INTERACTIVE
MULTIMODAL
INFORMATION
MANAGEMENT

IM2 Winter Institute 2007

Abstract list, February 19- 22, 2007, Morat

Title: Dialog Machines: A toneless world!

Duration: +2 hours
Speakers: Jean Hennebert (UniFr)

Schedule:

Monday 19, 14:15 – 15:30	Jean Hennebert (UniFr)
Monday 19, 16:00 – 17:00	Jean Hennebert (UniFr)

Abstract: Voice-activated telephony services are appearing at a growing pace. On the one side, the standardization of the technology has eased the development of such services, with, for example the introduction of VoiceXML. On the other side, IT managers are more and more motivated to replace the old-fashion touchtone based systems. Voice-activation finds applications in phonebanking, ticketing, call-routing, information kiosks, entertainment, etc. If the technical deployment has been made simpler over the last decade, the design of such voice-activated telephony services remains tricky. More specifically, the "dialog" that takes place between the system and the user must be carefully designed according to some usability principles. The objective of this tutorial is to give an overview of the key elements that enter in the conception of such dialog machines, putting in evidence the essential guidelines that make such applications usable. A part of the tutorial will also describe how, in practice, one can tune the performance of the system once it has been launched. Aspects of dialog design dedicated to telephony speaker verification will be also presented.

Title: "Introduction to user evaluation methods and usability engineering"

Duration: 2 hours
Speakers: Denis Lalanne, Enrico Bertini (UniFr)

Schedule:

Monday 19, 17:00 – 18:00	Denis Lalanne, Enrico Bertini (UniFr)
Wednesday 21, 08:15 – 09:15	Denis Lalanne, Enrico Bertini (UniFr)

Abstract: This tutorial will briefly present the usability engineering lifecycle and the major methods for evaluating user interfaces, quantitatively or qualitatively, with or without user involvement.

Title: "Metadata and annotations of multimodal meeting data -- content, representation, and access"

Duration: 4 hours
Speakers: A. Popescu-Belis, M. Flynn, S. Marchand-Maillet

Schedule:

Tuesday 20, 09:15 – 11:30	Andrei Popescu-Belis (ISSCO, UniGE)
Tuesday 20, 16:00 – 16:45	Stéphane Marchand Maillet (CVML, UniGe)
Tuesday 20, 16:45 – 18:00	Mike Flynn (IDIAP)

Abstract: This tutorial will present an overview of the multimodal data that is available to IM2 participants, and in particular the IM2/AMI meeting corpus, its annotations, their representation and the tools to access them. The goal of the tutorial is to increase the integration among IM2 partners through the use of common data for multimodal processing and retrieval.

Outline:
1. [APB, 1h] The IM2/AMI meeting corpus: available data, metadata and annotations (d-m-a). Description of the content and abstract structure of d-m-a, in particular the existing annotation dimensions. Access to d-m-a through the MMM server. Suggestions for adding new annotations.
The rest of the tutorial will be devoted to representation issues, motivations for using them, and tools to produce and to access them.

2. [APB, 1h] Database representation based on annotation dimensions. Conversion from NXT annotation files. Normalized representation of metadata. Conversion from implicit file structure. Tools and distribution formats.

3. [SMM, 45'] An entity/relationship approach to metadata and annotation representation.

4. [MF, 1h15'] A client-hub-server model for annotation production and consumption. Communication protocol for obtaining and providing annotations. Related model: annotation triples. Converting metadata and annotations to the triple-based representation.

Title: How invariants changed vision

Duration: 1 hour
Speaker:
Luc van Gool, ETHZ

Schedule:
Thursday 22, 08:15 – 09:15 Luc van Gool (ETHZ)

Abstract:
Matching is a generic problem in computer vision. It appears in virtually all vision tasks. Whether one wants to build 3D models, track targets, or recognize objects, being able to efficiently and effectively compare patterns is key. A major difficulty in all this is that any straightforward comparison is bound to fail, as it will pick up differences that are actually irrelevant. These may be due to variations in viewpoint or illumination, for instance. Invariants are features that remain the same under specific groups of transformations, and thereby flag similarity in spite of such variations. This strand of research has become especially influential in combination with the idea of confining the domain within which the invariants would be extracted. The resulting invariant regions have penetrated much of the modern computer vision literature.

Title: Face Imaging Processing

Duration: 2 hour
Speakers:
Jean-Phlippe Thiran (EPFL)
Sébastien Marcel (IDIAP)

Schedule:
Wednesday 21, 13:30 – 15:30 Jean-Phlippe Thiran (EPFL), Sébastien Marcel (IDIAP)

Abstract:
This tutorial will present the state of the art in face image processing, including an introduction to biometry, face detection, face recognition, facial expression recognition, as well as a short presentation of some of the activities in face image processing inside IM2.

Title: Exploiting the relationships between entities
Subtitle: Can structure replace content?

Duration: 3 hours
Speakers:
S. Marchand-Maillet, UniGE, A. Vinciarelli, IDIAP

Schedule:
Thursday 22, 09:15 – 10:15 S. Marchand-Maillet (CVML, UniGE), A. Vinciarelli (IDIAP)
Thursday 22, 10:30 – 12:30 S. Marchand-Maillet (CVML, UniGE), A. Vinciarelli (IDIAP)

Abstract:
Classical information retrieval and management is directed towards the content of documents at hand. In this tutorial, we investigate alternative techniques that allow to infer information from the structure underlying the organisation of items.
This topic spans the specificity of web-information retrieval where the hyperlink structure is used to infer information such as "authority". In this part, we will review and analyse the classical HITS and PageRank algorithms, which have led to concrete search systems.
Inter-relationships may also be exploited between the users of a system, leading to recommender systems, which act based on user actions and generally irrespective of the actual content of recommended items. In this part we review classical collaborative filtering strategies.
Finally, structure may be inferred from a group of actors with the help of Social Network Analysis. We show how the interaction pattern between individuals involved in multimedia recordings provides information about the underlying data structure that can be difficult to obtain using other features.

The aim of this tutorial is to introduce participant to processing data that are "hidden" in classical structures. These algorithms are popularised by the new trend on collaborative actions and data.