

IM2 Newsletter

www.im2.ch

IM2, c/o IDIAP Research Institute, Centre du Parc,
Rue Marconi 19, P.O. Box 592, 1920 Martigny
info@im2.ch - www.im2.ch

Contents

COVER STORY

- An IM2 IP Head as General Chair of HRI2011 1

FOCUS

- Completed Thesis, Bruno Dumas – DIVA, UniFr 2
- Pomelo, a startup issued from IM2, on the way for the full kick 2
- Completed Thesis, Sarah Favre – Idiap Research Institute 3
- Completed Thesis, Dinesh Babu Jayagopi – Idiap Research Institute 3

INSIDE IM2

- News
- Selected publications

Event

HRI 2011

March 6-9, 2011
Ecole Polytechnique Fédérale de
Lausanne (EPFL)

6th Annual Conference for Basic and
Applied Human-Robot Interaction
Research

<http://hri2011.net/>

An IM2 IP Head as General Chair of HRI2011

MARCH 6-9, 2011, EPFL-LAUSANNE

Aude Billard (Head of the «Integrated Multimodal Processing» IP) is the General Chair of the HRI2011 conference which is the 6th Annual Conference for Basic and Applied Human-Robot Interaction Research. The conference will take place on March 6-9, 2011 in Lausanne, Switzerland, and the venue will be at the École Polytechnique Fédérale de Lausanne (EPFL).

Scientists from across the world submit their best work and attend HRI to hear the latest theories, data, and videos from the world's best HRI researchers. Each year, the HRI conference highlights a particular area.

The theme of HRI 2011 is Real World HRI. This theme is intended to highlight HRI in which basic scientific research is further tested in real world settings or applied to questions that arise in real world settings.

One central aspect of this type of research, in contrast to other realms of applied research, is that it is theoretically driven and feeds back to

our theoretical understandings. As such, real world research fortifies our understanding of people, robots, and interaction between the two.

HRI is a single-track, highly selective annual international conference that seeks to showcase the very best interdisciplinary and multidisciplinary research in human-robot interaction with roots in social psychology, cognitive science, HCI, human factors, artificial intelligence, robotics, organizational behavior, anthropology and many more, and we invite broad participation

There is funding available to partially support travel to HRI2011 for Student Volunteers. Students receiving support will be expected to volunteer time to help with on-site registration and with general duties, as needed, from 5 March (early on-site registration, tutorials and workshops) to 9 March, 2011

Conference registration is now available at the registration site: <http://inform.epfl.ch/index.php?form=HRI2011>

Aude Billard
aude.billard@epfl.ch



Cover Story

Completed Thesis, Bruno Dumas – DIVA, UniFr

FRAMEWORKS, DESCRIPTION LANGUAGES AND FUSION ENGINES FOR MULTIMODAL INTERACTIVE SYSTEMS

On December 2nd 2010, Bruno Dumas successfully defended his PhD thesis at the Computer Science Department of the University of Fribourg.

Supervised by Denis Lalanne (in the DIVA group of Prof. Ingold), this PhD thesis sought to reduce the chasm between interaction modalities and fusion of their data.

Fusion of multimodal input was approached in a global way in this research: first, from the point of view of the architecture of a multimodal system as a whole, then, from the point of view of multimodal dialog modeling, and finally from an algorithmic point of view. The architectural angle focused on necessary features of an architecture to allow beneficial integration of a fusion engine, using usability and expressivity as attributes to characterize qualities and drawbacks of different architectures.

Three different architectures targeting the creation of multimodal interfaces were subsequently studied, which lead to the creation of a framework for creation of multimodal interfaces, named HephaisTK. On the second angle, dialog modeling oriented, eight guidelines for creation of multimodal dialog modeling languages were defined, then the SMUIML language was designed as an example of a language following those guidelines.

Finally, the algorithmic angle studied multimodal fusion itself through the implementation of two fusion algorithms: an algorithm based on meaning frames and an algorithm based on hidden Markov models (HMMs).



An evaluation of the performances of those two algorithms with the help of a proposed integrated benchmarking tool was also conducted.

Through theoretical as well as practical study of these three angles – architecture, modeling, algorithms – the different issues of multimodal fusion were defined and clarified. The work achieved in the context of this PhD thesis is meant to serve as a stepping stone for the study of advanced concepts based on multimodal fusion, such as adaptation to user.

Bruno Dumas
bruno.dumas@unifr.ch

Pomelo, an EPFL startup issued from IM2, on the way for the full kick

EYE-TRACKING FOR THE SUPERMARKET

The startup company Pomelo, an offspring of the LASA laboratory at EPFL, has been selected for the second stage of the Venture Kick program in Zurich last week.

The eye-tracking technology, developed within the frame of IM2 (MPR + IP1) is at the core of the business idea which aims at providing eye-tracking information as a means for analysing visual behaviour.

The IM2 project has allowed the development of the innovative technology and thus provided Pomelo with a leading edge over



the competitors in the field of eye-tracking.

The Venture Kick program has evaluated 8 business ideas from all around Switzerland and has awarded 20'000 CHF to Pomelo and one of its competitors. Both teams will also have the chance to participate in the final financing round where they will compete for a prize of 100'000 CHF. The final round will take place in Zurich in a couple of months.

Basilio Noris
basilio.noris@epfl.ch

The WearCam , a lightweight, portable, head-mounted eye gaze-tracking device, designed both for adults and children

Completed Thesis, Sarah Favre – Idiap Research Institute

SOCIAL NETWORK ANALYSIS FOR AUTOMATIC ROLE RECOGNITION

On December 8th 2010, Sarah Favre successfully defended her PhD thesis at EPFL. Supervised by Alessandro Vinciarelli (IM2 IP Head), her dissertation was entitled "Social Network Analysis for Automatic Role Recognition".

In her PhD thesis, Sarah Favre has applied Social Network Analysis (a corpus of mathematical techniques developed by sociologists) for the automatic recognition of roles in social interactions. Roles are a key aspect of social interaction, sociologists have shown not only that people play roles each time they interact, but also that roles shape behaviour and expectations of interaction participants. While the interest of the computing community for the analysis of social interactions has increased in the last few years, only a few groups have worked on the role recognition problem.

During her thesis, Sarah thus aimed at filling this gap by investigating the problem of automatic role recognition in a wide range of interaction settings, including production environments, e.g. news and talk-shows, and spontaneous exchanges, e.g. meetings.



The role recognition approach she proposed includes two main steps. The first step aims at representing the individuals involved in an interaction with feature vectors accounting for their relationships with others. This step includes three main stages, namely segmentation of audio into turns (i.e. time intervals during which only one person talks), conversion of the sequence of turns into a social network, and use of the social network as a tool to extract features for each person.

The second step uses machine learning methods to map the feature vectors into roles.

Sarah has carried out experiments over roughly 90 hours of material, which is not only one of the largest databases ever used in literature on role recognition, but also the only one including different interaction settings. The achieved accuracy, i.e. the percentage of data correctly labeled in terms of roles, is roughly 80% in production environments and 70% in spontaneous exchanges (lexical features have been added in the latter case).

Sarah Favre
sarah.favre@idiap.ch

Completed Thesis, Dinesh Babu Jayagopi – Idiap Research Institute

COMPUTATIONAL MODELING OF FACE-TO-FACE SOCIAL INTERACTION USING NONVERBAL BEHAVIORAL CUES

Dinesh Babu Jayagopi (Idiap Research Institute) successfully defended his PhD thesis at EPFL on January 25, 2011. His dissertation was entitled "Computational Modeling of Face-to-Face Social Interaction using Nonverbal Behavioral Cues". The members of his doctoral jury were Pearl Pu Faltings (EPFL, president of the jury), Anton Nijholt (University of Twente), Fabio Pianesi (FBK), Jean-Philippe Thiran (EPFL), and Daniel Gatica-Perez (Idiap, thesis director). Dinesh's work was supported by IM2.

The computational modeling of face-to-face interactions using nonverbal behavioral cues is an emerging and relevant problem in social computing. Studying face-to-face interactions in small groups helps in understanding the basic processes of individual and group behavior; and improving team productivity and satisfaction in the modern workplace. Apart from the verbal channel, nonverbal behavioral cues form a rich communication channel through which people infer - often automatically and unconsciously - emotions, relationships, and traits of fellow members.

The thesis investigates two individual social constructs - dominance and status - in small groups. The work attempts to understand how dominance perceived by external observers could be estimated by different nonverbal audio and video cues, and affected by annotator variability, the estimation method, and the exact task involved. In the second part, perceived dominance and role-based status are jointly studied

to understand whether dominant people are the ones with high status and whether dominance and status in small-group conversations be automatically explained by the same nonverbal cues. Speaking activity, visual activity, and visual attention cues were employed for both the works.



The thesis also investigates group social constructs using both supervised and unsupervised approaches. The supervised framework defines group cues by aggregating individual cues over time and person, and uses them to classify group conversational contexts - cooperative vs competitive and brainstorming vs decision-making. The unsupervised framework automatically discovers group interaction patterns using probabilistic topic models. An objective evaluation of our methodology involving human judgment and multiple annotators, showed that the learned topics indeed are meaningful, and also that the discovered patterns resemble prototypical leadership styles - autocratic, participative, and free-rein - proposed in social psychology.

Dinesh Babu Jayagopi
djaya@idiap.ch

News

Best Paper Award for an IM2 researcher

Mobile and Ubiquitous Multimedia

Daniel Gatica-Perez (Idiap Research Institute) and Raul Montoliu (University Jaume I, Castellon, Spain) received the Best Paper Award at the 9th Int. Conf. on Mobile and Ubiquitous Multimedia (MUM), Limassol, Cyprus, Dec. 1-3, 2010, for their paper «Discovering Human Places of Interest from Multimodal Mobile Phone Data».

The work is a result of the ongoing collaboration between Daniel's group and Nokia Research Center Lausanne.

Daniel Gatica-Perez
gatica@idiap.ch

IM2 researcher on the Organisation Committee of NIPS

NIPS, December 10, 2010

Daniel Gatica-Perez (Idiap Research Institute) co-organized a workshop with Louis-Philippe Morency (University of Southern California) and Nigel Ward (University of Texas, El Paso) on Modeling Human Communication Dynamics, at the Neural Information Processing Systems Conference (NIPS) on December 10, 2010 in Whistler, Canada.

The workshop featured invited speakers and posters on computational modeling of human communicative behavior.

The list of speakers included Marian Stewart Bartlett (University of California, San Diego), Jeff Bilmes (University of Washington), Dan Bohus (Microsoft Research), Justine Cassell (CMU), and Noah Goodman (Stanford University).

The workshop was supported by the PASCAL2 European Network of Excellence.

More information at:
<http://projects.ict.usc.edu/hcd2010/>

Daniel Gatica-Perez
gatica@idiap.ch

Klewel (Idiap spin-off) records the President of the Swiss Confederation

Rolex Center, November 17, 2010

The EPFL purchased a multimedia capture station from Klewel to record and webcast lectures and events on their campus.

The solution has been used the first time in the Rolex Learning Center for a conference about sciences and development, on November 17, 2010.

The main invited speaker was Mrs Micheline Calmy-Rey, President of the Swiss Confederation. More information can be found at: <http://tinyurl.com/klewel-im2-rolex>

Klewel is also webcasting an EPFL Software Engineering lecture every week. More information is available at <http://tinyurl.com/klewel-im2-epfl>

Mael Guillemot
mael.guillemot@klewel.com

2010 IMD startup competition winners

Koemei, Idiap Spin-off, selected

Every year, IMD searches for the most promising startups it can find to work with its MBA students. Over the past 13 years over 180 early-stage ventures have benefited from this collaboration.

On December 16, 2010, IMD representatives (Jim Pulcrano and Benoit Leleux) announced the twenty startups that have been selected to work with their MBA and EMBA students in 2011. Among them, Koemei (Idiap Spin-off) will participate in the venture.

Valerie Devanthery
valerie.devanthery@idiap.ch

Selected publications

Mining group nonverbal conversational patterns using probabilistic topic models..

D. Jayagopi and D. Gatica-Perez
IEEE Transactions on Multimedia,
12(8), 790 – 802, Dec. 2010.

A fit of distraction.

A. L. Masson and D. Lalanne

In Proc. of the ACM CHI Conference on Human Factors in Computing Systems 2011.

Omnidirectional Object Duplicate Detection.

P. Vajda, I. Ivanov, L. Goldmann, and T. Ebrahimi,

In the 14th IEEE Digital Signal Processing (DSP) Workshop and 6th IEEE Signal Processing Education (SPE) Workshop (DSPE 2011), Sedona, Arizona, USA, 2011.

Efficient Video Coding based on Audio-Visual Focus of Attention.

J.-S. Lee, F. De Simone, and T. Ebrahimi

In the Journal of Visual Communication and Image Representation, Special Issue on Emerging Techniques for High Performance Video Coding..

The ACLD: Speech-based Just-in-Time Retrieval of Multimedia Documents and Websites.

A. Popescu-Belis,, J. Kilgour, P.Poller and A. Nanchen

In Proceedings of SSCS 2010 (ACM Multimedia Workshop on Searching Spontaneous Conversational Speech), Florence, p.45-48.

PyGml–Creation and evaluation of a Portable Gestural Interface.

M. Schwaller, D.Lalanne and O.A.Khaled

In proc. of 6th Nordic Conference on Human-Computer Interaction (NordCHI 2010), Reykjavik (Iceland), October 16 - 20 2010 , pp.773-776.

Enriching the Design and Prototyping Loop: A set of tools to support the creation of activity-based pervasive applications.

P. Bruegger, A. Lisowska, D.Lalanne and B. Hirsbrunner

In the Journal of Mobile Multimedia (JMM), Vol.6 No.4, Dec. 2010, pp. 339-360.

An Introduction to Sequence Analysis for Human Behavior Understanding

H. Salamin and A. Vinciarelli.

In Human Behavior Understanding», A.A.Salah and T.Gevers (eds.), Springer Verlag (2011).