Issue N° 1 / April 2006

Editor : Céline Aymon, celine.aymon@idiap.ch

DIRAC

Detection and Identification of Rare Audiovisual Cues

Newsletter

www.diracproject.org

DIRAC c/o IDIAP Research Institute, Simplon 4, P.O. Box 592, CH-1920 Martigny

info@diracproject.org - www.diracproject.org

Contents

News

2006-04-15

DIRAC

After years of discussions, preparations, presentations, and negotiations, we can present you the new Integrated Project DIRAC (standing for the Detection and Identification of Rare Audiovisual Cues), awarded recently by the 6th EC programme in Cognitive Systems.

Cover Story

DIRAC Kick-off Meeting

FUNDED BY THE 6TH EUROPEAN FRAMEWORK PROGRAM

On January 19, 2006, the official kickoff of a new 5 year EC Integrated Project DIRAC was held in Martigny, Switzerland. Several representatives of each of the 9 DIRAC partners gathered together.

Beyond the social aspects of direct contacts to initiate dynamic collaborations, this kick-off meeting offered all the participants a good opportunity to share a common view of the project, while discussing in more detail the project objectives in terms of research and training.

On the research side, each institution presented an overview of its contribution to DIRAC for the next 18th months and a meeting was already scheduled to deepen the discussion held at the kickoff. Training program and opportunities to welcome a new partner in the project were also discussed.

Some administrative information was provided as to ensure a smooth reporting and management of this integrated project.

Finally, participants also focused on the current set up of the DIRAC website along with its intranet. DIRAC's various communication tools were presented including its Newsletter.

Meeting participants (left to right) : KUL: Kurt Kornelis, ETHZ: Tobias Jaeggli, OL: Joern Anemueller, OHSU: Misha Pavel, ETHZ: Bastian Leibe, ETHZ: Luc Van Gool, HUJI: Daphan Weinshall, IDIAP: Nancy-Lara Robyr, KUL: Rufin Vogels, CTU: Thomas Svoboda, CTU: Thomas Pajdla, IDIAP: Hynek Hermansky, LIN: Frank Ohl

DIRAC's partners:

IDIAP · Carl von Ossietzky Universität Oldenburg · Katholieke Universiteit Leuven · mp · OHSU · IfN · ETH Eidgenössische Technische Hochschule Zürich Swiss Federal Institute of Technology Zurich

In the DIRAC Newsletter we will take the opportunity to present you each of the DIRAC partners along with its affiliate, the university of Maryland. They will share with us their interests and strength to DIRAC's research.

Please visit our website to have already some information about them.

Detection and identification of Rare Audio visual Cues (DIRAC) is an Integrated project funded by the EC's 6th Framework Program, managed by IDIAP (CH).

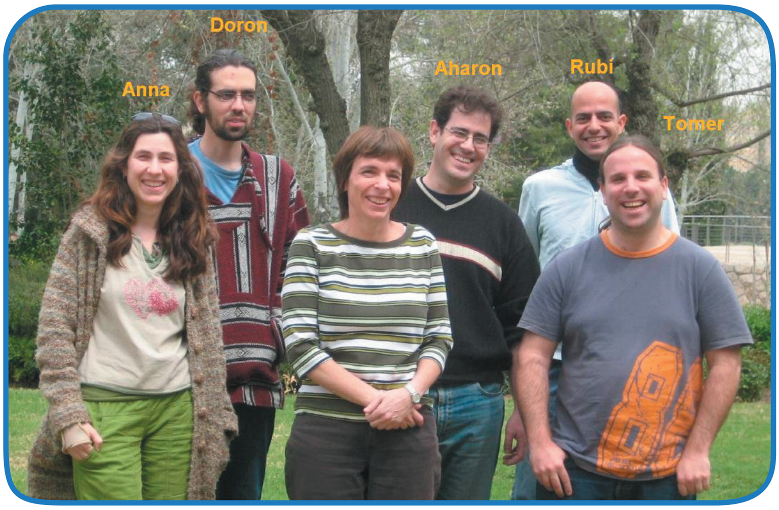
IDIAP · Information Society Technologies

Hebrew University

At the Hebrew University, the focus of research is learning and categorization. Most of our work, which is outlined below, follows one of two tracks: the design of algorithms, and human studies. In DIRAC's project, we will focus our research on the following three workpackages: Signal and Features (WP1), Visual representations (WP3), Learning and categorization (WP4). Five people, in addition to the PI (Daphna Weinshall) are involved at various level in this research.

1) Algorithm design is either aimed at categorization directly, or at aiding categorization by providing better data for training and testing. Some of the questions relate to the computation of rare events, such as the detection of motion discontinuities. A closely related question concerns the relevance of features to the task, which is the main focus of the work on learning distance functions.

More specifically:



Doron Feldman is developing motion segmentation algorithms. Segmentation is based on motion discontinuities only (not intensity edges). A great effort is put on precise localization, as motion flow computation is least accurate exactly in the areas of motion discontinuities. Precise automatic segmentation can aid later categorization processes that attempt to build object's representation automatically. This work is closely related to the focus of WP1, and we plan on integrating this work with parallel efforts on tracking and visual object detection.

Aharon Bar-Hillel is working on object class recognition and localization, where training is done using un-segmented images. We have developed an efficient algorithm to learn object classes, represented by star-like Bayesian networks. Hopefully we will integrate this work, and in particular our machinery for the efficient computation of semi-semantic structures, with the work on object detection done at ETH within WP1 and WP3.

Our current efforts focus on the automatic extraction of object class hierarchies, and the relation to the hierarchies observed in human cognition, where objects at the "basic level" of categorization are given some primacy. At a later stage, we may use classical machine learning and pattern recognition machinery, such as support vector machines, for the classification of sub-ordinate categories within the parent basic category (such as the distinction between types of motorbikes). This kind of feature sharing will be used to identify related objects, and build a hierarchy of categories. It would be interesting to see whether our hierarchy resembles the hierarchy of natural objects in human categorization. Another interesting question is whether we observe an intermediate level primacy, such as observed in human categorization.

Tomer Hertz is working on the learning of distance functions and kernels, using equivalence constraints between pairs of data-points as training data. The approach is semi-supervised, using unlabeled data in addition to some partially labeled data. Our goal is to improve clustering and categorization, by providing an improved distance measure that can be subsequently used by any clustering algorithm.

An earlier focus of this work has been the identification of relevant and irrelevant features, in the context of a specific task. The identification of rare events, which are task relevant, may be quite helpful in the successful identification of such features. Our current efforts are aimed at the design of a new kernel learning algorithm. Kernel learning is an adaptation mechanism that can be incorporated into other kernel-based classification methods. Our initial experiments show the benefit of kernel learning in a learning-to-learn scenarios, or when faced with the problem of small sample size.

2) Our research involving human studies probes human categorization and the perception of rare events:

Anna Sorkin is developing a Virtual Reality (VR) environment which is designed to be viewed via a Head Mounted Display (HMD) device, to give the participants a sense of immersion and a greater sense of reality. Participants in this study explore a multi-modal virtual city. Along the way they are presented with a number of inconsistent (or rare) events, characterized by some unusual color, unusual location, or un-usual sound. The participants are asked to identify anything unusual. Our goal is to characterize the kind of circumstances that prevent people from identifying correctly unusual events. We also study the saliency of the various cues, and its relation to successful detection.

Rubi Hammer studies human categorization jointly with Tomer Hertz. In our earlier work we have studied human categorization from equivalence constraints, showing that people use positive constraints (relations between objects from the same class) more effectively and more intuitively that negative constraints (relations between objects from different classes). Our first goal is to study the same phenomenon in young children, age 5-8, in order to probe a hypothesized connection between our observations and the hierarchical structure of natural objects in human categorization. We will also study eye movement in such tasks, to see if we can observe different strategies employed by participants who are using positive vs negative constraints. Finally, we may extend the study to the auditory domain, to see if similar behavior is observed in the auditory modality.

The difference between positive and negative constraints may have implications to animal categorization studies done by the Magdeburg group within WP4. We may want to study if there are differences in performance when animals are trained with the two different types of constraints. In a more interesting direction we may focus on cells which respond to specific features, and ask whether they respond differentially when the features are relevant or irrelevant to the task.

In summary, the above describes the current research activities, and questions which will direct our work within DIRAC in the next 1-2 years.

Issue N° 1 / April 2006

Editor : Céline Aymon, celine.aymon@idiap.ch

Newsletter

Introduction to DIRAC

Today's computers can do many amazing things but there are still many "trivial" but important tasks they cannot do well. Machines are trained to detect expected signals and rely heavily on large amounts of training data that are not available for rare but high-utility classes. Therefore, machines often fail in detection and recognition of such information-rich unexpected rare events.

In nature, surviving species are the ones which are capable of successfully detecting and identifying deviations from "normal". It is impressive how are humans able to find structure and form concepts like object classes and action or speech sound categories, without much supervision. They are also able to draw conclusions from subtle variations, while at the same time discarding much more substantial ones if not relevant. Clearly, for machine detection and classification of complex and/or rare events we need different, more structured models which rely less on training data.

The innate invariance-extracting cognitive mechanisms that evolved in successful species serve as an inspiration in DIRAC, which aims at designing and developing of an environment-adaptive artificial cognitive system, which will detect, identify and classify information-rich, rare events from data derived by multiple active information-seeking audio-visual sensors. This means, to move from interpretation of all incoming data to reliable rejection of non-informative inputs, from passive acquisition of a single incoming stream to active search for the most relevant information in multiple streams, and from a system optimized for one static environment to autonomous adaptation to new changing environments.

The detection and identification of information-rich rare and relevant events in audio-visual information streams was chosen as the focus of the project. Designing and building the envisaged systems entail multidisciplinary research on a broad set of cognitive principles, including:

Acquisition of sensory stimuli
- acquisition of visual and auditory stimuli in biology
- visual and audio sensors, appropriate to extract specific types of information (adaptive binaural microphones, omni-directional video sensors)

Representations of audio and visual signals
- peripheral and central audio and visual neural processes, with emphasis shifting from single to multiple targets, and from static to dynamic/variable scenes, categorization and adaptation (plasticity) in cortical processes

Learning and categorization
- use of context, bio-consistent combined top-down and bottom-up adaptive models, and task-dependent supervised, weakly supervised, and even unsupervised categorization

Sensory information fusion
- sensory information fusion in natural biological systems
- multi-stream processing with information fusion (creation of sub-streams, reliability of information in sub-streams, linear and nonlinear fusion techniques)

DIRAC's multidisciplinary challenge

An intensive cooperation between cognitive and engineering sciences would clearly be beneficial. Unfortunately, examples of cooperation between cognitive and engineering sciences are still few and far between. We also realized that besides the current boundary between cognitive sciences and engineering, yet another dangerous and unnatural divide exists between auditory and visual research, in spite of similarities between these two sensory modes.

Under DIRAC, we have assembled a group of scientists with the explicit goal to overcome these boundaries in order to advance the art of information extraction from audio and visual cognitive signals. Researchers in the project combine expertise in physiology of mammalian auditory and visual cortex and in audio/visual recognition engineering. All DIRAC's partners agree that understanding cognitive functions and emulating the selected ones in information extraction by machine is a way to more efficient technology. In contrast to the open loop character of most current engineering solutions, and consistent with the closed perception and perception-action loops found in natural cognitive systems, our emphasis will be on emulating structures and functionalities observed in processing of cognitive data in nature

1. with active elicitation and acquisition of relevant data based on feedback from the other modules of the system,
2. with autonomous adaptation to new and changing environments, where the context constrains the models which are applied and limits the search space, thus allowing for solving ambiguities, compensating for missing data and for correcting errors, and
3. with reliable rejection of non-informative data.

The general model for the collaboration between cognition-oriented research groups and engineering is that the outputs from cognitive studies guide the engineering strategies and the successful engineering strategies validate physiological observations. In the course of this research, cognitive and engineering sciences will mutually reinforce each other. We target new engineering systems that will implement available cognitive knowledge for their better functioning. However, progress in engineering technologies can be expected to also foster new insights in the study of cognition.

Besides the fulfillment of our scientific and technological goals, another important result of DIRAC will be a real and long lasting collaboration between cognitive sciences and engineering, resulting in a new generation of researchers, trained in both the cognitive disciplines and in engineering.

DIRAC will significantly contribute to moving the art of audiovisual machine recognition from the classical signal processing/pattern classification paradigm to human-like information extraction, thus forming foundation for a new generation of efficient cognitive information processing technologies.

News

Part of the work in Dirac is on the interpretation and 3D reconstruction of city scenes. The goal is to build wearable, interactive devices that understand what's going on in such complicated environments. The 3D city models will provide the necessary context to safely guide people around in such hectic environments. Even if in its initial stages, this work has already won Dirac researchers an oral presentation at the prestigious IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) 2006, in New York. Nico Cornelis, Kurt Cornelis, and Luc Van Gool will present their work `Fast compact city modeling for navigation pre-visualisation' in the big apple.

Present and Past Events

CogSys II *2006-04-12 and 13*

Radboud University Nijmegen, NL

CogSys
Cognitive Systems

The CogSys II conference is organized to share the progress that is being made in the research community around the sponsored projects within the area of Cognitive Systems (area FP6 2003-IST-2, call IST-2002-2.3.2.4).

From the Mission Statement of the Directorate E Unit E5 – Cognition: The area of Cognitive Systems aims to develop "artificial cognitive systems than can interpret information (images, text, speech, video footage) and other forms of sensor data, and act purposefully and autonomously towards achieving goals. These systems should learn and develop through individual or social interaction with their environment. The work should provide an enabling technology that applies across domains such as natural language understanding, image recognition, automated reasoning and decision support, robotics and automation, sensing and process control, and complex real-world systems. The work should furthermore borrow insights from the bio-sciences, and yield innovative insights about perception, understanding, interaction, learning and knowledge representation".
http://www.socsci.ru.nl/CogSys2/index.html

euCognition

The European Network for the Advancement of Artificial Cognitive Systems.

On January 1st 2006 the European Network for the Advancement of Artificial Cognitive Systems project started for a duration of 36 months.

The principal goal of the euCognition network is to advance the emerging area of cognitive systems by fostering truly inter-disciplinary interaction, spanning a greater spectrum of people, perspectives, and applications than is possible in a single R&D project.

The objective is to support the research community that is already involved in FP6 projects in cognitive systems and to help other individuals from research institutes and companies become involved in this initiative. This will be achieved by facilitating interaction between projects and collaboration between individuals on a variety of fronts, ranging from workshops, conferences, courses, exchanges of staff and students, development & dissemination of training material, access to development platforms, research planning, and the creation of an extensive dynamic web-based repository resource to facilitate research, education, and outreach to the greater community.

The ultimate aim is to leverage added-value from existing work through interaction and to use this to encourage further contributions from new participants. A key objective of the network is to foster interaction between all the many different scientific sectors involved in this multi-disciplinary area and to help create truly inter-disciplinary perspectives.

For more information on this Network of excellence, please visit its website at:
http://www.eucognition.org

euCognition
www.euCognition.org

DIRAC's Publications

(http://www.diracproject.org/publications/)

Prof Hynek Hermansky presented orally and with a poster the DIRAC project at the CogSys II conference. Please find below an abstract of his presentation:

Detecting Information-rich Events with Multiple Sensors

Hynek Hermansky
CogSys II, Radboud University, Nijmefen, NL, April 12-13 2006

Abstract: Today's computers can do many amazing things but there are still many "trivial" but important tasks they cannot do well. In particular, current information extraction techniques perform well when event types are well represented in the training data but often fail when encountering information-rich unexpected rare events. The talk discusses key challenges and some possible research directions of the newly awarded EC Integrate Project DIRAC that addresses this crucial machine weakness.

The project aims at designing and developing an environment-adaptive autonomous artificial cognitive system, which will detect, identify and classify information rich, rare events from data derived by multiple active information-seeking audio-visual sensors. This means, among other things, to move from interpretation of all incoming data to reliable rejection of non-informative inputs, from passive acquisition of a single incoming stream to active search for the most relevant information in multiple streams, and from a system optimized for one static environment to autonomous adaptation to new changing environments.

Partners in the project combine expertise in physiology of mammalian auditory and visual cortex and in audio/visual recognition engineering with the aim to move the art of audiovisual machine recognition from the classical signal processing/pattern classification paradigm to human-like information extraction, thus forming foundation for a new generation of efficient cognitive information processing technologies.
Type of presentation: Oral and poster

Second-order Separation of Multidimensional Sources with Constrained Mixing System.

J. Anemüller
Proceedings of the sixth International Conference on Independent Component Analysis and Blind Signal Separation (ICA), pages 16-23, Charleston, SC, March 5-8 2006.

Abstract:
The case of sources that generate multidimensional signals, filling a subspace of dimensionality K, is considered. Different coordinate axes of the subspace ("subspace channels") correspond to different signal portions generated by each source, e.g., data from different spectral bands or different modalities may be assigned to different subspace channels. The mixing system that generates observed signals from the underlying sources is modeled as superimposing within each subspace channel the contributions of the different sources. This mixing system is constrained as it allows no mixing of data that occurs in different subspace channels. An algorithm based on second order statistics is given which leads to a solution in closed form for the separating system. Correlations across different subspace channels are utilized by the algorithm, whereas properties such as higher-order statistics or spectral characteristics within subspace channels are not considered. A permutation problem of aligning different sources' subspace channels is solved based on ordering of eigenvalues derived from the separating system. Effectiveness of the algorithm is demonstrated by application to multidimensional temporally i.i.d. Gaussian signals.