Issue N° 6 / July 2007

Editor : Patrick Tacchini, patrick.tacchini@gmail.com

DIRAC
Detection and Identification of Rare Audiovisual Cues

Newsletter

Contents

News

BCI Meets Robotics: Challenging Issues in Brain-Computer Interaction and Shared Control
November, 19-20, 2007
KU Leuven, Belgium

Journal Paper

• Properties of Shape Tuning of Macaque Inferior Temporal Neurons Examined Using Rapid Serial Visual Presentation, De Baene, W., Premereur E., & Vogels R Journal of Neurophysiology, 97, 2900-2916

Cover Story

Brain mechanisms of audio-visual category transfer for processing of rare events

Imagine you take a relaxing walk through a forest and let your mind go. But all of a sudden you have noticed a little something moving between the trees. It was too short to be readily identified and the whole event was unexpected. What you, a "natural" cognitive system, now can do so much better than any existing "artificial" cognitive system, is making sense of that event. Making sense of a rare unexpected event might imply you take it as a sign of potential danger that is to be avoided. Conversely, you might reason it was a harmless bird having jumped from one twig to another. You might decide to take caution, simply ignore or further investigate. In natural situations the potential interpretations of a rare event and appropriate actions to take are innumerable. Natural cognitive systems have evolved optimized mechanisms of dealing with the rare and unexpected that are currently unparalleled by anything artificial systems can do. One of the research tracks of DIRAC investigates brain mechanisms of rare event processing in natural cognitive systems (animals and humans) to ultimately make them available for artificial cognitive systems also.

It is well known that many animals and humans respond to detection of a rare event with a number of stereotyped responses, e.g. an orientation response re-directing the subject's sensory organs (eyes, ears, nose, vibrissae, etc.) towards the putative localization of the event's origin. Also, a number of stereotyped brain processes are known in physiological research to accompany detection of rare events, like the "P300" component of event-related brain potentials or the "mismatch negativity" and others. Researches at the Leibniz Institute for Neurobiology in Magdeburg, Germany, investigate brain mechanisms in animals and humans during rare event processing taking yet a different viewpoint: The question is asked, what the mechanisms are, following the detection of a rare event, by which the animal or human subject makes sense of that event. Often, especially when nothing relevant can be identified as a consequence of that event, an animal or human might simply ignore it. From an evolutionary perspective, ignoring a rare event might be a most economic and favourable strategy of dealing with such a situation. But we all have experienced those cases in which someone (including animal subjects) might have shown "the right instinct" with respect to a particular rare event, although this was not based on a clear-cut identification and categorization of that event. This is one realm of cognitive performance where machines still perform so much poorer than humans or animals. What makes rare event processing difficult from an engineering point of view is that the event is typically not well represented in "world model" of a cognitive system; it is simply "not in the data base".

Fig. 1. Audio-visual category transfer to foster appropriate rare event processing. A Mongolian gerbil, prepared for multichannel electrophysiological recording of brain activity, has perceived a fast blinking LED as a rare unexpected visual event for the first time in his life. His crossing of a little hurdle in his cage indicates that he transfers the abstract concept of rhythm categories (fast rhythm signal potential unpleasant current on the floor grid, slow rhythms signal safety) from the auditory modality (using tone beeps) to the visual modality.

to be continued on page 3

DIRAC c/o IDIAP Research Institute, Simplon 4, P.O. Box 592, CH-1920 Martigny
www.diracproject.org
info@diracproject.org - www.diracproject.org

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).

DIRAC
Detection and Identification of Rare Audiovisual Cues

Issue N° 6 / July 2007

Editor : Patrick Tacchini, patrick.tacchini@gmail.com

Newsletter

Katholieke Universiteit Leuven, Department ESAT/PSI-VISICS

The Centre for the Processing of Speech and Images (PSI) (http://www.esat.kuleuven.be/psi) is part of the Electrical Engineering Department (ESAT) of the Katholieke Universiteit Leuven, Belgium. PSI carries out research on the acquisition, processing and generation of audio-visual signals with the application of this research to real world problems. PSI has been involved in several European research projects, has good contacts with research groups world-wide, works together with several national and international companies and has been the starting point of the spin-off companies ICOS (visual inspection systems), Eyetronics (3D modelling), Medicim (Medical Imaging) and GeoAutomation (mobile mapping). Together, these spin-offs employ more than 800 people by now. PSI is divided into 3 subgroups (Industrial Image Processing, Medical Image Processing and Speech Processing) that work closely together in several projects. The industrial image processing group VISICS is involved in the DIRAC project.

VISICS (VISion for Industry Communications and Services) is the biggest group within PSI, with over 25 staff members. The group focuses on several topics, which include 3D acquisition, remote sensing, visual robot navigation, texture analysis, virtual and augmented reality, and object recognition and categorization. VISICS has extensive experience with European projects and has coordinated several in the past. Under the 6FP VISICS was involved in CLASS (STREP), Reveal-This (STREP), DIRAC (IP), EPOCH (NoE), and Pascal (NoE).

KUL team

Recent accomplishments of the VISICS-team are a webservice that turns uploaded images into 3D models of the scene they contain, available for free download (http://www.arc3d.be). This webservice is based on a longstanding research effort by the group on the 3D reconstruction from uncalibrated image sequences. This work has won the group the David Marr Prize in 1998. Furthermore VISICS has experience in the visual recognition of objects and object classes. It has laid the foundations for affine invariant regions and also has proposed some novel interest point detectors and descriptors, that underlie several of the existing object recognition approaches.

For DIRAC, the VISICS group contributes expertise on 3D scene modeling and camera tracking, mainly in collaboration with the team in Prague. VISICS also brings in its expertise in the area of object class detection. It has e.g. already developed pedestrian and car detectors, in close collaboration with ETH Zürich, which it plans to improve. The main contribution however, will be in feeding such detections back into the process of scene understanding in order to make this old dream of computer vision possible, and this with a single moving camera. Such processes, where semantic interpretations guide lower-level processes were coined cognitive loops. Cognitive loops form an important part of DIRAC's toolbox to arrive at the overarching goals of detecting rare events and at the same time train systems with small sample sets.

Building stronger and stronger models of the environment is seen as key to the detection of more and more rare events, where the semantic level of what is to be considered 'rare' is systematically increased throughout the project.

Together with the DIRAC partner ETH Zürich, VISICS received the Best Video Award for their work on cognitive loops which was presented at the International Conference on Computer Vision and Pattern Recognition, June 17-22 2006, New York, USA. The awarded video demonstrates how a city modeling module devised by VISICS is combined with a car recognition module devised by ETH Zürich so that each module feeds information to the other module, allowing the combined system to overcome the weaknesses of each individual module. The implementation of such a cognitive loop between modules led to a drastic increase in robustness, efficiency and accuracy of the resulting 3D city model.
After this most encouraging start of the work, the same philosophy was extended to include pedestrian detection and most recently, people and car tracking. A recent joint paper by VISICS and ETH Zürich has won the Best Paper Award at the International Conference on Computer Vision and Pattern Recognition (CVPR) in Minneapolis (USA), 18-21 June, 2007. The fact that the paper was selected among more than 1200 submissions to the conference testifies to the relevance attributed to this novel line of research by the computer vision community.

VISICS also leads the Training and Education Program which goal it is to ensure that the knowledge and expertise generated in DIRAC is dissipated to a wider scientific community. This year DIRAC will have its very first summer workshop. To organize this event DIRAC has joined forces with IST-project CoSy because of its similarity in goals: to investigate the cognitive processes of learning and understanding environments using data as retrieved by different sensor inputs.
The summer workshop is sponsored by the PASCAL Network of Excellence and will carry the name "Summer Workshop on Multi-Sensory Modalities in Cognitive Science". It is scheduled to take place August 25-29 in Studienzentrum Gerzensee, Switzerland.

The VISICS team is currently composed of:
Luc Van Gool, Luc Van Eycken, Maarten Aerts, Mario Ausseloos, Geert Caenen, Nico Cornelis, Bert DeKnuydt, Annitta Demessemaeker, Stefaan De Roeck, Michael De Smet, Chris Engels, Paul Konijn, Thomas Koninckx, Xinhai Liu, Markus Moll, Wim Moreau, Alexander Thomas, Tinne Tuytelaars, Filip Van den Borre, Johan Van Rompay, Christoph Strecha, Frank Verbiest, Maarten Vergauwen, Patricia Waege, Geert Willems, Egemen Özden, Qiong Yang

VISICS is looking for additional researchers to reinforce its activities! People interested should not hesitate to contact Luc Van Gool (vangool@esat.kuleuven.be).

DIRAC
Detection and Identification of Rare Audiovisual Cues

Issue N° 6 / July 2007

Editor : Patrick Tacchini, patrick.tacchini@gmail.com

Newsletter

Brain mechanisms of audio-visual category transfer for processing of rare events

Hence, what can a cognitive system do at this point beyond merely signalling that such an event has been detected?

Antje Fillbrandt and Frank Ohl from the Leibniz Institute for Neurobiology in Magdeburg have designed an experiment with a particular rodent species, the Mongolian gerbil, to investigate brain activity underlying processing of a rare unexpected event (Fig. 1). In their experiment, the rare event is an LED blinking at a fast rate (5 Hz) for a few seconds. A naïve gerbil would show a behavioural orientation response and corresponding brain potentials in visual and central brain areas and might quickly return to whatever it has done before the event came on. This is different from behaviours and brain processes that can be observed in gerbils who had previously experienced that repetitive auditory stimuli (e.g. beeping tones) can be grouped into two categories: Auditory stimuli that are repeated at a fast rate (5 Hz) signal an unpleasant electric current through the metal floor grid following the tone beep series, while stimuli repeated at a slow rate (<1 Hz) signal safety. The animals also learn that they can totally avoid the unpleasant electric stimulus if they jump over a little hurdle separating two compartments in the cage before the tone beep series stops. How does this experience change the situation during the processing of rare unexpected visual event?

First, let us note that this situation is in many respects similar to the above sketched natural situation of rare event processing by an animal or human subject in the forest. The blinking LED is truly a rare and unexpected event (the animal subject has never experienced it before in its life). It is therefore not clear whether this event might have a special meaning for the animal or not. In particular it is not clear whether any inferences from what has been learned about auditory stimuli might be drawn for this situation. It turns out, however, that some gerbils seem to make such an inference and show the hurdle crossing response while other individuals don't. The first group of gerbils transfers the rhythm categories (slow vs. fast) learned in the auditory modality to visual stimuli, the second group doesn't. Note that at this point (when this behaviour of the gerbils does not yet yield any particular pleasant or unpleasant consequences) one cannot speak of appropriate or inappropriate behaviour; the situation merely demonstrates different strategies that different subjects employ. While the experimenter can of course set the rules which strategy would be optimal, the more interesting question is: What are the brain mechanisms that gave rise to transferring or not transferring the categories established in one sensory modality to the other when information about the appropriate strategy is not available, i.e. when the rare event is truly unclassifiable?

Fillbrandt and Ohl have analyzed the cross-talk between auditory and visual cortex in various stages of the experiment (Fig. 2). They used a technique, first developed in econometrics, and now used in many engineering fields, to study the direction of causation between two (or more) time series of data (Granger causality analysis). This technique yields a measure (normalized directed transfer function) of how much the neuronal activity measured in one brain area might "cause" activity in another brain area. In addition, this measure is expressed as a function of frequency, identifying potential preferred frequency bands in which this directed interaction takes place. Information about preferred frequency bands will foster the mechanistic understanding of the interaction on a biophysical basis in later states of the data analysis. These functions derived from first data obtained in the experiment are displayed in the panels of Fig. 2 as red and black curves indicating the amount of directed transfer from auditory to visual cortex or vice versa, respectively. Data from naïve gerbils indicate that no significant directed transfer, and hence no cross-talk, between auditory and visual cortex exist; this is true for both auditory and visual stimulation (top row). In contrast, gerbils that have formed the auditory categories but do not transfer them to visual stimuli show significant directional cross-talk (bottom row). The direction of this interaction is always from the directly exited cortex (auditory cortex during auditory stimulation, visual cortex during visual stimulation) to the corresponding not-excited cortex and does not show conspicuous peaks at particular frequencies. This is the characteristic of a passive system. Gerbils that do transfer the auditory categories to visual stimuli show a clear directional interaction with a clear peak in the gamma range (30-60 Hz) of brain activity at least from the cortex of the excited sensory modality to the other cortex but sometimes also in both directions (left and right, middle row).

The described experiment is a "laboratory version" of audio-visual category transfer that can be invoked by natural cognitive systems to solve the problem of rare event processing in one modality. The analysis is the first description of physiological correlates of such a process. Further studies, in collaboration with partners at IDIAP, Oldenburg University, Hebrew University of Jerusalem and Oregon Health and Science University aim at, confirming the preliminary results, revealing the details of its mechanistic implantation in natural brains, providing an algorithmic formulation of neuronal processing strategies and making them available for artificial cognitive systems.

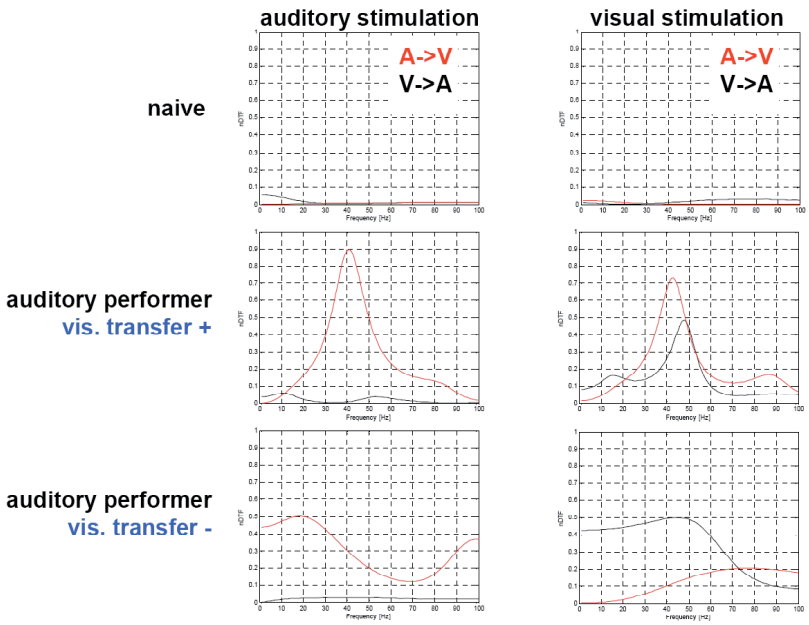

Fig. 2. Analysis of electrophysiological data during audio-visual category transfer. Representative data about directional brain interaction from auditory to visual cortex (A->V) or from visual to auditory cortex (V->A) in various phases of the experiment. For explanation see text

DIRAC
Detection and Identification of Rare Audiovisual Cues

Issue N° 6 / July 2007

Editor : Patrick Tacchini, patrick.tacchini@gmail.com

Newsletter

DIRAC's Conference Call/ Events

BCI Meets Robotics: Challenging Issues in Brain-Computer Interaction and Shared Control November, 19-20, 2007 KU Leuven, Belgium

www.maia-project.org/workshop-2007.php

Last years have witnessed advances in Brain-Computer Interfaces (BCI), but how far is this new field from controlling robotics devices?

The goal of the workshop is to introduce recent advances in brain-computer interfaces on the one hand, and on shared control and task recognition on the other. This workshop will give a new perspective on how humans and men cooperate to fulfill a challenging task. The concept of adaptive shared autonomy will be introduced and its relevance for BCI applications will be illustrated.

The presentations will consist of a series of invited talks and poster presentations. Also, the European MAIA project will report their achievements in non-invasive brain-controlled wheelchairs.

Abstract Submission

Please send a one-page abstract (including figures and references, no less than 200 words) to workshop@maia-project.org

Posters will be selected depending on relevance to the workshop topic, quality, and novelty.

Important Dates

Deadline for abstract submission: 21/09/2007

Notification of acceptance: 05/10/2007

Deadline for early registration: 26/10/2007

Conference dates: 19-20/10/2007

Organizing Committee

Prof. José del R. Millán, IDIAP Research Institute, Martigny,

Switzerland (co-chair)

Prof. Marnix Nuttin, Katholieke Universiteit Leuven, Belgium (co-chair)

Prof. Maria Grazia Marciani, IRCCS Fondazione Santa Lucia, Rome, Italy

Dr. Sara Gonzalez Andino, Geneva University Hospital, Switzerland

Prof. Fabio Babiloni, University of Rome "La Sapienza", Italy

Ac.Prof. Kimmo kaski, Helsinki University of Technology, Finland

For more information, please visit
www.maia-project.org/workshop-2007.php

DIRAC's Publications

Journal Paper

Properties of Shape Tuning of Macaque Inferior Temporal Neurons Examined Using Rapid Serial Visual Presentation,
De Baene, W., Premereur E., & Vogels R
Journal of Neurophysiology, 97, 2900-2916

We used Rapid Serial Visual Presentation (RSVP) to examine the tuning of macaque inferior temporal cortical (IT) neurons to 5 sets of 25 shapes each that varied systematically along pre-defined shape dimensions. A comparison of the RSVP technique using 100 ms presentations with that using a longer duration showed that shape preference can be determined with RSVP. Using relatively complex shapes that vary along relatively simple shape dimensions, we found that the large majority of neurons preferred extremes of the shape configuration, extending the results of a previous study using simpler shapes and a standard testing paradigm (Kayaert et al. 2005). A population analysis of the neuronal responses demonstrated that, in general, IT neurons can represent the similarities among the shapes at an ordinal level, extending a previous study that used a smaller number of shapes and a categorization task (Op de Beeck et al. 2001). However, the same analysis showed that IT neurons do not faithfully represent the physical similarities among the shapes. The responses to the two-part shapes could be predicted, virtually perfectly, from the average of the responses to the respective two parts presented in isolation. We also showed that IT neurons adapt to the stimulus distribution statistics. The neural shape discrimination improved when a shape set with a narrower stimulus range was presented, suggesting that the tuning of IT neurons is not static but adapts to the stimulus distribution statistics, at least when stimulated at a high rate with a restricted set of stimuli.

Conference Papers

To see the abstract of conference papers, please visit www.diracproject.org/publications

Hierarchical Neural Networks Feature Extraction for LVCSR system,
Fabio Valente, Jithendra Vepa, Christian Plahl Christian Gollan, Hynek Hermansky, Ralf Schl.
Interspeech 2007, Antwerpen, Belgium, August 27-31 2007

Multi-stream Features Combination based on Dempster-Shafer Rule for LVCSR System,
Fabio Valente, Jithendra Vepa, Hynek Hermansky,
Interspeech 2007, Antwerpen, Belgium, August 27-31 2007.

Wide-Band Perceptual Audio Coding based on Frequency-domain Linear Prediction,
Petr Motlicek, Vijay Ullal, Hynek Hermansky,
International Conference in Acoustic, Speech and Signal Processing (ICASP 2007). Honolulu, Hawai, USA, April 15-20 2007,

Dynamic 3D Scene Analysis from a Moving Vehicle,
B. Leibe, N. Cornelis, K. Cornelis, L. Van Gool,
IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'07), Minneapolis, Minnesota, June 18-23, 2007,

Combination of Acoustic Classifiers based on Dempster-Shafer Theory of evidence,
Fabio Valente and Hynek Hermansky,
International Conference in Acoustic, Speech and Signal Processing (ICASP 2007). Honolulu, Hawai, USA, April 15-20 2007

Subordinate Class Recognition Using Relational Object Models
Bar-Hillel A., Weinshall D.
Proc. of Neural Information Processing Systems (NIPS), Vancouver, Canada, December 4-9, 2006

Monocular Tracking with a Mixture of View-Dependent Learned Models
Jaeggli T., Koller-Meier E., Van Gool L.,
IV Conference on Articulated Motion and Deformable Objects (AMDO 2006), July 2006.

Multi-label Image Segmentation Via Max-sum Solver
B.Micusik B., Pajdla T.
IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'07), Minneapolis, Minnesota, June 18-23, 2007

A Minimal Solution to the Autocalibration of Radial Distortion
Kukelova Z., Pajdla T.
IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'07), Minneapolis, Minnesota, June 18-23, 2007

Posters

To see the abstract of conference papers, please visit www.diracproject.org/publications

Detection of Inconsistent Audio-Visula Events in Virtual Reality
Anna Sorkin, Avi Peled, Daphna Weinshall
European Conference on Visual Perception

Category Learning from Positive and Negative Pairwise Relations
Rubi Hammer, Tomer Hertz, Shaul Hochstein, Daphna Weinshall
European Conference on Visual Perception