12193-011-0061-9)

Publication date

2011

Document Version

Final published version

Published in

Journal on Multimodal User Interfaces

[Link to publication](#)

Citation for published version (APA):

Salah, A. A., & Dutoit, T. (2011). Enriching the user experience with multimodal interfaces. *Journal on Multimodal User Interfaces*, 4(2), 59-60. <https://doi.org/10.1007/s12193-011-0061-9>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Enriching the user experience with multimodal interfaces

Albert Ali Salah · Thierry Dutoit

Published online: 17 June 2011

© The Author(s) 2011. This article is published with open access at Springerlink.com

In the age of technical specialization, research into multimodal interfaces presents particular logistic challenges in bringing together technical expertise in different modalities, in multimodal fusion, and in integration. Data collection and usability studies are time consuming and require a further set of skills, as well as manpower. To meet these challenges, the eNTERFACE'10 Workshop on Multimodal Interfaces brought together 65 researchers from 18 countries in Amsterdam, for the duration of four weeks. This special issue features three research papers that were initiated as projects at the eNTERFACE'10,¹ and extended during the six months that followed the workshop.

The eNTERFACE Workshops have been places for initiating international, long-term collaborations. The paper "Automatic Fingersign-to-Speech Translation System" is written by such a group, where some of the members have been working together since the first eNTERFACE in 2005. In this paper, the authors describe a system for helping the communication of visually-impaired and hearing-impaired people, by converting speech to fingerspelling and fingerspelling to speech. The multi-national nature of the team is revealed in the experimental setup, where the authors report

person-dependent and person-independent results with four different languages (Turkish, Czech and Russian for fingerspelling input, and Turkish, English and Russian for speech recognition and synthesis). For speech synthesis, such a multitude of languages presents additional hurdles, and different approaches are adopted for each language. The resulting system works in real-time.

In the past few years, we see a marked trend for exploring user interfaces that take affective and social signals into account. The paper by Dibeklioğlu et al. describes an affect-responsive interactive photo-frame application that offers its user a different experience with every use. It relies on visual analysis of activity levels and facial expressions of its users to select responses from a database of short video segments. This ever-growing database is automatically prepared by an offline analysis of user-uploaded videos. The resulting system matches its user's affect along dimensions of valence and arousal, and gradually adapts its response to each specific user. In an extended mode, two such systems are coupled and feed each other with visual content. The strengths and weaknesses of the system are assessed through a usability study, where a Wizard-of-Oz response logic is contrasted with the fully automatic system that uses affective and activity-based features, either alone, or in tandem.

The work described by Reidsma et al. resembles the affective photo frame from an interaction perspective, but has a much more ambitious goal: To design a virtual human that can sustain a continuous interaction with a person. The benefit of a virtual human over recorded sequences of a real human is that the signals produced by the virtual character can be synthesized to match the interacting person. However, there are great challenges here. The multimodal analysis of the behavior of the interacting person should be fast enough, and includes for instance classification of listener responses. On the synthesis part, scheduling and planning

¹The eNTERFACE'10 was organized by A.A. Salah and Th. Gevers at the University of Amsterdam, and seven projects were conducted in total. The full proceedings of eNTERFACE'10 can be downloaded from the Workshop website, <http://enterface10.science.uva.nl/>.

A.A. Salah (✉)
ISLA, University of Amsterdam, Science Park 904, 1098XH,
Amsterdam, The Netherlands
e-mail: a.a.salah@uva.nl

T. Dutoit
Department of Electrical Engineering, University of Mons,
TCTS Lab, 31 Boulevard Dolez, 7000, Mons, Belgium
e-mail: thierry.dutoit@umons.ac.be

Fig. 1 Stills from the eINTERFACE'10 Workshop at the University of Amsterdam



of multimodal behavior for the virtual human is required, so that appropriate reactions to listener responses can be generated, and a graceful and natural turn-taking pattern can be enabled.

Taken together, these three papers demonstrate that the interaction between two communicating parties (be it human or virtual) is not unlike a well-timed and precisely coordinated dance, requiring rapid processing and a lot of anticipation on both sides. This clearly portends that improvements in hardware and algorithms will continue enabling more so-

phisticated approaches to multimodal user interfaces, and new applications will become possible.

We would like to thank all the authors who submitted manuscripts for this Special Issue, as well as our reviewers for their valuable suggestions and rigorous evaluations of the papers.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.