Issue N° 7 / October 2007

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

DIRAC

Detection and Identification of Rare Audiovisual Cues

Newsletter

Contents

Cover Story

DIRAC summer workshop

August 24, 2007. A group of more than 40 PhD students, postdocs and professors from different backgrounds in computer vision, speech and sound processing, neurophysiology, artificial intelligence, and robotics gathered in Studienzentrum Gerzensee, Switzerland. This off-the-beaten track location about half an hour away from the capital Bern was the venue of a summer workshop on multi-sensory modalities in cognitive science. This five-day event was organised in collaboration with the IST-project CoSy (http://www.cognitivesystems.org) and sponsored by the PASCAL Network of Excellence (***http://www.pascal-network.org/***). Tutorial sessions, group discussions and leisure activities were the three main ingredients of this summer workshop.

The goal of the tutorial sessions is to stimulate cross-fertilization of knowledge between researchers in cognitive science and prepare current students for a range of research in cognitive engineering they may attempt during their careers. Therefore, the tutorial sessions covered the main sensory modalities in cognitive sciences and were given by lecturers well-known in their field of expertise.

Six lecturers were invited by the DIRAC project. Hynek Hermansky discussed several speech processing techniques which are based on evolving understanding of the role of spectrally localized dynamic temporal cues in human auditory perception. A general introduction to visual object recognition from local features was given by Tinne Tuytelaars and Bastian Leibe. Tomas Pajdla built on this lecture to explain the main components of image matching and camera tracking which are useful tools for self localization, scene modeling and recognition. A different viewpoint was presented by Rufin Vogels who discussed the neural mechanisms of visual object recognition and categorization. Jörn Anemuller gave an overview on the fundamentals and applications of perception and processing of spatial sound.

The CoSy project provided four additional lecturers. Benjamin Kuipers described recent progress on tackling the problem how higher-level concepts that make up the foundation of commonsense knowledge (e.g. objects, actions, goals, etc) can be learned from unguided experience at the pixel level. The course given by Matthias Scheutz looked at the nature, role, and utility of building computational models of cognitive functions and introduced the main cognitive architectures. Craig Boutilier provided the participants with a survey of logical representational and computational methods for Markov decision processes. Finally, Frederic

to be continued on page 3

DIRAC c/o IDIAP Research Institute, Centre du Parc, Av. des Prés-Beudins 20, P.O. Box 592, CH-1920 Martigny
info@diracproject.org - www.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° 7 / October 2007

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

Newsletter

OREGON HEALTH AND SCIENCE UNIVERSITY OGI SCHOOL OF SCIENCE AND ENGINEERING

Ironically, we are often confronted with rare and unexpected events – the issue is which rare events are important enough to care about. The question of finding important, unique events is the modus operandi of the team in the Department of Biomedical Engineering at the Oregon Health and Science University. In fact, OHSU, located in the picturesque city of Portland, is itself is a unique, unusual institution that combines avant-garde technology, research and education with the state-of-the art clinical care. The BME department was created by a merger of an Academic Health Center with the Oregon Graduate Institute, an engineering school. In addition to BME, OHSU has strong programs in the Adaptive Systems Lab, the Center for Spoken Language Understanding, the Department of Medical Informatics. A unique symbol of our innovative approach to bridging the chasms among various disciplines is the aerial tram that connects the BME department with the rest of the university.

Its multidisciplinary nature makes OHSU an ideal place to investigate problems at the intersection of engineering, computer science, cognitive science and neuroscience. Although the strengths of this group are predominantly theoretical, the majority of our projects are motivated by practical, socially relevant problems.

OHSU Team

One such issue is the care for elders – one of the most challenging healthcare and economic problem looming to confront the society due to the exponentially increasing proportion elders. A typical elder, Simone, is determined to stay in her old house living independently as long as possible. Her daughter, Eva, an informal caregiver, would like to know that her mom is doing well today, i.e., had a breakfast, took her medications, visited her friends and is no more depressed than in her younger years. But Eva has little time since she has a full time job and takes care of her children. A feasible solution, investigated by OHSU, is based on a system that continuously monitors Simone, assesses her health and cognition, and most importantly, detects any anomalies, and unexpected events.

This monitoring system infers cognitive functionality from sensors that observe unobtrusively elders' behaviors during the course of their normal activities of daily living. Data from passive motion sensors, contact switches, radio frequency identification (RFID) tags and video cameras are used to determine the health and cognitive function of elders and patients with chronic diseases living in their own homes. These data are fused by the state of the art algorithms that learn the elders' normal patterns and are, therefore, poised to detect and identify anomalies and unexpected events. One of the ultimate goals of these efforts is to detect any departure from the normal that would be an indicator of early diagnosis of cognitive decline or other neurodegenerative disease that may have devastating effects. In collaboration with IDIAP and a number of DIRAC partners, OHSU is planning to enhance the current approach by adding a combination of visual and auditory sensors. The current data collection is a part of a large longitudinal study funded by the National Institute for Health.

One behavior that appears to have a rich ensemble of potentially diagnostic characteristics is gait. A good clinician may use her impression of a patient's gait to infer aspects of health. Researchers at OHSU and elsewhere, working on gait and posture, have recently discovered that there is a much tighter relationship between aspects of gait and cognitive function than previously thought. In fact, if a person is asked to perform a challenging cognitive task while walking, he may slow down and even his balance control may deteriorate. It follows therefore, that changes in the gait characteristics of an elder may be significant behavioral markers signaling changes in his cognitive function.

In collaboration with ETH Zurich, the OHSU team has been developing techniques that could extract and estimate relevant gait parameters using unobtrusive sensing techniques such as infrared sensors or video cameras. The unobtrusive nature of the sensor technology, whereby the elders are not required to wear special reflecting devices, are prone to errors due to a variety of environmental conditions such as lighting and shadows. These are surmounted by statistical state estimation techniques in combination with the physical models of the walking individual. The ultimate goal of this research is to develop techniques that would enable us to detect subtle changes and unexpected behaviors before they lead to catastrophic events. This work is funded by several sources, including the Intel Corporation under the Behavioral Assessment and Intervention Commons.

Since the unobtrusive behavior-based approaches have limitations, the OHSU team is exploring ways to use electrophysiological sensing to make further inferences about the cognitive processes. With rapid advancements in sensor technology, the collection of signals like those generated by the cardiovascular system or brain waves on a continuous basis is becoming a real possibility. The sensing and processing of electroencephalographic data enables OHSU's researchers to detect and identify instantaneous cognitive states and cognitive processes. The resulting information is then used to control the amount and kind of information presented by the task so as to improve cognitive performance. This approach to augmenting cognition is the result of a successful collaboration with the Honeywell Corporation.

In a related project, a human observer's visual system is used to detect rare targets in a stream of rapidly presented images, e.g., 10 per second. All the human observer does is to watch these images – the detection of the rare targets is performed automatically, by the analysis of the EEG signals. This project that provides insights into humans' detection of rare events is a result of a close collaboration with Honeywell and Columbia University, and represents a new direction in the field of brain-computer interfaces.

OHSU's theoretical work has been focused on the development of techniques for fusing information from multiple streams without having explicitly labeled training data. OHSU's team, inspired by the physiological and behavioral data bearing on aspects of biological fusion, has been working on a hierarchical, high-dimensional representationof adaptive fusion mechanisms. This work is part of a collaboration with Hebrew University (HUJI) and Leibniz Institute for Neurobiology (LIN).

OHSU's team for these projects includes Misha Pavel, Stuart Hagler, Holly Jimison, Tamara Hayes, Deniz Erdogmus, Catherine Huang and Guoping Wang, with informal contributions from as number of other OHSU's colleagues.

DIRAC c/o IDIAP Research Institute, Centre du Parc, Av. des Prés-beudins 20, P.O. Box 592, CH-1920 Martigny,
info@diracproject.org - www.diracproject.org

DIRAC
Detection and Identification of Rare Audiovisual Cues

Issue N° 7 / October 2007

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

Newsletter

DIRAC summer workshop

CONTINUED FROM PAGE 1

Kaplan described the work done on developmental algorithms which make robots eager to explore their surroundings in search of new things to learn.

Individual group discussion sessions were organised daily after the tutorial sessions to allow participants to exchange ideas on the various topics that were covered earlier that day. Hereto, students were divided into different groups in such a way that each group consisted of a balanced mix of people from different backgrounds and from the two projects, DIRAC and CoSy. A senior researcher was assigned to each group to guide the discussions when necessary. A plenary session was scheduled after the individual group discussions which allowed students to ask questions to the lecturers. The main goal of these group discussion sessions was to identify where the different topics may lead us in the future and to see how they can be geared towards a common goal, encouraging integration between the different disciplines in cognitive sciences.

Leisure time was foreseen in the afternoon and evening so that the workshop participants could try out some of the facilities available at the studycentre (table soccer, billiards, table tennis, indoor gym), go for a stroll or cycling in the rural surroundings or - for the bravest among them - go swimming in the lake nearby. Furthermore, an excursion to Bern was organised as well. The participants were guided through the streets of this UNESCO World Heritage site by two experienced guides who proudly told about the Swiss culture and the history of the capital city. These leisure activities provided the participants with a pleasant environment allowing them to get to know better other people with similar interests and related scientific backgrounds.

More information about the first DIRAC summer workshop can be found at the website ***«http://www.diracproject.org/workshop-2007»***. Participants of the summer workshop who are interested in obtaining the original slides of the presentations can download them from the ftp site ***«ftp://ftp.esat.kuleuven.be/psi/visics/mausselo/DIRAC»***. Pictures taken at the summer workshop can be found at ***«http://www.esat.kuleuven.be/~konijn/gs/»***.

One of the main goals of the DIRAC training program is to foster discussions between researchers from different scientific backgrounds, to see how different branches in cognitive sciences can be brought closer together in order to strive towards a common goal. Therefore, this five-day workshop in Gerzensee is only the first in a series of summer schools which will be established within the DIRAC framework. The second summer workshop is foreseen to take place next year.

The Winner of SAIS best AI Master's Thesis award 2007

DIRAC is proud to announce that Jie Luo, from the Royal Institute of Technology and employee at IDIAP, on the DIRAC project, has won the prize for best AI Master's Thesis 2007 from the Swedish AI Society.

The submitted theses were read and evaluated by a group of expert reviewers.

A combination of four criteria based on maturity of content, scientific value, writing style, and relevance of the content to the AI area, were used to determine the winner. All submitted theses were of very high quality but the two chosen (the one from Jie Luo and the one from Boris Schäfer) stood out when evaluating all criteria.

Please see http://www.sais.se/ for more details on the award.

DIRAC
Detection and Identification of Rare Audiovisual Cues

Issue N° 7 / October 2007

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

Newsletter

DIRAC's Publications

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

Journal papers

Differential Effects of Iontophoretic in Vivo Application of the GABAA Antagonists Bicuculline and Gabazine in Sensory Cortex

Kurt S, Crook JM, Ohl FW, Scheich H, Schulze H

We have compared the effects of microiontophoretic application of the GABAA-receptor antagonists bicuculline (BIC) and gabazine (SR95531) on responses to pure tones and to sinusoidally amplitude-modulated (AM) tones in cells recorded extracellularly from primary auditory cortex (AI) of Mongolian gerbils. Besides similar effects in increasing spontaneous and stimulus-evoked activity and their duration, both drugs elicited differential effects on spectral tuning and synchronized responses to AM tones. In contrast to gabazine, iontophoresis of the less potent GABAA-antagonist BIC often resulted in substantial broadening of frequency tuning for pure tones and an elimination of synchronized responses to AM tones, particularly with high ejecting currents. BIC-induced effects which could not be replicated by application of gabazine were presumably due to the well-documented, non-GABAergic side-effects of BIC on calcium-dependent potassium channels. Our results thus provide strong evidence that GABAA-mediated inhibition in AI does not sharpen frequency tuning for pure tones, but rather contributes to the processing of fast temporal modulations of sound envelopes. They also demonstrate that BIC can have effects on neuronal response selectivity which are not due to blockade of GABAergic inhibition. The results have profound implications for microiontophoretic studies of the role of intracortical inhibition in sensory cortex.

Conference papers

Classification with Postive and Negative Constraints: Theory, Computation and Cognition

Rubi Hammer, Tomer Hertz, Shaul Hochsteing and Daphna Weinshall

2nd International Symposium on Brain, Vision and Artificial intelligence (BVAI), October 10-12 2007, Naples, Italy

We tested the efficiency of category learning when participants are provided only with pairs of objects, known to belong either to the same class (Positive Equivalence Constraints or PECs) or to different classes (Negative Equivalence Constraints or NECs). Our results in a series of cognitive experiments show dramatic differences in the usability of these two information building blocks, even when they are chosen to contain the same amount of information. Specifically, PECs seem to be used intuitively and quite efficiently, while people are rarely able to gain much information from NECs (unless they are specifically directed for the best way of using them). Tests with a constrained EM clustering algorithm under similar conditions also show superior performance with PECs. We conclude with a theoretical analysis, showing (by analogy to graph cut problems) that the satisfaction of NECs is computationally intractable, whereas the satisfaction of PECs is straightforward. Furthermore, we show that PECs convey more information than NECs by relating their information content to the number of different graph colorings. These inherent differences between PECs and NECs may explain why people readily use PECs, while many of them need specific directions to be able to use NECs effectively.

Learning Distance Function by Coding Similarity

Aharon Bar-Hillel and Daphna Weinshall

Proceedings: International Conference on Machine Learning (ICML), July 5-9 2008, Helsinki Finland

We consider the problem of learning a similarity function from a set of positive equivalence constraints, i.e. 'similar' point pairs. We define the similarity in information theoretic terms, as the gain in coding length when shifting from independent encoding of the pair to joint encoding. Under simple Gaussian assumptions, this formulation leads to a non-Mahalanobis similarity function which is efficient and simple to learn. This function can be viewed as a likelihood ratio test, and we show that the optimal similarity-preserving projection of the data is a variant of Fisher Linear Discriminant. We also show that under some naturally occurring sampling conditions of equivalence constraints, this function converges to a known Mahalanobis distance (RCA). The suggested similarity function exhibits superior performance over alternative Mahalanobis distances learnt from the same data. Its superiority is demonstrated in the context of image retrieval and graph based clustering, using a large number of data setsAcoustic Feature Selection for Speech Detection Based on Amplitude Modulation Spectograms

Acoustic Feature Selection for Speech Detection Based on Amplitude Modulation Spectrograms

Denny Schmidt Jörn Anemüller

Fortschritte der Akustik: DAGA 2007, March 19-22 2007, Stuttgart Germany

As acoustic devices possess evermore computing power, signal processing is influenced increasingly by machine learning techniques. E.g., hearing aids detect different listening situations by extracting several features (spectrum, modulations...) and feeding them into a classifier. The question is how to determine features that result in best classification performance and good generalization to new signals. Here, a feature selection strategy for automatic speech/non-speech classification based on the support vector machine algorithm (SVM) is presented. Input for the selection algorithm is the psychoacoustically motivated amplitude modulation spectrogram (AMS) presented by Kollmeier and Koch (1994). Classification is performed using clean speech signals in the speech class and «clean» signals from noise-like acoustic scenes in the non-speech class. Results are presented regarding number of modulation frequencies required for speech detection, corresponding classification accuracy and generalization to novel data. Relevant modulation frequencies for speech detection are identified and related to psychophysical evidence.

Two Minimal Problems for Cameras with Radial Distortion

Z. Kukelova and T. Pajdla

Seventh Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras (OMNIVIS), October 20, 2007, Rio de Janeiro, Brazil

Epipolar geometry and relative camera pose computation for uncalibrated cameras with radial distortion has recently been formulated as a minimal problem and successfully solved in floating point arithmetics. The singularity of the fundamental matrix has been used to reduce the minimal number of points to eight. It was assumed that the cameras were not calibrated but had same distortions. In this paper we formulate two new minimal problems for estimating epipolar geometry of cameras with radial distortion. First we present a minimal algorithmfor partially calibrated cameras with same radial distortion. Using the trace constraint which holds for the epipolar geometry of calibrated cameras to reduce the number of necessary points from eight to six. We demonstrate that the problem is solvable in exact rational arithmetics. Secondly, we present a minimal algorithm for uncalibrated cameras with different radial distortions. We show that the problem can be solved using nine points in two views by manipulating polynomials by a sequence of Gauss-Jordan eliminations in exact rational arithmetics. We demonstrate the algorithms on synthetic and real data.

Posters

States in the Ongoing Cortical Activity Carrying Information in Discrimination Learning of Differential Electrical Stimulation

Deliano M, Scheich H, Ohl FW

Cortical Activity Associated with Discrimination Learning of Auditory, Audio-Visual or Intracortical Electric Stimuli After Appe

Deliano M, Ilango AM, Fillbrandt A, Wetzel W, Ohl FW

DIRAC c/o IDIAP Research Institute, Centre du Parc, Av. des Prés-beudins 20, P.O. Box 592, CH-1920 Martigny,
info@diracproject.org - www.diracproject.org