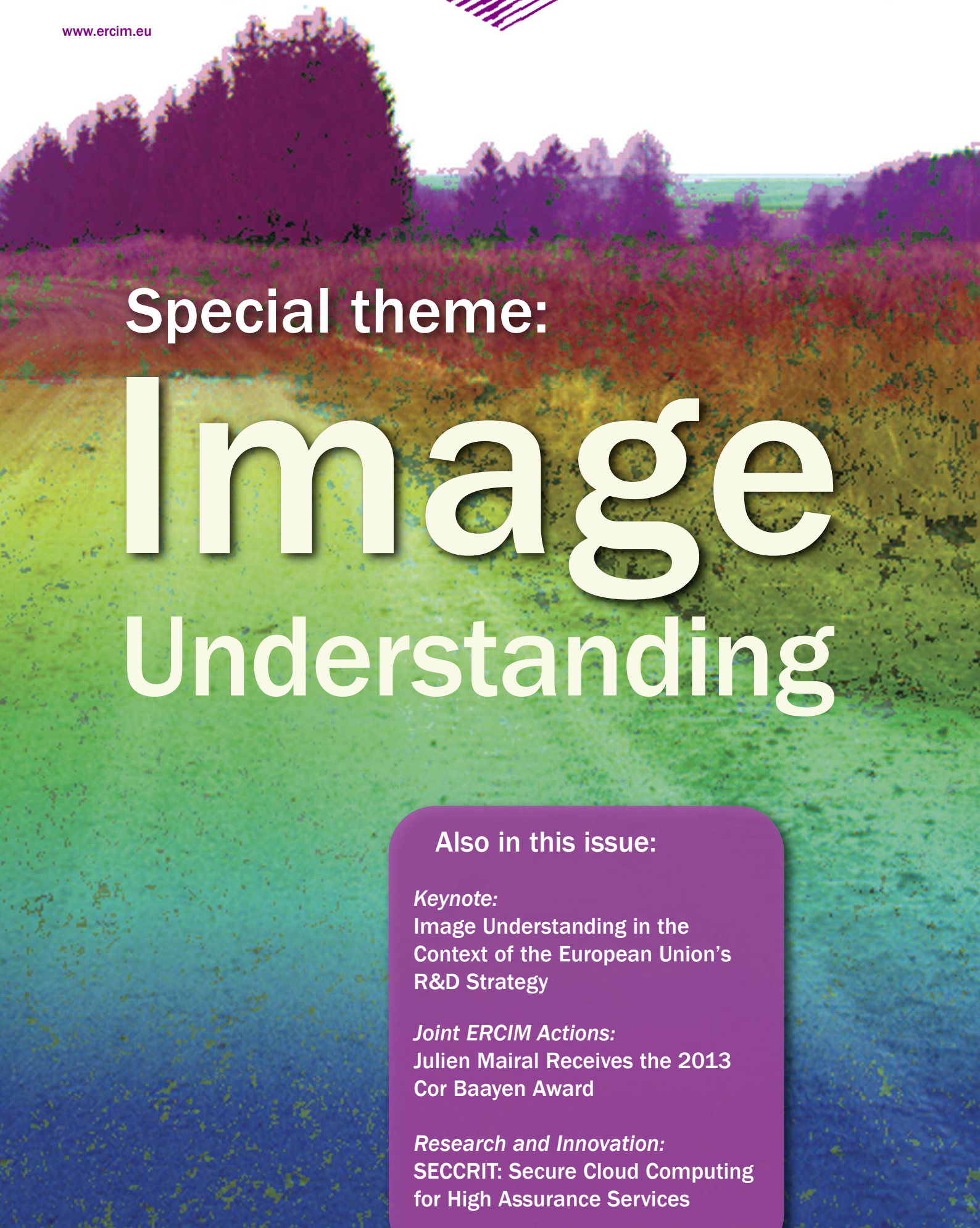
Number 95 October 2013

ERCIM NEWS

www.ercim.eu



Special theme:

Image Understanding

Also in this issue:

ERCIM News is the magazine of ERCIM. Published quarterly, it reports on joint actions of the ERCIM partners, and aims to reflect the contribution made by ERCIM to the European Community in Information Technology and Applied Mathematics. Through short articles and news items, it provides a forum for the exchange of information between the institutes and also with the wider scientific community. This issue has a circulation of about 8,500 copies.

ERCIM News is published by ERCIM EEIG
BP 93, F-06902 Sophia Antipolis Cedex, France
Tel: +33 4 9238 5010, E-mail: contact@ercim.eu
Director: Jérôme Chailloux
ISSN 0926-4981

JOINT ERCIM ACTIONS

KEYNOTE

SPECIAL THEME

The special theme section "Image Understanding" has been coordinated by Michal Haindl, Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, and Josef Kittler, University of Surrey, UK

RESEARCH AND INNOVATION

This section features news about research activities and innovative developments from European research institutes

EVENTS, BOOKS, IN BRIEF

Event Reports

Announcement

Books

In Brief

Joint ERCIM eMobility and MobiSense Workshop

by Desislava Dimitrova and Torsten Braun

The 7th ERCIM Workshop on eMobility, held for the second time together with the MobiSense workshop, took place on 4 June 2013, in St Petersburg, Russia. The workshop venue was the former imperial palace with its fantastic decor and view to the Neva River, which set a special atmosphere for the participants. Like previous editions the workshop was co-located with the 11th edition of the International Conference on Wired/Wireless Internet Communications (WWIC 2013) and was co-funded by ERCIM. The workshop targets research related to the impact of user and device mobility on multiple technical areas in communications such as network architecture, communication protocols and algorithm performance but also on more user-oriented areas such as application development and quality of service/experience. The workshop focused on two major research topics, namely, localization and communication solutions for wireless sensor networks.

The joint workshop included a keynote talk and two technical sessions of peer reviewed and invited papers. The invited talk, given by Prof. Peter Reichl from Telecom Bretagne, addressed less technical but nevertheless highly important research for the development of wireless systems. Prof. Reichl discussed how the relationship between system performance parameters and user satisfaction can be correctly represented by mathematical models. He supported his presentation with an analysis of studies that investigated user satisfaction of a content delivery system.

The first technical section on "Localization" included four papers addressing issues in the area of indoor localization covering a diversity of approaches from dead reckoning to algorithms based on WiFi and cellular signals. In particular, Kun-chan Lan (National Cheng Kung University, Taiwan) discussed filtering schemes that can be applied to improve the localization estimate obtained from motion sensors in personal mobile devices. Dr. Enrica Zola (Polytechnic University of Catalonia, Spain) presented two papers from her Telematics group. The first presentation focused on time-based localization approaches with WiFi signals that can offer feedback on human mobility patterns, while the second talk discussed the importance of such mobility patterns for handovers in cellular networks. Finally, Dr. Desislava Dimitrova (University of Bern, Switzerland) gave an overview of the challenges arising in localization systems with passive signal overhearing, comparing the cases of WiFi and cellular (GSM) signals.

The second technical session on "Wireless Sensor Networks" addressed a broader set of topics related to communication between devices with often limited power supply. For instance, Prof. Jorge Granjal (University of Coimbra, Portugal) discussed the critical topics of security support for wireless sensor networks (WSN) and, in particular, how the challenge of performing energy-heavy authentication and encryption computations with energy-limited

Enrica Zola from Polytechnic University of Catalonia, Spain

Peter Reichl from Technical University Vienna, Austria

sensor nodes can be resolved. He then presented the work of a PhD student, who is proposing a novel middleware that can support machine-to-machine (M2M) communications. The second part of the session included work on a novel energy-efficient MAC Protocol for WSNs by Sarwar Morshed (University of Twente, The Netherlands) and a presentation of the Mobility2.0 project by Dr. Geert Heijenk (University of Twente, The Netherlands). Mobility 2.0 is an EU FP7 project in the framework of "ICT for Green Cars-2012" and targets the development of commuting applications for electrical vehicles taking into account the specific requirements of the end users but also of the vehicle.

The event was accompanied by a social event, which was held in the city of St Petersburg allowing the participants to get familiar not only with the city but also with typical Russian cuisine. We are grateful to the local organizers, under the guidance of Prof. Yevgeni Koucheryavy (Tampere University of Technology, Finland) for their excellent work.

The workshop proceedings (ISBN: 978-3-9522719-4-0) can be downloaded in pdf from the workshop website, where all presentations can also be found, following the corresponding program entry.

Link:
Workshop website:
http://wiki.ercim.eu/wg/eMobility/index.php/Second_Joint_ERCIM_eMobility_and_MobiSense_Workshop

Please contact:
Torsten Braun, E-mobility Working Group chair
University of Bern, Switzerland
E-mail: braun@iam.unibe.ch

IDEALIST - An International ICT Partner Search System and Network of National Contact Points

by Givi Kochoradze

Within the framework programs of the European Commission there exist particular projects that make tangible contributions to the development of European science and technology. One such project is "IDEALIST". IDEALIST, which actually encompasses a cycle of projects, has, since its inception in 1996, established itself as a strong and efficient learning network for the ICT National Contact Points (NCPs). IDEALIST is open to all ICT NCPs and ICT specialists, regardless of their individual status within the network.

The IDEALIST 2014 team.

A total of eight IDEALIST projects have been implemented since 1996: Idealfit; Idealist-East; Idealist5FP; Idealist34; Idealist Extend; Idealist7FP; Idealist2011; and Idealist2014. The first project started with the participation of NCPs from a few members and associated states. The latest IDEALIST project has now reached 85 NCPs from 75 different EU and non-EU countries. IDEALIST is continuing to attract NCPs from the entire world. The National Contact Points (officially appointed by national governments) continue to play a key role in the Framework Programs, assisting grant applicants and recipients in on-going projects. Their tasks are:

- to inform and raise awareness about the funding opportunities of the Framework Programs (FPs);
- to advise and assist potential applicants in the preparation, submission and follow-up of grant proposals;
- to provide support during the execution of projects – especially for management and administrative aspects;
- to gather quick, informal feedback from grant applicants and recipients about perceived difficulties in the implementation of the Framework Program, and hence immediate input to the EC to improve specific aspects of the FP.

IDEALIST is implementing advisory services to provide grant applicants with comprehensive information about ways of participating in the work program itself and about other FP7 and alternative ICT funding programs. To this end, a new IDEALIST project in 2014 will develop and implement a coordination mechanism for strengthening cooperation among ICT NCPs in the FP7 and in the Horizon 2020 program. The aim of the IDEALIST 2014 project is to increase the visibility of NCPs and improve the overall quality of NCP services across the Europe in the ICT area. The objectives are:

- Reinforcement of the ICT NCP network by promoting trans-national cooperation;
- Identification and promotion of good NCP practices;
- Enhancement and sharing of NCP knowledge on ICT research objectives;
- Provision of training and twinning measures tailored to the specific needs of ICT NCPs;
- Raising awareness of NCP services including the implementation of practical initiatives to benefit cross-border audiences, such as a quality checked partner search and a pre-proposal and full proposal service for all kinds of funding applicants;
- Establishment of trans-European networking among FP participants by offering brokerage services around large European events and small targeted brokerage meetings;
- Improved cooperation with other networks, organizations and initiatives;
- To put special emphasis on the transition to ICT in Horizon 2020.

Due to the European Commission's strong interest in strengthening international cooperation, we also expect that efforts will be made to enhance cooperation with ICT contact points in the International Cooperation Partner Countries (ICPC).

One of the main initiatives of the IDEALIST 2014 project is cooperation with international ICT associations in Europe. This backbone for international collaboration will guarantee quick information exchange. The IDEALIST project has already established closer cooperation with so-called "Coordination and. Support Actions" and other NCP networks with complementary missions. Closer cooperation between the other ICT funding initiatives has also begun. Last year, ERCIM signed a Memorandum of Understanding (MoU) with the IDEALIST project. The memorandum encompasses cooperation in partner search, sharing information about various events, invitation of ERCIM representatives to major EC workshops (eg the European Commission's ICT event in Vilnius which will be held in November 2013). MoUs were also agreed upon with the Central and Eastern European Network (CEENet) and Informatics Europe. It is expected that the new IDEALIST project will play an important role in encouraging the involvement of European scientific organizations, associations, and several existing projects in EC framework programs.

Link: http://www.ideal-ist.eu

Please contact:
Givi Kochoradze, EC ICT NCP in Georgia
International Center for Advancement of Research Technology and Innovation, Georgia
Tel: +995 599 292516, E-mail: gcp@ip.osgf.ge

ERCIM Security and Trust Management Workshop

by Rafael Accorsi and Silvio Ranise

ERCIM's Working Group "Security and Trust Management" (STM) is responsible for organizing the annual "International ERCIM STM Workshop", which provides an exciting setting to catch up with the state-of-the-art in this area. The ninth edition of this workshop took place in Egham, UK, in September 2013, in conjunction with ESORICS 2013, the "European Symposium on Research in Computer Security".

The pervasive nature of emerging information and communication technologies (ICT) brings with it new problems, including:

- New threats and the potential for existing security vulnerabilities to be exploited.
- New and unanticipated application scenarios may be accompanied by novel security threats.
- The increased virtual and physical mobility of users increases their interactions whilst rendering obsolete the notion of security perimeters.
- Privacy is also a major concern in the current ambient intelligence paradigm: devices can be interacting with users anywhere and at any time, allowing sensitive information to be gathered.

These issues create a new demand for reliable trust relationships among users, service providers, and even devices - owing to the increasing popularity of bring-your-own-device policies. End-users, technology producers, scientific and governance communities perceive these deeply intertwined problems at different levels of concern and urgency.

The ERCIM "Security and Trust Management" (STM) Working Group focuses on a series of activities that aim to foster European research and development in the areas of security, trust, and privacy in ICT. One of these activities is the annual ERCIM STM Workshop, which is affiliated with the European Symposium on Research in Computer Security (ESORICS). The workshop's main goal is to promote progress and novel research on all theoretical and practical aspects of security and trust in ICT by acting as a forum for researchers and practitioners from academia, industry, and government.

As in previous years, the review process for this ninth workshop was very competitive. We received 46 submissions, of which 15 were selected. These, together with two keynotes delivered by Claire Vishik (Intel Corporation, U.K.) and Michael Huth (Imperial College, U.K.), provided a very rich and exciting program that gave rise to lively discussions during both the workshop sessions and the breaks.

The invited speakers approached the problems around security and trust management from several perspectives (industry, research, and practice) to give a comprehensive and novel view of the challenges ahead for the scientific and technical communities in the field. Claire Vishik reported on the Intel's experiences on building trusted systems. Her invited speech summarized the lessons taken from the first generation of Trusted Computing, as well as the upcoming challenges for research and practice. Michael Huth presented a principled and well-founded approach to trust evidence in distributed systems. He reported on the results of an industry-funded project that provides for the verifiable numerical aggregation of trust evidence for policy-based access control. Together, these keynotes presented complementary views on the challenges of providing trust mechanisms in practical systems, as well as evidence that the provision of such mechanisms can make modern computing applications more trustworthy.

The 15 selected papers identify problems and propose solutions concerning a wide range of security and trust issues in current and future ICT applications related to privacy, cryptographic protocols, authorization and trust management policies, business processes, watermarking, networks, web and mobile applications. The proposed solutions are both theoretical and practical. The former are justified by applying rigorous mathematical techniques or formal method analysis tools, eg, to verify cryptographic protocols or to support the selective disclosure of sensitive data in trust management. The latter propose interesting extensions to existing techniques, eg, browser plug-ins to protect mobile devices from SSL vulnerabilities, or advocate greater attention to socio-technical security factors for evaluating user behaviour in choosing WiFi networks. We believe the contributed papers together with the two invited talks offer an interesting and inspiring picture of the state-of-the-art and indications on future directions in the field of security and trust. We encourage the reader to take a look at the workshop proceedings for a more in-depth overview of these and related problems (see link below).

In summary, we believe that the confluence and combination of rigorous techniques to alleviate certain technical and human-related security problems provide a great perspective for the development of hybrid techniques capable of greatly increasing the trust of both developers and users in the security of deployed ICT technologies. For all these reasons, we are very much looking forward to attending the tenth edition of the ERCIM STM workshop (affiliated with ESORICS) which will take place next year in Wroclaw, Poland.

Links:
ERCIM STM WG: http://www.iit.cnr.it/STM-WG/
ERCIM STM Workshop:
https://sites.google.com/site/sectrustmgmt2013/Home
Workshop proceedings: http://link.springer.com/book/10.1007/978-3-642-41098-7/page/1

Please contact:
Pierangela Samarati, STM WG coordinator
Università degli Studi di Milano, Italy
E-mail: pierangela.samarati@unimi.it

Rafael Accorsi, Business Process Security Group,
University of Freiburg, Germany
E-mail: rafael.accorsi@iig.uni-freiburg.de

Silvio Ranise, Centre for Information Technology
Fondazione Bruno Kessler, Trento, Italy
E-mail: ranise@fbk.eu

Julien Mairal Receives the 2013 Cor Baayen Award

Julien Mairal from Inria has received the 2013 ERCIM Cor Baayen Award. Honorary Mentions are given to Alexander Bertrand, KU Leuven and Ben Glocker, Microsoft Research Cambridge. The Cor Baayen Award is awarded annually by ERCIM to a promising young researcher in the field of Informatics and Applied Mathematics.

Julien Mairal obtained his PhD in 2010 at Inria under the joint supervision of Francis Bach and Jean Ponce. His PhD thesis was concerned with new and efficient algorithms for sparse coding and dictionary learning with applications to image processing and computer vision.

Julien's most significant contribution is a stochastic optimization algorithm for a problem know as dictionary learning in signal processing and machine learning.Given a dataset represented by high-dimensional vectors, the task consists of approximating each data point by a linear combination of a few elements from a learned dictionary. The original dictionary learning formulation, which was introduced a decade ago by Olshausen and Field from the neuroscience community, has recently gained a large interest in signal processing and machine learning, showing promising results for numerous applications. The proposed algorithm was shown to be much more scalable and computationally efficient than existing approaches, which is of acute importance currently with the emergence of massive datasets and the need to analyze them. This work is available through an open-source software, which is now widely used.

Julien also developed state-of-the-art image restoration algorithms for image denoising and image demosaicking (reconstructing colour images from raw data from CCD sensors), by combining two successful paradigms: sparse representations of image patches, and non-local image self-similarities. His third contribution was to develop new formulations of dictionary learning for classification or regression tasks, whereas traditional applications were devoted to signal or image reconstruction. This proved particularly useful in computer vision, where it allows discriminative representations of natural image patches to be found.

Finally, Julien and his collaborators worked on solving optimization problems related to structured sparse estimation, which has received recently a lot of attention in statistics and machine learning. Structured sparse models differ from classical ones in that variables are grouped together according to an a priori structure, a hierarchy for example. Computing penalized likelihood estimators is challenging in this context, but Julien and his collegues obtained efficient and scalable optimization techniques by drawing novel connections between network flows and sparse estimation. This allows the use of structured sparse models on a large scale, which was not previously possible.

Julien's PhD work was published in top conferences (CVPR, ECCV, ICCV, ICML, NIPS) and journals (JMLR, PAMI) of computer vision and machine learning, and was honoured by three French prizes for a best PhD thesis. It is also widely cited, with ten papers having received over 100 citations according to Google Scholar, and an amazing h-index of 19 for such a young researcher.

2013 Cor Baayen Award:
Winner: Julien Mairal
Honorary Mentions: Alexander Bertrand, Ben Glocker
Finalists: Igor Boehm, André Chailloux, Marek Cygan, Jose Luis Fernandez-Marquez, Gerard Hoekstra, Mustafa Misir Anna Monreale, Stefanie Nowak, Mathias Payer, Sini Ruohomaa, Floor Sietsma, Pavel Vácha, Rossano Venturini.

Cor Baayen Award

The Cor Baayen Award is awarded each year to a promising young researcher in computer science and applied mathematics. The award was created in 1995 to honour the first ERCIM President.

The award consists of a cheque for 5000 Euro together with an award certificate. The selected fellow is invited to the ERCIM meetings in autumn. A short article on the winner, together with the list of all candidates nominated, is published in ERCIM News.

Eligibility

Nominees must have carried out their work in one of the "ERCIM countries", currently Austria, Belgium, Cyprus, Czech Republic, Finland, France, Germany, Greece, Hungary, Italy, Luxembourg, Norway, Poland, Portugal, Spain, Sweden, Switzerland, the Netherlands and the United Kingdom

The selection of the Cor Baayen award winner is the responsibility of the ERCIM Human Capital Task Group, who may consult expert opinion in reaching their decision.

Cor Baayen Award 2014

The Call for nominations is expected to be published in February 2014, the deadline for submitting a nomination is probably 30 April 2014.

More information: http://www.ercim.eu/activity/cor-baayen-award

Image Understanding – An EU Perspective

Libor Král,
Head of Unit A2 – Robotics in Directorate General Communication Networks, Content & Technology, European Commission,
http://cordis.europa.eu/fp7/ict/robotics/
E-mail: libor.kral@ec.europa.eu

Vision is a key factor in natural evolution, whether aiding a jellyfish to detect light or providing the supra-human ability of an eagle to spot its prey from miles away. The desire to enhance human vision through artificial means dates back several centuries, but it is not until the last twenty years or so that machine and computer vision took off. Today, as evidenced by the contributions to this ERCIM edition, it is a vibrant research field.

Why should the European Union be interested in image technology? There are two main reasons. Firstly, vision systems represent an important socio-economic activity. Sales by European suppliers of sophisticated (e.g. application-specific or reconfigurable) vision systems and of products like smart cameras, software and components reached the €2 billion mark according to a 2012 survey by the European Machine Vision Association. Vision systems are an enabling technology with a multiplier effect on sectors important to the seventh EU research framework programme (FP7), like health, agriculture, space and security. Robotics, with a projected annual economic impact of several trillion euros worldwide by 2025 [1], is a prime example of a future sector whose growth will depend heavily on image understanding. According to the 2020 EUROP (European technology platform for robotics) Strategic Research Agenda (SRA), sensing is what sets robots apart from other types of machines. The SRA technology targets include increasing the resolution and range of 3D sensors and exploiting novel sensing mechanisms.

Secondly, the EU has focussed on the scientific potential of computer vision for just over one decade, marked by the launch of the ECVISION coordination network in March 2002 and of a small batch of research projects in cognitive vision. The term "cognitive vision" encapsulates an attempt to achieve more robust, resilient, and adaptable computer vision systems by endowing them with a cognitive faculty: the ability to learn, adapt, weigh alternative solutions, and develop new strategies for analysis and interpretation. This initiative broadened to cognitive systems and robotics in 2004.

Since then, FP7 has invested some €700 million into over 100 cognitive systems and robotics projects. The main scientific and technical research challenges for these projects lie in innovative methods of perception, understanding and acting in real world situations. Autonomous moving robots set radical new challenges for computer vision, e.g. for on-the-fly scene analysis and for person and object recognition in real-life environments, where only few prior assumptions can be made by the vision tools. Projects address topics such as urban scene recognition, modelling a city pedestrian's viewshed (line of sight), robust sensor deployment in the ocean for seabed mapping or in hazardous terrains for search & rescue, low-cost sensors for navigation or for shape and gesture recognition, networked vision, e.g. continuous tracking of a moving object or person from one set of cameras to another, and compact or modular and fast image processing for real-time inspection of large installations by mobile platforms like minicopters or crawling robots.

What is next at the EU level? Proposed research in the coming Horizon 2020 EU research framework programme will aim to increase significantly the scale of deployment and technological readiness of industrial and service robotics through research projects and through new innovation-promoting use cases. This will put extra demands of robustness and cost-effectiveness on vision systems. Research opportunities will not only exist in robotics. Proposed research into digital content will promote unprecedented access to very large data sets ("big data"), putting new demands on image analysis, modelling and visualization tools. Simulation and testing systems will become far more realistic and more closely integrated with the hardware they simulate. 3D and augmented reality technologies will need to evolve even further for digital content creation, virtual learning or digital gaming.

In conclusion, the role of image understanding as an enabling technology has been highly productive in FP7 and should certainly continue in Horizon 2020, not only pursuing scientific and technical goals in a wide variety of research topics but also contributing to innovation through deployment projects.

Reference:

[1] "Disruptive technologies: Advances that will transform life, business, and the global economy" (McKinsey, 2013), available at:
http://www.mckinsey.com/insights/business_technology/disruptive_technologies

Introduction to the Special Theme

Image Understanding

by Michal Haindl and Josef Kittler

Vision is the most important sense on which the majority of organisms depend for life. Scene reflectance properties in various spectral bands provide invaluable information about an object's characteristics, including its shape, material, temperature, illumination and dynamism. This information, however, is very difficult to capture with an electronic device. A real visual scene to be captured is subject to variable illumination as well as variable observation conditions. Furthermore, single objects of interest can be partially occluded or shaded, may be positioned at various distances from the capturing device, data can be noisy and / or incomplete; thus successful interpretation of imaging sensor data requires sophisticated and complex analytical methods and computing power.

The wide availability of visual data and continuous advances in computer vision and pattern recognition techniques have stimulated a growing interest in novel applications, which successfully simulate human visual perception. This special issue presents a sample from the gamut of current research activities in this area, reflecting the wide spectrum of imaging and visualization (range observation) modalities and their combinations. These include the conventional static grayscale (Salerno et al) and colour images (Cardillo et al, Riesen et al, Li et al, Amato et al, Suk et al), satellite panchromatic and radar images (Craciun et al, van Lieshout), stereo images (Kadiofsky et al), magnetic resonance images (Roerdink, Murino et al), through dynamic video (Bak et al, Gaidon et al), and multi modal measurements such as spectral video and audio (van der Kreeft et al), range-thermal images (Guerrero), video-multi-beam LIDAR (Benedek et al), spectral-range video (Oikonomidis et al, Fotopoulos et al, Piérard et al, Rogez et al), and the frustrated total internal reflection images (Risse et al).

This issue unveils an impressive number of successful applications of visual scene understanding. Remote sensing applications are targeted in papers by Craciun et al (boat detection) and van Lieshout (field detection), document analysis is addressed in Riesen et al (handwritten word recognition in ancient manuscripts) and in Salerno et al (interferences removal). Biological studies feature in Risse et al (small translucent organisms detection and real time visualization of their internal structures) and Suk et al (leaf recognition based on visual contour features). Driver assistance systems are the topic of the contribution by Kadiofsky et al, while video shot detection is discussed in Kreeft et al and Li et al. Mixed reality is pursued by Benedek et al, and security by Bak et al (person identification) and Gaidon et al (human activities recognition). The fashionable topic of content-based image retrieval is discussed in Cardillo et al and Amato et al. Hand tracking is addressed in Oikonomidis et al. The most popular applications in the issue are medical applications. This group includes the papers of Guerrero (mental or physical disabilities detection), Roerdink (brain diagnosis), Murino et al (neurodegenerative diseases), Fotopoulos et al and Piérard et al (multiple sclerosis), and Rogez et al (assisted living).

The articles herein reflect recent trends in tackling diverse problems in visual information analysis in realistic, less restrictive conditions. They indicate that moving from fixed laboratory acquisition setups to much more challenging, dynamically changing exterior conditions, where all critical parameters, such as illumination, distance, viewpoint, shape and surface properties can vary simultaneously, requires either sophisticated invariant representation or adaptive machine learning approaches. The work here also suggests that adopting contextual cognitive reasoning and multidimensional data models offers an effective way to deal with the challenges of ever-increasing scene complexity.

Please contact:
Michal Haindl
Institute of Information Theory and Automation AS CR, Czech Republic
E-mail: haindl@utia.cas.cz

Josef Kittler, University of Surrey, UK
E-mail: j.kittler@surrey.ac.uk

Boat Extraction in Harbours From High Resolution Satellite Images Using Marked Point Processes

by Paula Crăciun and Josiane Zerubia

Earth observation satellites represent a significant resource when it comes to acquiring data about the Earth. Satellite data is used in a range of fields, including environmental monitoring, map updating and meteorology. Since the launch of the first Earth observation satellite, the resolution of the optical sensors installed on board has greatly improved, thus, nowadays, panchromatic images can be acquired at a resolution equal or lower than 0.7 [m] (ie GeoEye, Pleiades). This makes it possible to recognize small objects, such as boats and cars.

When it comes to the problem of object extraction, two conceptually different techniques are usually used. The first, a pixel-based approach, takes only the spectral information of individual pixels into account. It has proved its usefulness in the case of low or medium resolution imagery, but has found its limits in high resolution imagery owing to the rich information content in such data. The main difficulties appear when trying to cope with objects that are unevenly illuminated, partially occluded or blended with the background. The second type of technique considers objects, rather than individual pixels during the extraction. It also takes shape and textural information into account, along with the spectral data.

A possible framework for object-based solutions is offered by stochastic geometry. The quest is to find a pattern of objects that fits the image, which is considered to be a realization of a random process. Three main ingredients are needed to build a spatial pattern analysis tool: a prior to give a general aspect of the searched solution, a conditional probability that depends on the data, and an optimization method to get the best solution to the problem which consists of maximizing a probability or minimizing an energy function [1].

In this regard, statisticians working in image processing have introduced marked point processes as a tool to solve pattern analysis problems. By developing specialized models, different kinds of objects in complex scenes can be handled. The marked point processes are used as priors. Interactions between objects are defined, allowing for the incorporation of global properties into the model. The shape of the objects is modelled by constraints on the marks of the point process.

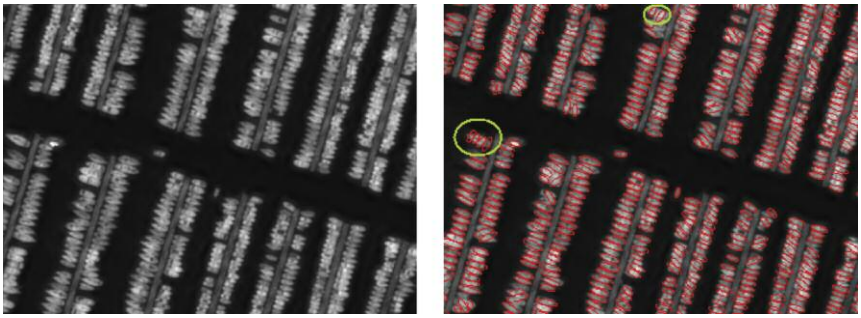


Figure 1. (left) Image of boats in harbors ©CNES; (right) Extraction results ©INRIA-AYIN.

The conditional probability depending on the data is given by the likelihood that a configuration of objects fits the actual image. Finally, by maximizing a Bayesian criterion, such as the Maximum A Posteriori using an optimization method, one obtains the resulting configuration, eg the configuration which maximizes the product between the prior and the likelihood.

One particularly complex case is the extraction of boats in harbours. The objects of interest are usually clustered around the docks, vary in size and orientation, and tend to be very close to each other. Here, objects are modelled as stochastic ellipses. Constraints on the prior do not allow for object overlapping and neighbouring objects should have similar orientations. The conditional probability depending on the data is determined individually for each object in a given configuration by computing a contrast distance between the interior of the ellipse and part of its surrounding background.

Additional prior information is embedded into the model by identifying the local orientations of the objects [2]. By determining how each dock in the harbour is oriented, it is possible to make inferences about the orientation of the boats. Figure 1 shows results for boat extraction in a harbour.

The areas highlighted in green present extraction errors due to the fact that the objects in the image are larger than the sizes specified in the model.

Marked point processes offer a reliable tool and moreover, a general framework for extracting objects from high resolution optical satellite images. Their usability in dynamic images (ie for object tracking) is currently being investigated.

Link: https://team.inria.fr/ayin/

References:

[1] S. Ben Hadj, F. Chatelain, X. Descombes and J. Zerubia, "Parameter estimation for a marked point process within a framework of multidimensional shape extraction from remotely sensed images", Proc. ISPRS, XXXVIII, 2010

[2] P. Craciun and J. Zerubia, "Unsupervised marked point process model for boat extraction in harbors from high resolution optical remotely sensed images", Proc. IEEE ICIP, 2013

Please contact:

Josiane Zerubia,
AYIN research team, Inria, France
E-mail: Josiane.Zerubia@inria.fr

FIM: Frustrated Total Internal Reflection Based Imaging for Biomedical Applications

by Benjamin Risse, Xiaoyi Jiang, and Christian Klämbt

Video-based imaging of animal behaviour is commonly used in biomedical studies. Imaging small and translucent organisms, such as worms or larvae, however, tends to require sophisticated illumination strategies. We developed a novel technique to image the contact surface between organisms and substrate utilizing Frustrated Total Internal Reflection. This technique has a wide range of potential applications.

Imaging the locomotion of animals in behavioural experiments has become commonplace. Over the past several decades, video-based quantification has become increasingly important in understanding the mechanisms involved in circadian rhythms, learning, courtship and various diseases. Even genetic inferences are possible based on the observable characteristics [1].

However, owing to their small size and translucent appearance, many popular model organisms, such as Drosophila melanogaster and Caenorhabditis elegans, require sophisticated illumination strategies to acquire contrast-rich images. Furthermore, moisture or irregular subsurface media are frequently required to ensure natural movement or growth. Several image improvement techniques and new imaging modalities have been introduced recently, covering dye assays, illumination protocols and optical as well as non-optical imaging techniques.

We have devised a novel acquisition technique to indirectly determine the shape and position of organisms. The technique, which is based on Frustrated Total Internal Reflection (FTIR), measures the contact surface between the organism and subsurface (FTIR-based Imaging Method: FIM) [2]. The FTIR-principle is well known in the human-computer interaction community, since it is used in the construction of touch devices [3], but it also has biomedical applications. We use Drosophila larvae as a running example (Figure 1a), since conventional imaging of translucent larvae suffers from a number of difficulties. However, FIM can be used for a wide range of applications (examples of FIM-imaged planarian, C. elegans, fruit flies, and Arabidopsis roots are shown in Figure 1b).

Instead of illuminating the scene with incidental or transmitted light, the light source is integrated into a subsurface, such as an acrylic glass plate (Figure 2a). Because of the optical density of glass and air, providing the angle of incidence is greater than a critical angle, the light is completely reflected at the glass/air boundary (ie total internal reflection). The critical angle depends on the refractive indices of both materials. To improve the environmental conditions for larvae, a thin but moist agar layer is placed between the animals and the surface. Since the optical density decreases (ie the refractive indices increase) from glass over agar to the larva, the light can pass through the

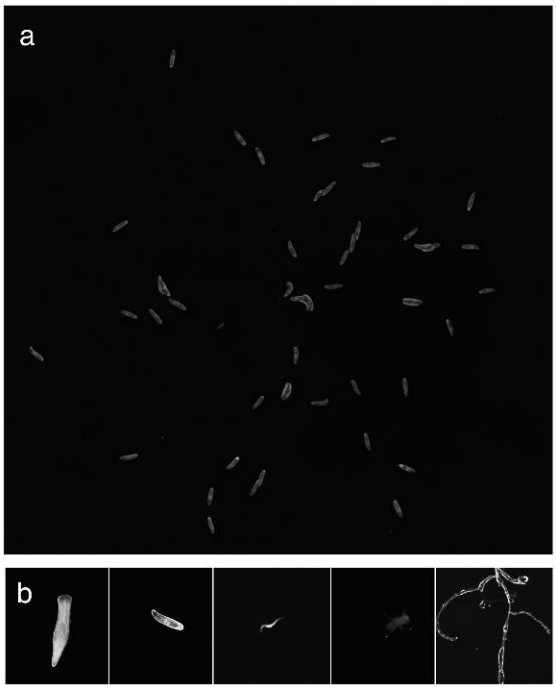


Figure 1: a): Resultant image of dozens of larvae: Raw image (neither image enhancement nor background subtraction) showing larvae of different sizes. b) FIM offers a wide range of applications: Planarian, Drosophila larva, C. elegans, adult Drosophila fly, Arabidopsis roots (from left to right)

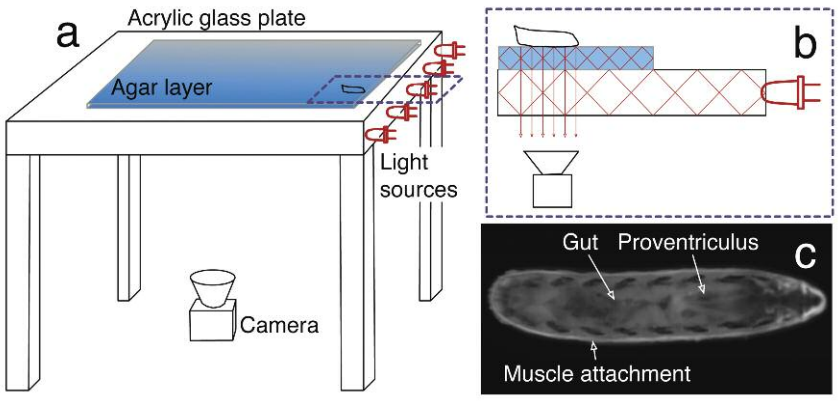


Figure 2: a) System overview: Light sources are integrated into the tracking stage, which is covered by an agar layer. The camera captures the reflected light from underneath. b) FTIR principle: Light (red lines) passes from the glass into the agar and is frustrated at the agar/larva contact area, where it can pass through the layers and is captured by the camera. c) Resultant image of a single larva: Several internal organs are highlighted.

layers while the angle of incidence decreases during the transition between the materials. If the angle of incidence is below the critical angle, the light is said to be frustrated since the reflection is no longer total at that point and can be measured by a camera from underneath (Figure 2b). As a result, only light reflections caused by the objects of interest are visible, while the surface appears as an almost black background, leading to a superior foreground-background contrast (Figure 1a). Furthermore, most disturbing artefacts, such as surface irregularities and light reflections caused by the surface area can be eliminated on a physical level, thus reducing the requirements for subsequent image analysis algorithms to facilitate automatic quantifications.

Because of the high contrast sensitivity of FIM, even internal structures of the animals become visible using relatively low-resolution cameras. For example, organs including the muscle attachments, the gut and the proventriculus can be identified in Drosophila larvae (Figure 2c). These new features can be used in the subsequent automatic image analysis to generate more precise quantifications. For instance, these structures were used to automatically identify the orientation (ie the head) of the animals [2]. Note that the spatial and temporal resolution of FIM is only limited by the resolution of the camera: FTIR is based on the principles of light, thus even small structures, such as the footprints of adult fruit flies can be measured (Figure 1b) [2].

In summary, FIM offers several advantages over conventional imaging: superb image quality, almost noise and background free recordings and the ability to visualize internal structures. Because of this and its straightforward and cheap construction, and its inherent high spatial and temporal resolution, FIM is feasible for a wide range of biomedical applications.

References:
[1] M. B. Sokolowski: "Drosophila: Genetics meets behaviour," Nat. Rev. Genet., vol. 2, no. 11, pp. 879–890, Nov. 2001
[2] B. Risse, et al: "FIM, a Novel FTIR-Based Imaging Method for High Throughput Locomotion Analysis," PloS one, vol. 8, no. 1, p. e53963, Jan. 2013
[3] J. Han: "Low-Cost Multi-Touch Sensing through Frustrated Total Internal Reflection," Proceedings of 18th Annual ACM Symposium on User Interface Software and Technology, 2005.

Please contact:
Benjamin Risse and Xiaoyi Jiang
Department of Mathematics and Computer Science
University of Münster
Tel: +49 251 83 32704
E-mail: {b.risse,xjiang}@wwu.de

AXES - Finding Video Clips Using Speech and Image Recognition

by Peggy van der Kreeft, Kay Macquarrie and Martijn Kleppe

Searching for clips or segments of videos in large archives can be a daunting task. In which clip was a person mentioned and where in the clip is he or she shown? Even after you locate the correct video, you still need to watch the entire video to find that one segment containing the person that you are looking for. The novel technologies being developed by AXES make finding what you are looking for in large archives and libraries significantly easier.

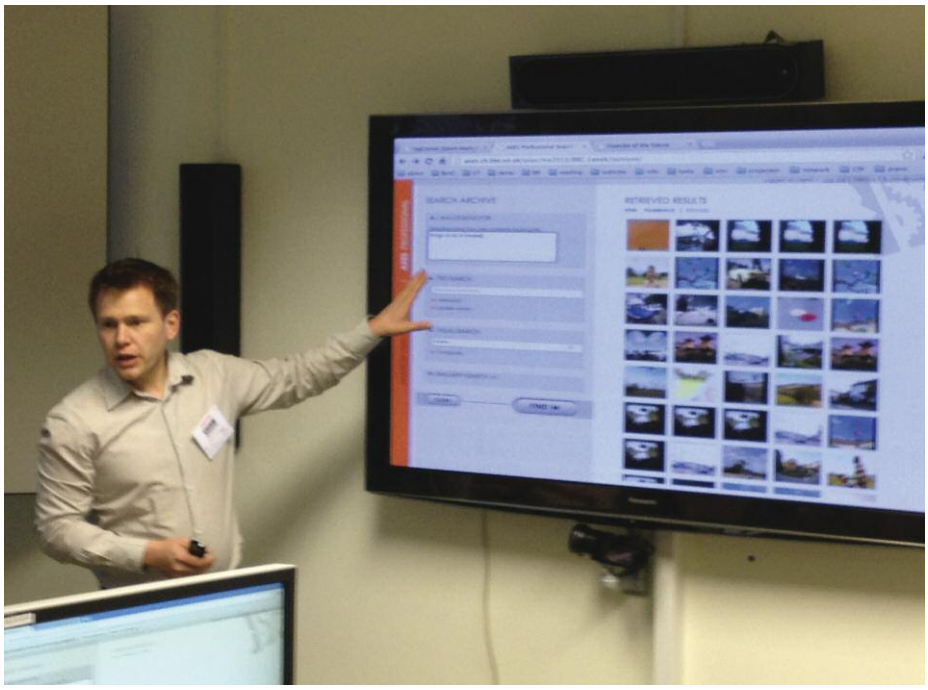
The aim of the AXES (Access to Audio-Visual Archives) project is to develop tools that provide various types of users with new engaging ways to interact with audiovisual libraries, helping them discover, browse, navigate, search and enrich archives.

The technologies used in the project involve multimodal analysis of people, places, objects and events. This includes recognition and identification at both a general and a rather specific level. For example, the system can find categories of people (athlete) or individuals (Angela Merkel); location categories (countryside) or specific places (Grand Place, Brussels); object categories (car) or specific objects (a logo or flag); and event categories (mountain climbing) or particular events (a speech by a politician). Visual analysis (including image similarity), audio (speech-to-text) and text analysis (metadata, OCR) across various languages, and advanced linking technologies are also seamlessly integrated.

Each component of the AXES system focuses on a particular aspect of the audiovisual analysis. For instance, the on-the-fly specific person retrieval [1] component provides a method for finding shots of a particular person in a large dataset without relying on metadata. This on-the-fly method involves pre-processing of the video corpus, making it first searchable for any person, and subsequently, based on a text query specifying the individual, learning a discriminative classifier for that person from face images downloaded from Google image search. It results in a ranking of the faces in the corpus, allowing retrieval of the person of interest.

The AXES video search engine is designed to allow easy integration of existing and novel search methods, to be adaptable for innovative user interaction models, and to provide a test bed for trialing these methods. The search engines and components have been demonstrated and extensively tested within the TRECVid international benchmarking initiative during the course of the project. [2]

The user accesses the system through one of the user interfaces, specifically geared towards different end-user groups. Three distinct user interfaces



Robin Aly of University of Twente explains visual search during user trials.

are being developed, one for each of the targeted user groups and based upon several user studies [3]. Specifically, we target (1) media professionals and archivists, working with archive systems and metadata on a daily basis; (2) researchers, including students, academics and (investigative) journalists looking for resource material; and (3) home users, the general public in search of information. The integrated system, built upon the open-source WebLab platform, is accessible as web services and all components work in the background without expert knowledge required from the users; the user interface steers the end user through the advanced and powerful system in a user-friendly and intuitive way.

The first prototype, AXES PRO, was developed during the first year and has been tested by media professionals. The second prototype, AXES RESEARCH, is currently in its final stages of development, with preliminary testing ongoing. Full user testing by the researcher group is scheduled for autumn 2013. The third prototype, AXES HOME, is planned to be deployed in 2014, with subsequent user testing. The underlying technologies being used by all three systems are largely the same, but the user options and interface differ based on carefully considered user requirements, resulting in an optimized user experience geared to each user group.

AXES opens up a whole new way of experiencing digital libraries and archives, reaching out to the end user and making the vast and rich amount of existing audiovisual content accessible. Visit the AXES website (http://www.axes-project.eu) for more details, video demonstrations, and related publications.

Link:
http://www.axes-project.eu

References:
[1] O. M. Parkhi et al.: "On-the-fly Specific Person Retrieval," International Workshop on Image Analysis for Multimedia Interactive Services, 2012
[2] R. Aly et al.: "AXES at TRECVID 2012: KIS, INS, and MED", in: TRECVid 2012, 26-28 Nov 2012, Gaithersburg, Maryland, USA
[3] M. Kemman et al.: "Dutch Journalism in the Digital Age," Icono 14, 11(2) 2013, pp. 163–181. doi:10.7195/ri14.v11i2.596.

Please contact:
Peggy van der Kreeft
Deutsche Welle, Germany
Tel: +32 495 664099
E-mail: peggy.van-der-kreeft@dw.de

Kay Macquarrie, Deutsche Welle, Germany
Tel: +49 30 4646 5656
E-mail: kay.macquarrie@dw.de

Martijn Kleppe, Erasmus Universiteit Rotterdam, the Netherlands
Tel: +31 10 408 2646
E-mail: kleppe@eshcc.eur.nl

AXES

The four-year research project AXES started in January 2011 and is co-funded within the EU FP7 framework. Thirteen partners are collaborating on this integrated project. ERCIM, the administrative coordinator, and KU Leuven (Belgium), the technical coordinator, ensure a consistent and collaborative operation. Cassidian (France), Technicolor (France), Inria (France), Fraunhofer IAIS (Germany), KU Leuven, Dublin City University (Ireland), University of Oxford (UK), University Twente (the Netherlands) develop and integrate the technical components. Finally, user partners BBC (UK), NISV (the Netherlands), Erasmus University Rotterdam (the Netherlands) and Deutsche Welle (Germany) describe user requirements, provide audiovisual content and perform user testing. The partners' complementary skill sets allow an in-depth and wide coverage of technologies resulting in a powerful, innovative set of audiovisual search and analysis tools.

Random Mosaics for Network Extraction

by Marie-Colette van Lieshout

Arak and Surgailis [1] introduced a class of random mosaics with remarkable mathematical properties. A collaborative project between researchers in Poland and at CWI shows that such models are useful for intermediate level image analysis because they can capture global aspects of an image without requiring a detailed description of the objects within it.

In a research project carried out in the context of the ERCIM MUSCLE Working Group, researchers at CWI and the Nicolaus Copernicus University in Toruń, Poland, studied random mosaic models[1] for intermediate level image analysis. These elegant mathematical models, which possess appealing properties such as consistency, have not been used in practice for a long time. The main deterrent to their use has been a naive Monte Carlo algorithm which iteratively makes small local changes to model realizations, converges slowly, and is difficult to program. The overall aim of our research project was to demonstrate that Arak-Surgailis type models can be applied by developing efficient Monte Carlo samplers. Initially, most attention was paid to segmentation and foreground/background separation [2]. More recent research at CWI has been focusing on extraction of fields in rural landscapes [3].

Figure 1, obtained from the collection of publicly released Synthetic Aperture Radar images at the NASA/JPL web site http://southport.jpl.nasa.gov, shows an agricultural region in Ukraine. A pattern of fields separated by tracks is visible, broken by some hamlets. Note that the tracks that run between adjacent fields show up in the image as whitish lines against the darker fields. In other words, a track is associated with a high image gradient. The objective is to extract the mosaic of fields.

To do so, we build a family of straight lines on which to draw mosaics by computing the gradient of the data image (after convolution with a radially symmetric Gaussian kernel to suppress noise) and select the most salient edges. To quantify how well a particular mosaic fits the data, we use the integrated absolute gradient flux along its edges (thresholded to discourage spurious ones). Note that as a sum of contributions from each edge, this function is local in nature, which is convenient from a computational perspective.

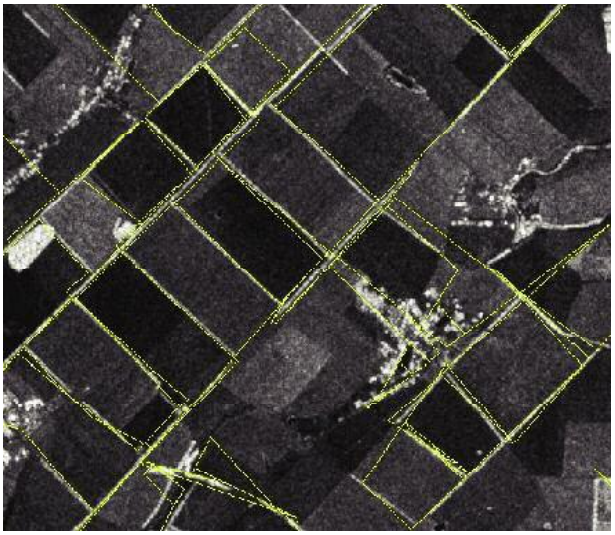


Figure 1: Mosaic (in yellow) overlaid on a Synthetic Aperture Radar image of fields in rural Ukraine.

To extract the fields, we use simulated annealing with a geometric cooling schedule. The result is shown in Figure 1 by the yellow lines. Note that all fields bordered by tracks are detected. A few false positives occur near the hamlets and are connected to the track network. The precision of the line placement is clearly linked to that of the underlying edge set. For example, the border of a field near the middle of the top of the image is misplaced. Naturally, one could expand the collection of edges to increase the precision, but this would incur a higher computational cost.

From a mathematical point of view, we introduce consistent random mosaics with polygonal tiles that are not necessarily convex, thus allowing a wide variety of shapes. The construction is inspired by the Arak-Surgailis model and extends our previous work [3]. We give an equivalent construction in terms of an interacting particle system that can be used to prove basic model properties including local and spatial Markovianity. This dynamic representation is also the foundation on which to build efficient Monte Carlo samplers that combine local modifications which are useful for fine tuning with global changes that allow fast exploration of the space of mosaics.

This research was funded by the EC 6th Framework Programme, as well as the Polish and Dutch Research Councils.

Full details of the research can be found in [2].

Link:
http:/homepages.cwi.nl/~colette

References:
[1] T. Arak, D. Surgailis: "Markov fields with polygonal realizations", Probab. Theory Related Fields 80

[2] T. Schreiber, M.N.M. van Lieshout: "Disagreement loop and path creation/annihilation algorithms for binary planar Markov fields with applications to image segmentation", Scand. J. Statist. 37

[3] M.N.M. van Lieshout: "Discrete multi-colour random mosaics with an application to network extraction", Scand. J. Statist., to appear.

Please contact:
Marie-Colette van Lieshout
CWI, The Netherlands
Tel: +31 20 5924008
E-mail: M.N.M.van.Lieshout@cwi.nl

Computer-Aided Leaf Recognition Visual System

by Tomáš Suk, Petr Novotný and Jan Flusser

Plant identification is an important task in botany and related areas, such as agriculture, forestry, and nature conservation. It is also of interest of general public. While botanists usually have no problem identifying a species, non-specialists would often welcome a computer-aided system for species recognition. Creating such a system is a challenge that we have resolved using visual pattern recognition methods.

Generally, trees and shrubs can be recognized by their local features extracted from leaves, flowers, fruits and bark, and by global characteristics, such as height, crown shape, and branch structure. Our system uses only the leaves, since leaves alone provide enough discriminative information, are available from spring to autumn, and, last but not least, the user can collect them and classify the trees retrospectively.

The system consists of two parts: the database, which is dependent on the geographic location, and the universal search engine. We created our own dataset named Middle European Woody Plants (MEW 2012, see Figure 1), which is available at http://zoi.utia.cas.cz/node/662. It comprises native and frequently cultivated trees and shrubs of the Central Europe region. It contains 151 botanical species, at least 50 samples per species and a total of 10,000 samples (leaves) scanned in high resolution.

Figure 1: Samples from our dataset (different scales): First row - Ailanthus altissima, Betula pubescens, Buxus sempervirens, Crataegus monogyna, Euonymus europaea; Second row - Humulus lupulus, Prunus padus, Prunus spinosa, Quercus robur, Quercus rubra.

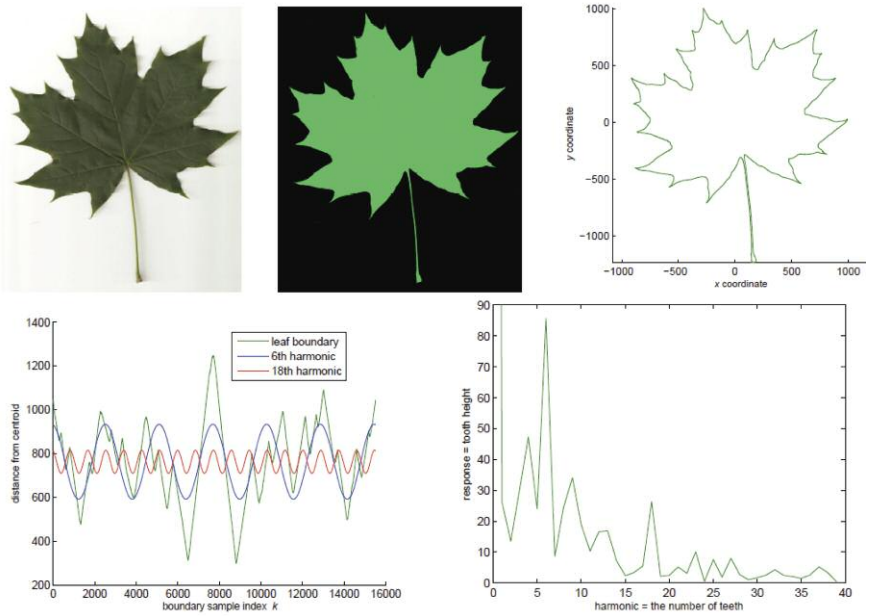


Figure 2: Processing of a leaf (Acer platanoides) - scanning, thresholding, boundary tracing, Fourier kernel functions and magnitude of descriptors.

Recognition is based solely on leaf contour, leaf size and classification of the leaf as either simple or compound. We avoided using leaf texture and colour because they can vary between individuals of a given species, and can vary depending on the season (phenology phase). Furthermore, describing these features requires working with very fine venation detail.

From a mathematical point of view, the crucial question was how to efficiently encode and characterize the leaf contours. The best results were obtained by Fourier descriptors (Figure 2). We achieved an 89% success rate in the experiment, in which the dataset was randomly divided into two halves. One half was used as a training set and the other half was tested against it. We also compared the performance of the automatic method with the performance of humans. We asked 12 computer science students to classify the leaves visually. The students were able to see the query leaf and simultaneously browse the database to compare the query with the training leaves. Unlike the algorithm, which works only with contours, the students worked with full colour images. Each test person classified 30 leaves. The mean success rate was 63%, significantly lower the success rate of the algorithm.

Our recognition system is publicly available as a web-based application, which allows the user to upload the query leaf and receive the answer with greater reliability and higher speed than

from a non-trained individual. We encourage readers to try the system at http://leaves.utia.cas.cz/index?lang=en

The system is a good tool for non-specialists; it is not aimed at professional botanists. The system can be easily adapted to other areas and continents just by replacing/extending the MEW database with a database of local plants, the search engine does not require any modification. Some publicly available databases which can be used for this purpose include:

- Flavia: 1,907 samples of 33 species from the Yangtze Delta, China
- The Swedish data set: 1,125 samples of 15 species from Sweden
- ICL (Intelligent Computing Laboratory): 16,851 samples of 220 species from China
- ImageCLEF (Cross Language Evaluation Forum): version 2011 includes 6,436 pictures of 71 species from the French Mediterranean area.

The main weakness in the current version of the system is that it requires high-quality leaf scanning. We plan to develop an advanced version which should also work with cell-phone photographs, and we are also considering an embedded application for smartphones.

Detailed information about the system can be found in [1] and [2].

References:
[1] P. Novotný, T. Suk: "Leaf recognition of woody species in Central Europe", Biosystems Engineering, 115(4):444–452, 2013
[2] T. Suk, J.Flusser, P. Novotný: "Comparison of leaf recognition by moments and Fourier descriptors", in Computer Analysis of Images and Patterns, vol. 8047 of LNCS, p. 221–228, Springer, 2013.

Please contact:
Tomáš Suk, Institute of Information Theory and Automation of the ASCR, Prague, Czech Republic
Tel: +420 266052231
E-mail: suk@utia.cas.cz

Automatic Recognition of Human Activities in Realistic Videos

by Adrien Gaidon, Zaid Harchaoui and Cordelia Schmid

Automatic video understanding is a growing need for many applications in order to manage and exploit the enormous – and ever-increasing – volume of available video data. In particular, recognition of human activities is important, since videos are often about people doing something. Modelling and recognizing actions is as yet an unsolved issue. We have developed original methods that yield significant performance improvements by leveraging both the content and the spatio-temporal structure of videos.

Video is a popular medium, spanning a wide range of sources and applications. One of the main challenges is how to easily find meaningful content, for instance, particular events in video-surveillance data and video archives; certain types of human behaviour for autonomous vehicles in the area of robotics; and gestures in human-computer interfaces.

Owing to the overwhelming quantity of videos, answering this question calls for tools that can automatically analyse, index, and organize video collections. The problem is, however, that computers are currently unable to understand what a video is about. The relationship between the bits composing a video file and its actual information content is complex.

Our research focuses on designing computer vision and machine learning algorithms to automatically recognize generic human actions, such as "sitting down", or "opening a door" (see Figure 1), which are often essential components of events of interest. In particular, we investigate the fundamental structure of actions and how to use this information to accurately represent real-world video content [1].

Figure 1: Two human actions and their temporal decomposition.

The main difficulty lies in how to build models of actions that are information-rich, in order to correctly distinguish between different categories of action (such as running versus walking), while at the same time being robust to the large variations in actors, motions, and imaging conditions present in real-world data.

We discovered that many types of action can actually be modelled as a simple sequence of action atoms, or "actoms", which correspond to meaningful atomic temporal parts [2]. We found a robust parameterization, our "Actom Sequence Model" (see Figure 2), which efficiently leverages an

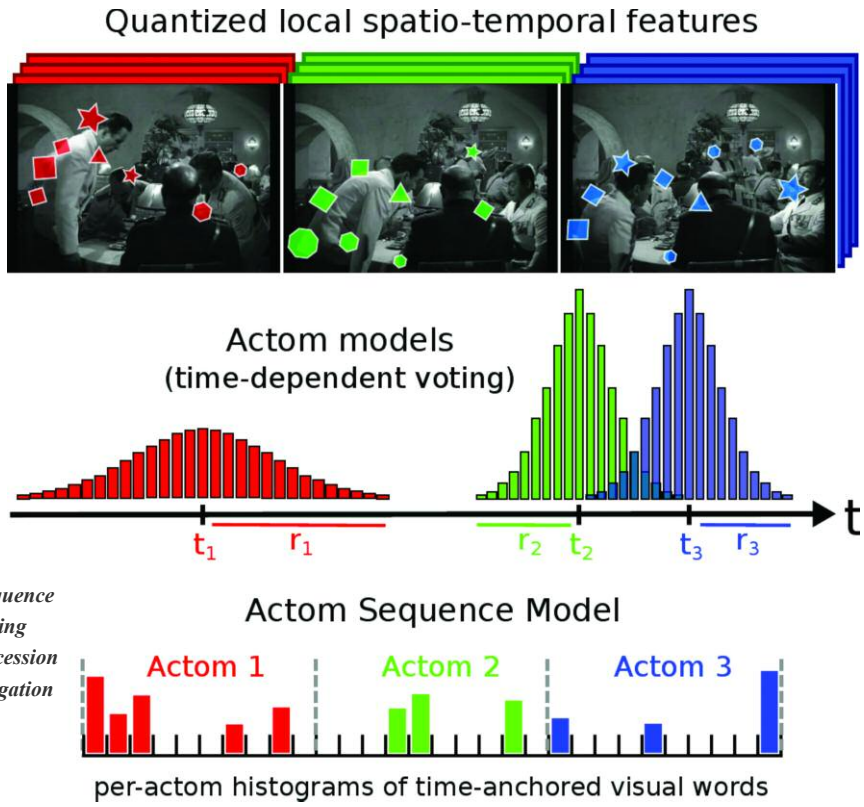


Figure 2: Illustration of our Actom Sequence Model (ASM) on three actoms of a sitting down action. ASM is based on the succession of temporal parts modeled by the aggregation of time-anchored local features.

action's temporal structure. Our method allows for more accurate action detection in long video sequences, which is particularly important when aiming to index events in large video archives such as the BBC motion gallery or the French INA repositories.

The previous approach is well suited to actions performed in a few seconds. Some activities, however, are not just a short sequence of steps, but have a more complex spatio-temporal structure, for instance sports activities like weight-lifting (see Figure 3). In these cases, we found that it is possible to extract elementary motions and how they relate to each other without human intervention [3]. This allowed us to represent complex activities using a tree-like data structure that can be efficiently computed and compared across videos. We showed that our approach can tag high-level activities in sports videos, movies, TV archives, and internet videos.

In all cases, we conducted thorough experiments on real-world videos from a wide array of sources: movies, TV archives, and amateur and internet videos (YouTube). We showed that our methods outperform the current state of the art on a variety of actions. These promising results suggest that our structured models could be applied in many different application contexts, for instance to recognize complex behaviours of autistic children for diagnostic assistance, to automatically browse video archives based on their contents instead of their meta-data, or to index web videos by generating video tags based on detected events.

This research was conducted during the PhD thesis of Adrien Gaidon, under the supervision of Zaid Harchaoui and Cordelia Schmid, at the joint Microsoft Research – Inria research center in Paris, France, and with the LEAR team, Inria Grenoble, France.

References:

[1] A. Gaidon: "Structured Models for Action Recognition in Real-world Videos", PhD thesis, http://tel.archives-ouvertes.fr/tel-00780679/

[2] A. Gaidon, Z. Harchaoui, C. Schmid: "Temporal Localization of Actions with Actoms", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013 http://hal.inria.fr/hal-00804627/en

[3] A. Gaidon, Z. Harchaoui, C. Schmid: "Recognizing activities with cluster-trees of tracklets", British Machine Vision Conference, 2012, Guildford, United Kingdom. http://hal.inria.fr/hal-00722955/en

Link:

https://lear.inrialpes.fr/people/gaidon/

Please contact:

Adrien Gaidon
Xerox Research Center Europe
adrien.gaidon@xrce.xerox.com

Figure 3: Automatically obtained motion decomposition of a weightlifting activity.

Egovision4Health - Assessing Activities of Daily Living from a Wearable RGB-D Camera for In-Home Health Care Applications

by Grégory Rogez, Deva Ramanan and J. M. M. Montiel

Camera miniaturization and mobile computing now make it feasible to capture and process videos from body-worn cameras such as the Google Glass headset. This egocentric perspective is particularly well-suited to recognizing objects being handled or observed by the wearer, as well as analysing the gestures and tracking the activities of the wearer. Egovision4Health is a joint research project between the University of Zaragoza, Spain and the University of California, Irvine, USA. The objective of this three-year project, currently in its first year, is to investigate new egocentric computer vision techniques to automatically provide health professionals with an assessment of their patients' ability to manipulate objects and perform daily activities.

Activities of daily living (ADL) represent the skills required by an individual in order to live independently. Health professionals routinely refer to the ability or inability to perform ADL as a measure of the functional status of a person, particularly in regards to the elderly and people with disabilities. Assessing ADL can help: 1) guide a diagnostic evaluation, 2) determine the assistance a patient may need on a day-to-day basis or 3) evaluate the rehabilitation process. Initial deployment of technologies based on wearable cameras, such as the Microsoft SenseCam (see Figure 1a) have already made an impact on daily life-logging and memory enhancement. We believe that egocentric vision systems will continue to make an impact in healthcare applications as they appear to be a perfect tool to monitor ADL. One unique wearable camera can potentially capture as much information about the subject's activities as would a network of surveillance cameras. Another important benefit is that the activities are always observed from a consistent camera viewing angle, ie in first-person view.

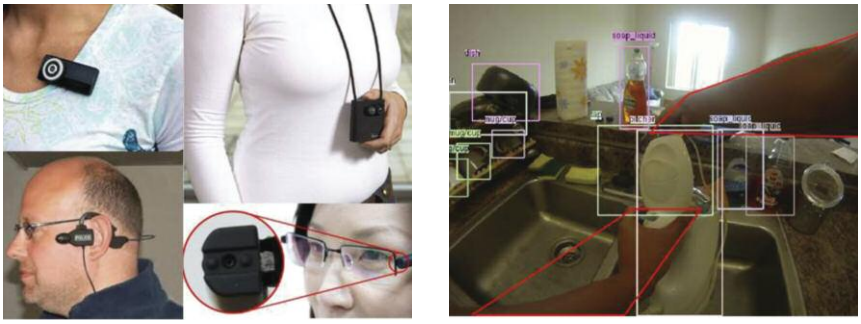


Figure 1 a (left) Examples of wearable cameras (Clockwise from top-left) lapel, neck-worn Microsoft SenseCam, glasses and head-worn camera. Figure 1b (right) Example of a processed image from [1].

Recent work on ADL detection from first-person camera views [1] (Figure 1b) demonstrated an overall performance of 40.6% accuracy was obtained in ADL recognition, and 77% when simulating a perfect object detector. In egocentric vision, objects do not appear in isolated, well positioned photos, but are embedded in a dynamic, everyday environment, interacting constantly with one another and with the wearer. This greatly complicates the task of detection and recognition, especially when an object is being manipulated or occluded by the user's arms and hands.

EgoVision4Health is addressing this problem. Our work is organized along three research objectives: 1) to advance existing knowledge on object detection in first-person views, 2) to achieve advanced scene understanding by building a long-term 3D map of the environment augmented with detected objects, and 3) to analyse object manipulation and evaluate ADL using detailed 3D models.

The analysis of "near-field" object manipulations - and consequently ADL recognition and assessment - could benefit greatly from having all the objects that are likely to be manipulated already located in the 3D environment. For example, if we want to determine whether a person is picking up a mug the wrong way due to an injury it seems important to know where the handle of the mug is, and how it is oriented in 3D. Another advantage of having all the objects already located around the subject is that we can categorize the scene and improve ADL recognition, eg cooking only happens in the kitchen. For a real breakthrough in ADL detection and assessment from a wearable camera, a thorough a priori understanding of the subject's environment is vital.

Since we expect Kinect-like depth sensors to be the next generation of cheap wearable cameras, we use a RGB-D camera as a new wearable device and exploit the 2.5D data to work in the 3D real-world environment. By combining bottom-up SLAM techniques and top-down recognition approaches, we cast the problem as one of "semantic structure from motion" [2] and aim at building a 3D semantic map of the dynamic environment in which the wearer is moving. We plan to model objects and body parts using 3D models and adopt the completely new approach of considering each ADL as the interaction of the 3D hands with 3D objects.

The hypothesis at the basis of our proposal is that the EU's well-established technology in mapping for robotics and the latest computer vision techniques can be cross-fertilized for boosting egocentric vision, particularly ADL recognition. Our goals are motivated by the recent advances in object detection, human-object interactions and ADL

detection [1] obtained by UC Irvine's group, as well as by the expertise of the University of Zaragoza in robust and real-time Simultaneous Localization and Mapping (SLAM) systems [3].

The tools currently available in each of these domains are not powerful enough alone to account for the diversity and complexity of content typical of real everyday life egocentric videos. Current maps, composed of meaningless geometric entities, are quite poor for performing high-level tasks such as object manipulation. Focusing on functional human activities (that often involve interactions with objects in the near-field), and consequently on dynamic scenes, adds to the challenging and interesting nature of this problem, even from a traditional SLAM perspective.

On the other hand, ADL detectors perform poorly in the case of small objects occluded by other surrounding objects or by the user's body parts. Research breakthroughs are thus required, not only in vision-based ADL recognition and SLAM, but also in exploiting the synergy of the combination.

EgoVision4Health is financed by the European Commission under FP7-PEOPLE-2012-IOF through grant PIOF-GA-2012-328288.

Links:

http://www.gregrogez.net/research/egovision4health/

http://cordis.europa.eu/projects/328288

References:

[1] H. Pirsiavash, D. Ramanan: "Detecting activities of daily living in first-person camera views", in proc. of IEEE CVPR, 2012, pp. 2847-2854

[2] N. Fioraio, L. Di Stefano, "Joint Detection, Tracking and Mapping by Semantic Bundle Adjustment", in proc. of IEEE CVPR, 2013, pp. 1538-1545

[3] J. Civera, A. J. Davison, J. M. M. Montiel: "Structure from Motion using the Extended Kalman Filter", Springer Tracts in Advanced Robotics 75, Springer 2012, pp. 1-125.

Please contact:

Grégory Rogez
Aragon Institute for Engineering Research (i3A),
Universidad de Zaragoza, Spain
E-mail: grogez@unizar.es

Applying Random Matrix Theory Filters on SenseCam Images

by Na Li, Martin Crane, Cathal Gurrin and Heather J. Ruskin

Even though Microsoft's SeneseCam can be effective as a memory-aid device, there exists a substantial challenge in effectively managing the vast amount of images that are maintained by this device. Deconstructing a sizeable collection of images into meaningful events for users represents a significant task. Such events may be identified by applying Random Matrix Theory (RMT) to a cross-correlation matrix C that has been constructed using SenseCam lifelog data streams. Overall, the RMT technique proves promising for major event detection in SenseCam images.

Microsoft's SenseCam is a lifelogging camera with a fisheye lens that is worn, suspended around the neck, to capture images and other sensor reading in an automatic record of the wearer's every moment. SenseCam can thus collect a large amount of data, even over a short period of time, with a picture typically taken every 30 seconds, and an average of 4,000 images captured in a typical day. Even though experience shows that the SenseCam can be an effective memory-aid device, serving to improve recollection of an experience, users seldom wish to refresh their memory by browsing large image collections. Hence, tools are required to assist in the management, organization and analysis of these large data sets, eg, to automatically highlight key episodes and, ideally, to classify these in terms of importance to the life logger. Our previous research has shown that SenseCam time series exhibit a strong long-range correlation, indicating that the series do not constitute a random walk, but are cyclical, with continuous low levels of background information picked up constantly by the device [1]. Further, we have shown that a cross-correlation matrix can be analysed to highlight key episodes, thus identifying boundaries between daily events [2].

However, due to the finite length of time series available to estimate cross correlations, the matrix contains "random" contributions. As a consequence, a percentage of noise or routine event inclusion in processing is inevitable. (This phenomenon can also be observed in other domains such as the analysis of financial data and wireless communications [3]). A well-known method for addressing this type of problem is to apply Random Matrix Theory (RMT).

The aim is to compare the properties of the cross-correlation matrix C with those of the random correlation matrix R, separating the content of C into two groups: (a) the part reflecting properties of R ("noise") and (b) the part that deviates from R (and contains information on major events). Figure 1 compares the probability distribution of a typical cross-correlation matrix with that for the random correlation matrix. We note the presence in the former of a well-defined "bulk" of eigenvalues, which fall within the bounds for the latter. We also note deviations for a number of the largest and smallest eigenvalues. This suggests that the cross-correlation matrix captures many major events from the data stream, but also contains substantive noise.

The deviations of the probability distribution of the cross-correlation matrix from RMT suggest that such deviations should also be observed in the corresponding eigenvector components. In order to interpret the meaning of the

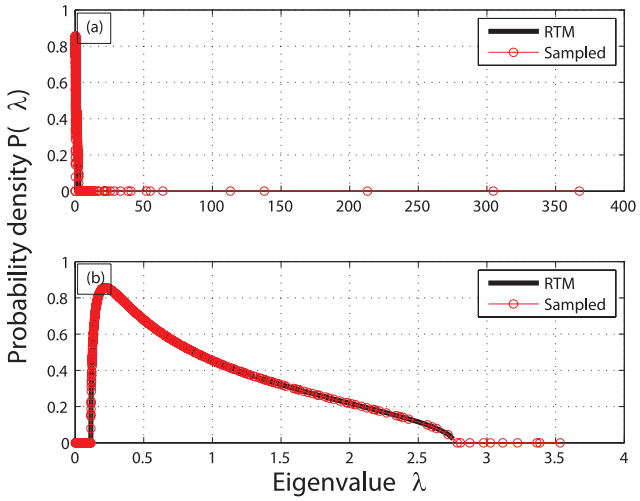


Figure 1: Eigenvalue Distribution for the Cross-Correlation Matrix C for SenseCam data, (a) Full spectral distribution and (b) Partial spectral distribution.

deviating eigenvectors, we note that the largest eigenvalue is an order of magnitude larger than the others. Thus, in order to analyse the contents of the remaining eigenvectors, the effect of the largest eigenvalue should first be removed. This is achieved through linear regression of the image number time series on the additive term common to all images in a similar way to that carried out by V. Plerou et al [3]. Next, we analyse the distribution of the eigenvector components of the largest, the second largest and third largest eigenvalues remaining. Figure 2 shows evidence of distinctive clustering for the distribution of eigenvector components. By examining the images, we find that each cluster involves similar light levels for a group of events. For instance in Figure 2 (a) the distribution of the eigenvector components for the largest remaining eigenvalue is related to clustering for Events 2, 6 & 12: Event 2 described the user going from outside into the office; the office is dark; switching on the lights thus caused a marked change in light levels. Broadly similar circumstances apply for Events 6 and 12; on the user walking into the building, the camera sometimes captured the lights, but at other times did not. On the user preparing to leave the office and collecting belongings, standing or stooping, the camera captured the lights from the ceiling. Several images also captured the user's bag (dark in colour), which again gives rise to contrast in the sequence of images. A similarly clustered distribution also emerges for the other deviating eigenvalues, shown in Figure 2 (b) and (c). While variations are small compared to those found for the second largest eigenvector, this does suggest that these smaller eigenvectors may carry additional information on supporting or lead in, lead out events.

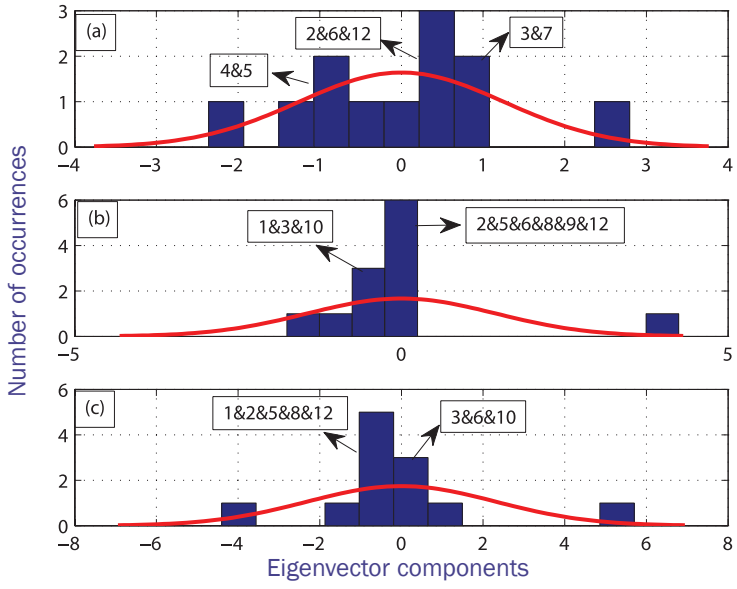


Figure 2: Distribution of eigenvector components: (a) the largest remaining eigenvalue; (b) the second largest remaining eigenvalue; and (c) the third largest remaining eigenvalue.

Overall, the RMT technique can serve as a fairly subtle tool for extraction of major events or key episodes and noise from the SenseCam cross-correlation matrix, with significant deviations from RMT predictions observed. Further, by examining the eigenvectors corresponding to the images, different light levels are identified to be a key source of deviation from randomness, in terms of contribution from the second largest eigenvector. In addition, similar events can be successfully and distinctively clustered. Subsequent deviating eigenvectors exhibit similar clustering and appear to presage or anticipate major events. Future work includes broadening both the user sample and the scenario range to explore the robustness of these findings.

N. Li would like to acknowledge generous support from the SCI-SYM Centre Small Scale Research Fund.

Links:

http://research.microsoft.com/en-us/um/cambridge/projects/sensecam/
http://www.computing.dcu.ie/~msc/

References:

[1] N. Li, et al.: "Automatically Detecting 'Significant Events' on SenseCam", ERCIM News 2011 (87) 2011
[2] N. Li, et al.: "Multiscaled Cross-Correlation Dynamics on SenseCam Lifelogged Images". MMM'13: Proceedings of the 19th International Conference on Multimedia Modeling, Springer Verlag, 1 2013, pp 490-501
[3] V. Plerou, et al.: "Random Matrix approach to cross correlations in financial data", Physical Review E (65) (6), 066126+.

Please contact:

Na Li, Martin Crane, Cathal Gurrin and Heather J. Ruskin
Dublin City University, Ireland
E-mail: na.li@dcu.ie, mcrane@computing.dcu.ie, cgurrin@computing.dcu.ie, hruskin@computing.dcu.ie

Multi-Modal Human Behaviour Analysis from Visual Data Sources

by Sergio Escalera Guerrero

The Human Pose Recovery and Behaviour Analysis group (HuPBA), University of Barcelona, is developing a line of research on multi-modal analysis of humans in visual data. The novel technology is being applied in several scenarios with high social impact, including sign language recognition, assisted technology and supported diagnosis for the elderly and people with mental/physical disabilities, fitness conditioning, and Human Computer Interaction.

The analysis of human actions in visual and sensor data is one of the most challenging topics in Computer Vision. Recently driven by the need for user friendly interfaces for the new generation of computers comprising smart phones, tablets and game consoles, the field of gesture recognition is thriving. Given the inherent difficulties of automatically analysing humans in standard images, alternative visual modalities from different input sensors, including 3D range sensors, infrared or thermal cameras, have attracted a lot of attention. The next challenge is to integrate and analyse these different modalities within the smart devices.

The Multi-modal Human Pose Recovery and Behaviour Analysis project, developed by the HuPBA group and the Computer Vision Centre at the University of Barcelona, has been running since 2009. It is a frontier research project addressing the development of novel computer vision and pattern recognition techniques capable of capturing human poses from data captured via digital images, depth maps, thermal data, and inertial sensors, under a variety of conditions including: changes in illumination, low resolution, appearance, partial occlusions, presence of artefacts, and changes in point of view. Moreover, the project aims to develop new machine learning techniques for the analysis of temporal series and behaviours. The main members of the group are: Albert Clapes, Xavier Pérez-Sala, Victor Ponce, Antonio Hernández-Vela, Miguel Angel Bautista, Miguel Reyes, Dr. Oriol Pujol, and Dr. Sergio Escalera. To date, in collaboration with different universities (Aalborg, Berkley, Boston, Carnegie Mellon), we have developed a range of new generation applications, including:

- Automatic sign language recognition: Using RGB data to detect hands and arms, track their trajectory and automatically translate gestures to words within a vocabulary of 20 lexicons from the Spanish Sign Language.
- Assisted technology for the elderly and people with mental and physical disabilities: Using RGB-Depth and inertial sensors (such as accelerometers) to detect and recognize users within an indoor environment in a non-invasive way, identifying daily activities and objects present in the scene, and providing feedback by means of audio-visual reminders, including alerting the family/specialist if a risk event is produced. This system was designed in collaboration with the Spanish Government.

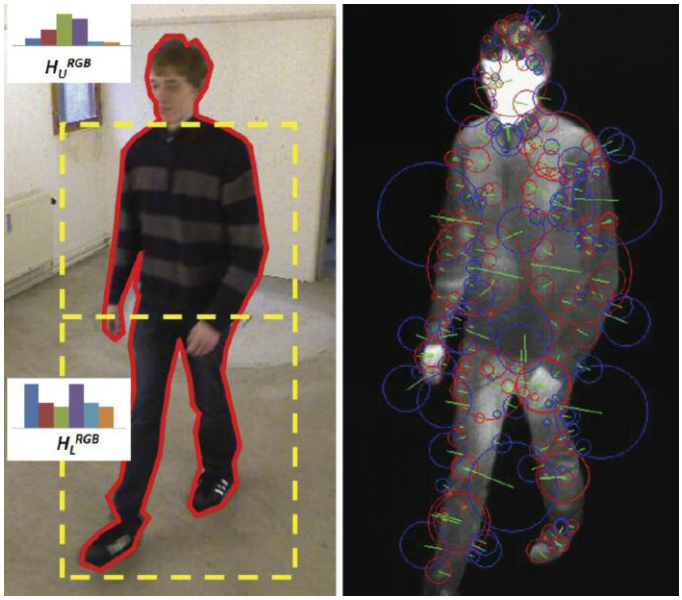


Figure 1: Features for biometry analysis in RGB and thermal data.

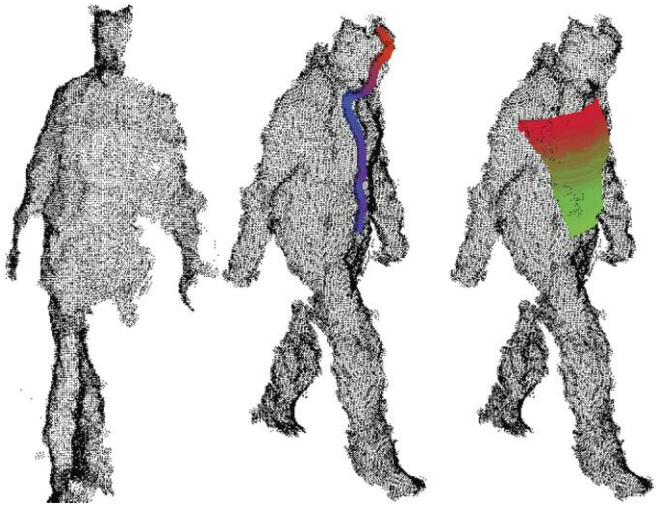


Figure 2: Features for biometry analysis in depth data.

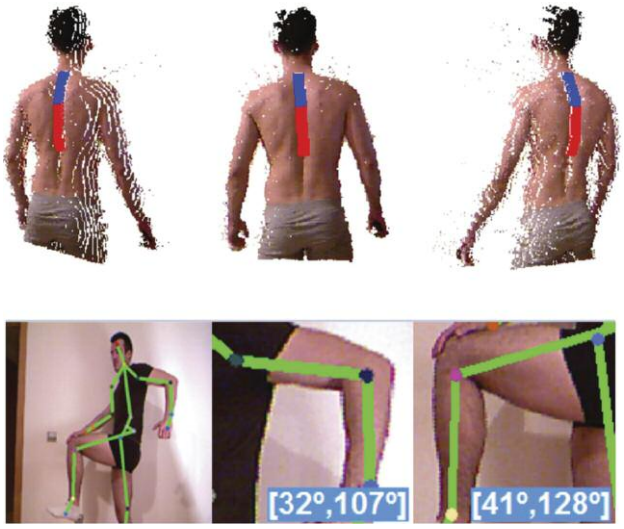


Figure 3: Automatic RGB-Depth analysis of a patient's spine and the range of movement estimation of different body articulations.

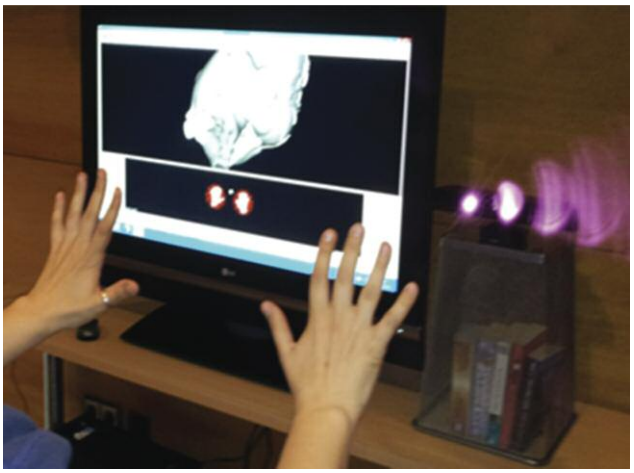


Figure 4: Example of recognized hand poses in a medical volume navigation scenario.

- Intelligent security 2.0: Using bi-modal RGB-Depth and tri-modal RGB-Depth-Thermal analysis to recognize users and object membership in different environments so that thefts can be automatically recognized even with no light in the scene. Figures 1 and 2 show an example of the automatically computed soft biometrics from RGB, thermal, and depth input data sources, respectively [1,2].

- Supported diagnosis in psychiatry: Using RGB-Depth computer vision techniques to automatically recognize a set of behavioural indicators related to Attention Deficit Hyperactivity Disorder (ADHD). As a result, our system is able to summarize the behaviour of children with ADHD diagnosis, giving support to psychiatrists in both diagnosis and evolution analysis during the treatment. This collaboration was performed with Tauli Hospital in Barcelona.

- Intelligent assistance in physiotherapy, rehabilitation, and fitness conditioning: Using RGB-Depth data from patients in order to detect particular body relations and range of movements, giving support in the diagnosis and evolution of physiotherapy and rehabilitation treatment of patients with muscle-skeletal disorders. The system was designed in collaboration with the Instituto de Fisioterapia Global Mezieres (IFGM) in Barcelona. Validations of the method in sport centres demonstrated its reliability in analyzing athletes' training programs and making recommendations for performance improvements. Figure 3 shows an example of the system analysis [3].

- Human Computer Interaction systems: We have applied human hand pose recognition techniques from depth data in order to define a new generation of human computer interaction interfaces. This design has been tested for retail, medical image navigation, and in living labs. Figure 4 shows an example of the system.

From a leadership point of view, our group won the Human Layout Analysis challenge in the Pascal 2010 VOC (http://pascallin.ecs.soton.ac.uk/challenges/VOC/voc2010/), and achieved the 3rd prize in the Kinect© 2012 challenge at the ICPR conference in Japan (http://gesture.chalearn.org/dissemination). Currently, we continue performing research on novel multi-modal descriptors and human behaviour analysis in time series. As a result of our research, we expect to increase the number of real applications in leisure, security and health, and transfer them to society.

References:
[1] Albert Clapés, Miguel Reyes, Sergio Escalera: "Multi-modal User Identification and Object Recognition Surveillance System", PRL, 2013.
[2] A. Møgelmose, C. Bahnsen et al.: "Tri-modal Person Re-identification with RGB, Depth and Thermal Features, Perception Beyond the Visible Spectrum", CVPR, 2013.
[3] M. Reyes, A. Clapés et al.: "Automatic Digital Biometry Analysis based on Depth Maps, Computers in Industry", 2013.

Please contact:
Sergio Escalera Guerrero
Universitat de Barcelona, Computer Vision Centre, Spain
Tel +34 93 4020853
E-mail: sergio@maia.ub.es,
http://www.maia.ub.es/~sergio/

Tracking the Articulated Motion of Human Hands in 3D

by Iason Oikonomidis, Nikolaos Kyriazis and Antonis A. Argyros

The FORTH 3D hand tracker recovers the articulated motion of human hands robustly, accurately and in real time (20Hz). This is achieved by employing a carefully designed model-based approach that capitalizes on a powerful optimization framework, GPU processing and the visual information provided by the Kinect sensor.

Humans use their hands in most of their everyday life activities. Thus, the development of technical systems that track the 3D position, orientation and full articulation of human hands from markerless visual observations can be of fundamental importance in supporting a diverse range of applications.

Developing such a system is a complex task owing to a number of complicating factors such as the high dimensionality of dexterous hand motion, the ambiguity in identifying hand parts because of their colour uniformity and the absence of observations when fingers or the palm occlude one another. Last but not least, hands often move fast and tracking needs to be performed by relatively low resolution cameras that are placed at a considerable distance from the scene.

To alleviate some of these problems, some successful methods employ specialized hardware for motion capture and/or visual markers. Unfortunately, such methods require a complex and costly hardware setup that interferes with the observed scene. Furthermore, there is limited potential to move such methods out of the lab. Several attempts have been made to address the problem by considering markerless visual data only. Existing approaches can be categorized into appearance- and model-based.

Appearance-based methods have much lower computational cost and hardware complexity but they recognize a discrete number of hand poses that correspond typically to the method's training set. Model-based implies that the appearance of an arbitrarily positioned and articulated hand can be simulated and compared to actual observations. 3D hand tracking is then performed by systematically searching for hand motions whose simulated appearance best matches the observed images. Such methods provide a continuum of solutions but are computationally costly.

The FORTH 3D hand tracker is an accurate, robust and real time model-based solution to the problem of 3D hand tracking. The input is provided by a Kinect sensor and comprises an RGB image and a depth map, which assigns a depth measurement to each RGB value (Figures 1a and 1b). Skin colour detection is used to isolate the hand in the RGB and depth images (Figure 1c). The adopted 3D hand model comprises a set of appropriately assembled geometric 3D primitives (Figure 1d). Each hand pose is represented as a vector of 27 parameters: three for global position; four for global orientation (quaternion representation); and 20 for the relative articulation of the fingers.

Having a parametric 3D model of a hand, the goal is then to estimate the 27 parameters that make the model most compatible to the visual observations. Compatibility between observations and hypotheses is judged by differentiating depth maps, pixel by pixel. Actual observations already contain depth maps. Hand hypotheses are converted to depth maps using 3D rendering. Comparable 3D rendering is made possible by the established 3D hand model and the knowledge of the camera parameters. Simply stated, we can approximate how a hand in 3D might look through the Kinect sensor. The sum of pixel-wise differences constitutes the objective function and is parameterized over hand configurations. Optimization of this function is performed with a variant of Particle Swarm Optimization (PSO). The result of this optimization is the output of the method for a given frame (Figure 1e). Temporal continuity is exploited to track hand articulation in a sequence of frames.

The computationally demanding parts of the process have been implemented so as to run efficiently on a GPU. The resulting system tracks hand articulations with an accuracy of 5mm at a rate of 20Hz on a quad core Intel i7 920 CPU with 6GB RAM and an NVidia GTX580 GPU. Better accuracy can be traded off for slower tracking rates.

Our work [1] (Figure 2a) is the first to demonstrate that a model-based

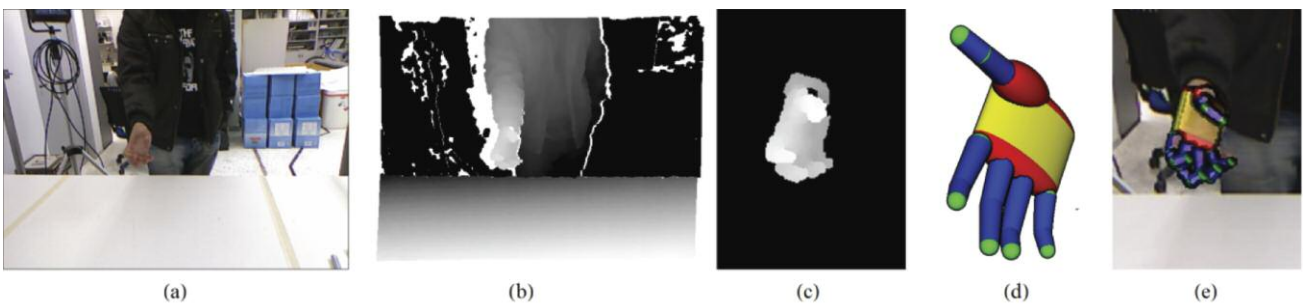


Figure 1: Graphical illustration of the operation of the FORTH 3D hand tracker. A Kinect RGB image (a) and the corresponding depth map (b). The hand is segmented (c) by jointly considering skin color and depth. The proposed method fits the employed hand model (d) to this observation recovering the hand articulation (e).

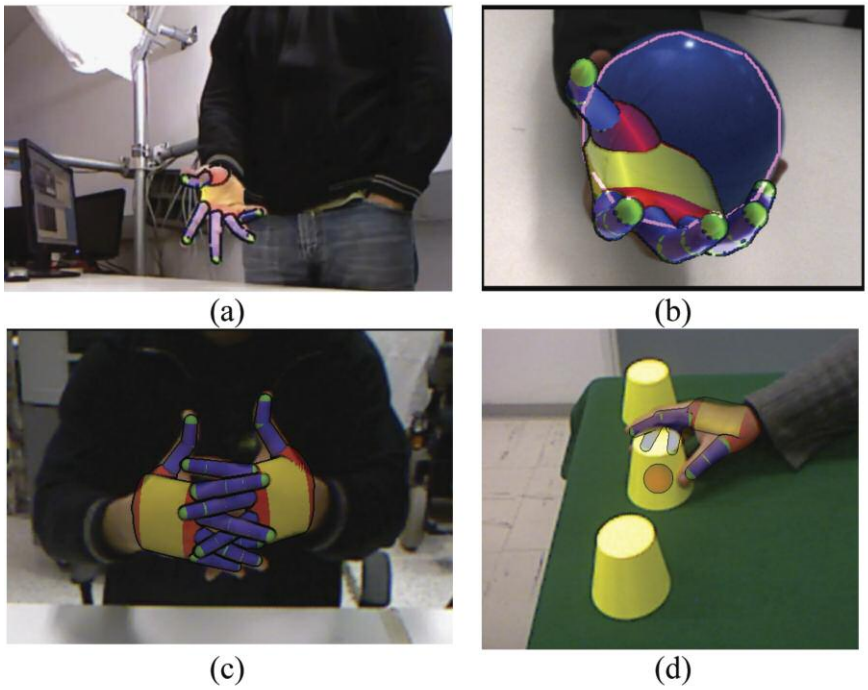


Figure 2: The framework developed has been employed to track (a) single hand (b) a hand interacting with an object (c) two strongly interacting hands and (d) the state of a complex scene where a hand interacts with several objects.

approach can produce a practical hand tracking system. The core of the tracking framework has also been employed to provide state-of-the-art solutions for problems of even higher dimensionality and complexity, e.g., for tracking a hand interacting with an object [2] (Figure 2b), for tracking two strongly interacting hands (Figure 2c) and for tracking the state of a complex scene where a hand interacts with several objects [4] (Figure 2d).

The FORTH 3D hand tracker has been implemented as a library and can be downloaded as a middleware of the OpenNI 2.0 framework, so that it can be used freely for research purposes. Several researchers around the world have already used this library to support their research. Additionally, several European, American and Asian companies have expressed interest in using it as enabling technology for developing applications in the fields including health, education, gesture recognition, gaming and automated robot programming.

An extension of our work was awarded the 1st prize at the CHALEARN Gesture Recognition demonstration competition, organized in conjunction with ICPR 2012 (Tsukuba, Japan, Nov. 2012).

This research is partially supported by the EU-funded projects GRASP, FP7-ROBOHOW.COG and WEARHAP.

Links:
FORTH 3D Hand Tracker web page:
http://cvrlcode.ics.forth.gr/handtracking/
FORTH 3D Hand Tracker OpenNI middleware library download:
http://www.openni.org/files/3d-hand-tracking-library/

References:
[1] I. Oikonomidis, N. Kyriazis, A. Argyros: “Efficient model-based 3d tracking of hand articulations using Kinect”, in BMVC 2011, http://youtu.be/Fxa43qcm1C4

[2] I. Oikonomidis, N. Kyriazis, A. Argyros: “Full DOF tracking of a hand interacting with an object by modeling occlusions and physical constraints”, in ICCV 2011, http://youtu.be/e3G9soCdIbc

[3] I. Oikonomidis, N. Kyriazis, A. Argyros: “Tracking the articulation of two strongly interacting hands”, in CVPR 2012, http://youtu.be/N3ffgj1bBGw

[4] N. Kyriazis, A. Argyros, Physically Plausible 3D Scene Tracking: The Single Actor Hypothesis. In CVPR 2013, http://youtu.be/0RCsQPXeHRQ

Please contact:
Antonis Argyros
ICS-FORTH, Greece
Tel: +30 2810 391704
E-mail: argyros@ics.forth.gr

KAD - An Intelligent System for Categorizing and Assessing the State of Patients with Multiple Sclerosis

by Spiros Fotopoulos and Dimitrios Kastaniotis

Neurological disorders can be reliably assessed using the low-cost Microsoft Kinect depth sensor to record human gait, coupled with our Kinect Assessment Disorders (KAD) system to process the information.

In recent years, computer science researchers, keen to develop more natural and non-intrusive means of human-machine interaction, have become interested in analysing human motion. Although some impressive methods have been proposed, the integration of such systems has been stalled by several factors: in particular the variability in scale, colour and texture of the captured scenes makes the problem almost intractable.

Recently, with the launch of low cost depth sensors, such as Microsoft's Kinect, the landscape of Human Machine Interface (HMI) has started to change. This breakthrough in the history of human-machine interaction has triggered the design of new algorithms both for the task of pose estimation in real time and for extracting high level information from actions.

Beyond the area of HMI, Kinect also finds applications in human body gait analysis [1], for example, recent research has shown that gender recognition is feasible using recordings from the Kinect sensor [2] by utilizing the poses estimated by the Kinect SDK pipeline together with a novel feature representation. Thus we took advantage of the Kinect sensor's capabilities in order to design novel techniques for processing human motion with the aim of assessing the state of patients with multiple sclerosis (MS).

Inspired by recent advances, our team, in collaboration with the Patras University Hospital (Department of Neurology), has been collecting data from patients diagnosed with MS. Using a special arrangement of Kinect sensors and a specialized computer program based on the Microsoft's Kinect SDK, we are recording depth data, as well as the skeletal poses estimated by the Microsoft SDK pipeline. The experimental setup is located at the University Hospital of Patras, where a number of patients are being recorded every week following a purpose-designed protocol.

Our team is focusing on the development of "KAD", a novel computer aided diagnosis system that can learn how to distinguish between different populations of patients by imitating a human expert's ability to assess gait. This can be achieved, for example, by applying the "two minute test", the established method used by specialists. This project started in early 2012, and it is expected to conclude by the end of 2015.

In order to distinguish between different populations of MS patients, we utilize state of the art pattern recognition techniques [3] that are appropriate for applying to a small number of samples per subject (or class) and that have been proven to work in similar tasks [3]. Our initial work on action recognition and gait-based gender recognition [2] already shows promising results. Namely, we have already proposed

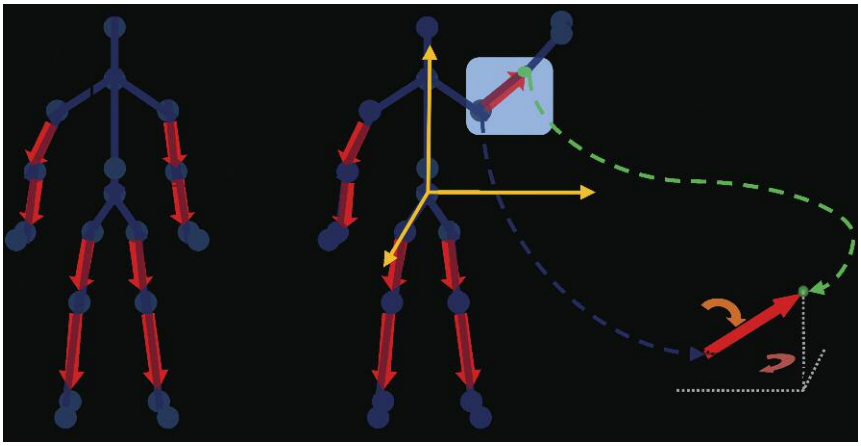


Figure 1: The proposed (view invariant) Euler representation of the eight skeletal primitives. The 'xyz' corresponds to the Kinect coordinate system, and the "rut" is a view invariant estimated by our algorithm.

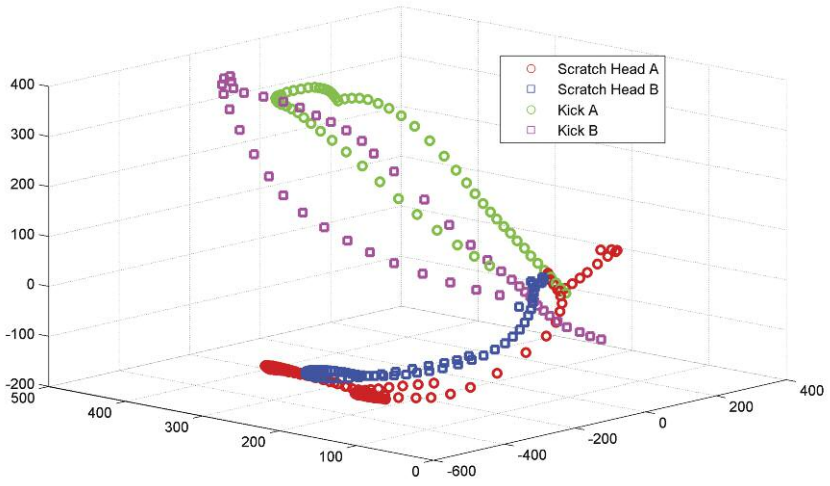


Figure 2: Trajectories of Euler angles for two subjects performing two different actions.

methods for capturing the dynamical characteristics of human motion by encoding the eight skeletal primitives depicted in Figure 1, using Euler angular representation. This results in the trajectories depicted in Figure 2, which seem to be very efficient at capturing the intrinsic parameters of human motion. These results lend strength to our hypothesis that intrinsic parameters of human gait can be captured, and this ability may allow us to differentiate among several types and states of MS patients.

Our project belongs to the field of Computer Assisted Diagnosis. In the future, we plan to extend KAD to other kinds of neurological disorders that affect human motion. In particular we plan to focus on the early diagnosis of several neurological diseases by using videos captured from a consumer device (Kinect) or even from gait sequences recorded in a laboratory.

References:

[1] M. Gabel et al: "Full body gait analysis with Kinect", Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE, pp.1964-1967, 2012

[2] D. Kastaniotis, et al: "Gait-based Gender Recognition using Pose Information for Real Time Applications", DSP 2013, in press

[3] Theodorakopoulos, et al: "Pose-based Human Action Recognition via Sparse Representation in Dissimilarity Space", J. Vis. Commun. Image R. (2013).

Please contact:
Spiros Fotopoulos
University of Patras, Greece
Tel: +30 2610 99 6058
E-mail: spiros@physics.upatras.gr

GAIMS: A Reliable Non-Intrusive Gait Measuring System

by Sébastien Piérard, Samir Azrour, Rémy Phan-Ba and Marc Van Droogenbroeck

Gait observation and analysis can provide invaluable information about an individual [1]. Studies that have interpreted gait using traditional imaging devices have demonstrated that it is difficult to make reliable measurements with colour cameras. GAIMS, our new system resulting from a multidisciplinary project born from collaboration between engineers and neurologists, aims at developing non-intrusive and reliable tools to provide quantitative measures of gait and interpretations of the acquired data. Following a current trend in imaging, it takes advantage of imaging sensors that measure distance instead of colour. While its principles are general, GAIMS is currently used for the diagnosis of multiple sclerosis (MS) and the continued evaluation of disease progression [2]. It is the first available system to fully satisfy the clinical routine and its associated constraints.

Gait is a complex process involving neurology, mechanics and muscle. Gait is unique to the individual [1], and depends on many factors including: age, sex, weight, height, the presence of disease, and medication. Because the gait is so dependent on these factors, it is possible to infer some information about an individual based on his gait.

While colour images have been used intensively for gait analysis and recognition, they are inadequate for deriving precise quantitative measurements and objective information about the walking person. Therefore, at the University of Liège, we have developed GAIMS, a non-intrusive gait analysis and imaging system based primarily on range laser sensors. In our setup, several range laser scanners, which can be seen as 3D cameras reduced to one line, are combined to reconstruct a horizontal slice of the scene, at a height of 15 cm above the floor. While we hope to be able to build denser 3D information utilizing all the information provided by 3D cameras in the future, in its current form GAIMS is able to reconstruct the trajectories of both feet, and to provide derived gait measures.

GAIMS has numerous advantages. It provides a range of information, such as the location of the feet, the walking speed, the spacing between the feet, the double support duration, etc, which can be combined with images to form a bridge between the observation and interpretation of complicated scenarios such as gait analysis. Knowledge about the location and dynamics of individuals within the scene is obtained without the need for the subjects to wear any sensor, marker, or reflector. GAIMS is therefore a valuable tool in applications such as trajectory analysis for behaviour interpretation or clinical evaluation. In addition, GAIMS operates over a large area (about 10m wide) and it is almost insensitive to ambient lighting.

For the interpretation, we have developed, on top of our system, a set of intelligent tools dedicated to the specific data type, that combine filtering, tracking, machine learning, and some heuristics. As the relationship between the factors influencing the gait and the gait itself is complex and still largely unknown, gait interpretation is a delicate task: an expert is usually able to

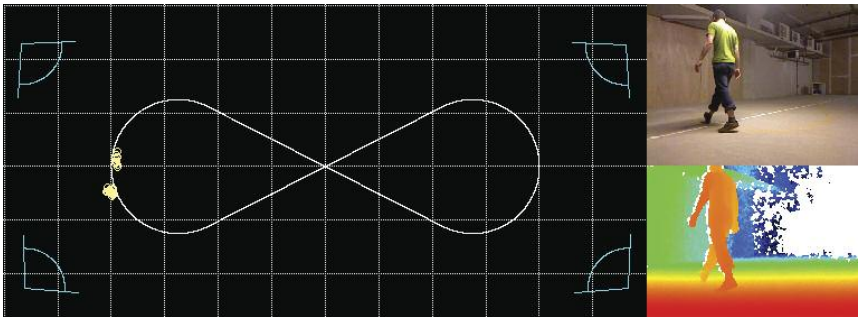


Figure 1: Illustration of GAIMS: Left: A bird's-eye view of the scene, with four sensors depicted in turquoise and the horizontal cross-section of the walking person's legs in yellow. Right: Synchronized colour and range images acquired by a kinect.

interpret the observations, but rarely able to explain how he proceeds.

GAIMS can help with the diagnosis of multiple sclerosis (MS) and for the continued evaluation of the disease's progression [2]. The ambulation impairments appear in the early stages of the disease and the magnitude of gait modification is a good indicator of disease activity. The current standardized tests used to assess gait in MS are still based largely on a measure of speed, and a simple stop watch is often the sole measuring instrument used by the neurologist. This is too limiting. GAIMS is the first powerful system that fully satisfies the clinical routine constraints and is compatible with existing standardized tests. Its effectiveness has been shown through a threefold validation procedure: (1) GAIMS has been demonstrated to detect subtle (intra-subject) gait alterations and to be more effective than a stop watch in detecting gait abnormalities. (2) It is possible to differentiate between healthy individuals and MS patients based on gait characteristics measured by GAIMS. (3) GAIMS can also measure significant differences between MS patients with different disease states.

While it has been successfully used for multiple sclerosis analysis, GAIMS is not limited to medical applications. For example, we have used it in conjunction with 3D cameras and projectors to create a 3D immersive and interactive system that gives the illusion of a 3D virtual world all around the user [3]. GAIMS has the power to tell you precisely where the user is located. This is useful for real time scene interpretation and interactive applications.

References:

[1] N. Boulgouris, D. Hatzinakos, and K. Plataniotis: "Gait recognition: a challenging signal processing technology for biometric identification", IEEE Signal Processing Magazine, 22(6):78– 90, Nov. 2005

[2] S. Belachew et al: "Multimodal evaluation of gait and stride dynamics in relapsing and progressive forms of multiple sclerosis", in proc. of the Belgian Royal Academies of Medecine, 1:66–69, 2012

[3] S. Piérard et al: "I-see-3D! An interactive and immersive system that dynamically adapts 2D projections to the location of a user's eyes", in proc. of IC3D, Liège, Belgium, December 2012.

Please contact:

Sébastien Piérard
INTELSIG Laboratory, Montefiore Institute, University of Liège, Belgium
E-mail: sebastien.pierard@ulg.ac.be

Mixed Reality by Understanding and Integrating Spatio-Temporal Data of a LIDAR and a 4D Studio

by Csaba Benedek, Zsolt Jankó, Dmitry Chetverikov and Tamás Szirányi

Two labs of SZTAKI have jointly developed a system for creation and visualization of mixed reality by combining the spatio-temporal model of a real outdoor environment with the models of people acting in a studio. We use a LIDAR sensor to measure an outdoor scene with walking pedestrians, detect and track them, then reconstruct the static part of the scene. The scene is then modified and populated by human avatars created in a 4D reconstruction studio.

Real-time reconstruction of outdoor dynamic scenes is essential in intelligent surveillance, video communication, mixed reality and other related applications. Compared to conventional 2D video streams, 3D data flows of real-world events (4D scenes) provide a more authentic visual experience and additional functionality. A reconstructed 4D scene can be viewed and analysed from any viewpoint. Furthermore, it can be virtually modified by the user. However, building an interactive 4D video system is a challenging task that requires processing, understanding and real-time visualization of a large amount of spatio-temporal data.

This issue is being addressed by "Integrated 4D", a joint internal R&D project of two labs of SZTAKI. We have built an original system for spatio-temporal (4D) reconstruction, analysis, editing, and visualization of complex dynamic scenes. The i4D system [1] efficiently integrates two very different kinds of information: the outdoor 4D data acquired by a rotating multi-beam LIDAR sensor, and the dynamic 3D models of people obtained in a 4D studio. This integration allows the system to understand and represent the visual world at different levels of detail: the LIDAR provides a global description of a large-scale dynamic scene, while the 4D studio builds a detailed model, an avatar, of an actor (typically, a human) moving in the studio.

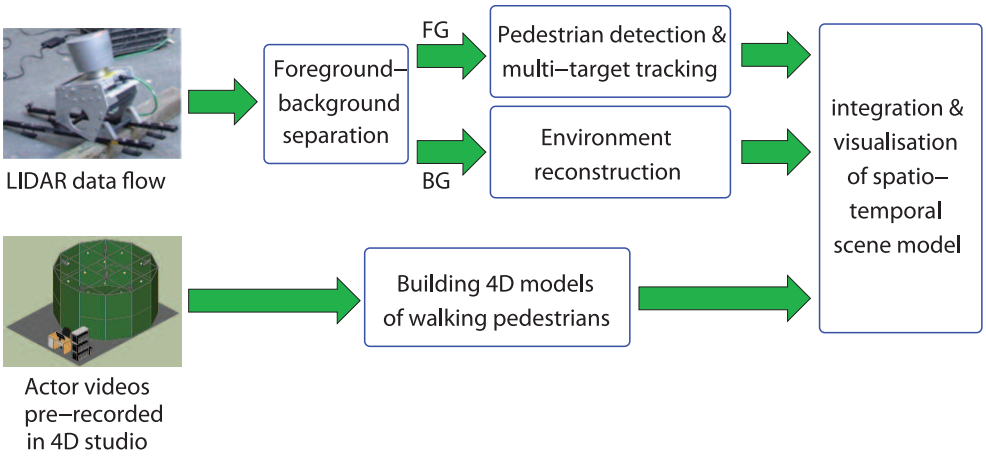


Figure 1: Flowchart of the i4D system. FG is the foreground, BG the background.

In our project, a typical scenario is an outdoor environment with multiple walking people. The LIDAR measures the scene from a fixed position and yields a time-varying point cloud. This data is processed to build a 3D model of the static part of the scene and detect and track the moving people. Each pedestrian is represented by a sparse, moving point cluster and a trajectory. A sparse cluster is then substituted by an avatar created in the 4D studio of MTA SZTAKI. This results in a reconstructed and textured scene with avatars that follow in real time the trajectories of the pedestrians.

Figure 1 shows a flowchart of the i4D system. The LIDAR data flow involves environment scanning for point cloud sequence acquisition, foreground and background segmentation by a robust probabilistic approach, detection and tracking of the moving pedestrians and generating motion trajectories, geometric reconstruction of the ground, walls and other field objects, and, finally, texturing the obtained 3D models with images of the scene. Technical details are given in [2]. Our 4D reconstruction studio creates textured dynamic avatars of people walking in the studio. The hardware and software components of the studio are presented in [3]. The integration and visualization block combines the graphical elements into a joint dynamic scene model, provides editing options and visualizes the final 4D model where the avatars move in the scene along the assigned trajectories.

Sample results of the 4D reconstruction and visualization process are shown in Figure 2. The pedestrian trajectories are obtained by motion-based segmentation and tracking in the dynamic point cloud. Each avatar follows the prescribed 3D trajectory. Its orientation and rotation to the proper direction are automatically determined from the trajectory.

The main novelty of the i4D system is that it integrates two different modalities of spatio-temporal perception operating at different scales. This may open a way towards real-time, free viewpoint, scalable visualization of large time-varying scenes, which is crucial for future mixed reality and multimodal communication applications. We plan to extend the project to LIDAR data collected from a moving platform and add modules for automatic field object recognition and surface texturing.

Link: http://web.eee.sztaki.hu/i4d/

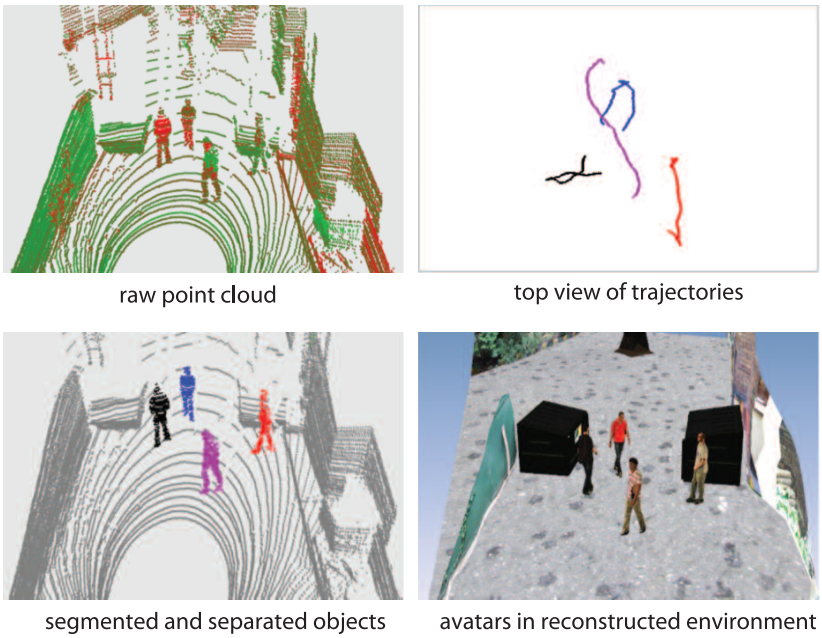


Figure 2: Sample results of object tracking and integrated dynamic scene reconstruction.

References:

[1] "Procedure and system for creating integrated three-dimensional model", in Hungarian, Hun PO: P1300328/4 patent pending, 2013.

[2] C. Benedek et al: "An Integrated 4D Vision and Visualisation System", Computer Vision Systems, Lecture Notes in Computer Science, vol. 7963, pp 21-30, 2013

[3] J. Hapák, Z. Jankó, D. Chetverikov: "Real-Time 4D Reconstruction of Human Motion", Articulated Motion and Deformable Objects, Lecture Notes in Computer Science, vol. 7378, pp. 250-259, 2012.

Please contact:

Csaba Benedek
SZTAKI, Hungary
E-mail: benedek.csaba@sztaki.mta.hu

Visual 3D Environment Reconstruction for Autonomous Vehicles

by Thomas Kadiofsky, Robert Rößler and Christian Zinner

In the foreseeable future it will be commonplace for various land vehicles to be equipped with 3D sensors and systems that reconstruct the surrounding area in 3D. This technology can be used as part of an advanced driver assistance system (ADAS) for semi-autonomous operation (auto-pilot), or for fully autonomous operation, depending on the level of technological maturity and legal regulations. Existing robotic systems are mostly equipped with active 3D sensors such as laser scanning devices or time-of-flight (TOF) sensors. 3D sensors based on stereo cameras cost less and work well even in bright ambient light, but the 3D reconstruction process is more complex. We present recent results from our visual 3D reconstruction and mapping system based on stereo vision, which has been developed within the scope of several research projects.

The system of 3D reconstruction and mapping based on stereo vision, also known as "Simultaneous Localization and Mapping" (SLAM), has largely been addressed with robotic platforms in indoor scenarios [1]. In such man-made environments, various simplifying assumptions can be made (horizontal floor, flat walls and perpendicular geometry). Thus, simple laser devices scanning a single horizontal plane are sufficient to create a 2D map (floor plan). When moving to outdoor environments, which is the working area for the vast majority of vehicles, this kind of 2D representation becomes insufficient and has to be extended to a 2.5D representation called elevation map. 3D sensors are now required to capture the scene as a whole and not just along a single scan line. Stereo vision systems have this ability.

Stereo vision systems project distance values within a solid angle onto an image plane – the result is a depth image. A range of factors in the design of a stereo vision system determine the quality and accuracy of the depth images, which together determine the feasibility of stereo for the intended application: The stereo matching algorithm has to perform at interactive frame rates and at high image resolution. The stereo geometry and camera resolution must provide sufficient depth resolution even for larger ranges. We use a trinocular camera setup with 2-megapixel-class cameras and a large stereo baseline of 1.1 m delivering useful distance data even at ranges >100 m (Figure 1). The stereo matching process is performed by the real-time stereo engine S3E [2].

Single views generated by 3D sensors suffer from limited fields of view and resolutions as well as occlusions.

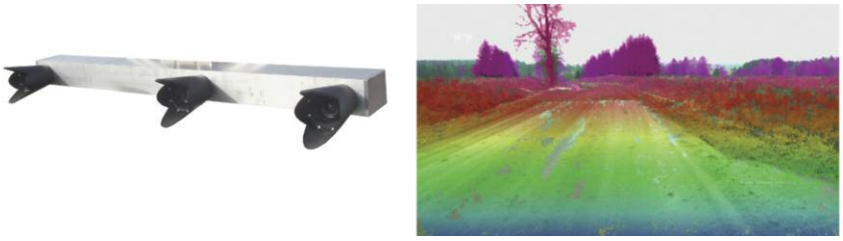


Figure 1: Stereo camera system for land vehicles, camera image overlaid with depth data (blue=near, red=far)

Figure 2: Top: camera image; bottom: map rendered from the camera's perspective.

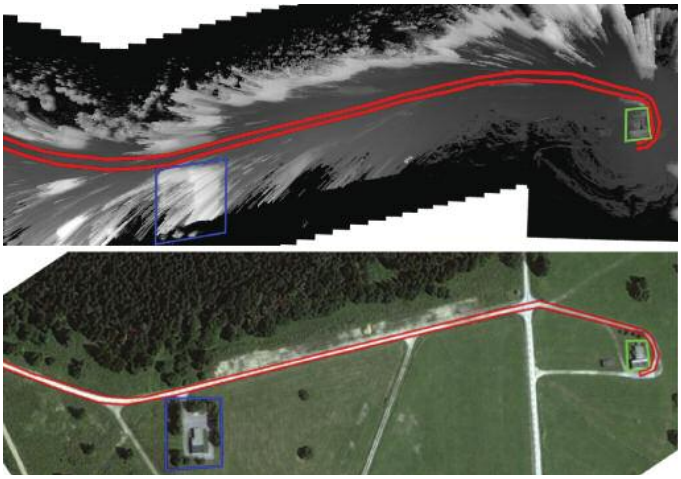


Figure 3: Top: map with elevation (dark=low, bright=high); bottom: corresponding orthophoto from Google maps

Therefore, obtaining a more complete reconstruction of a vehicle's surroundings requires the combination of multiple views from different positions and perspectives. The 3D data has to be processed and fused in a common coordinate frame, but it is typically captured relative to a sensor coordinate system. Hence, an important task is to recover the ego-motion of the vehicle between consecutive measurements and to localize it within the common frame. Currently we are using the following sensors for computing the trajectory:

- The intensity data of the stereo camera is used as input for a visual odometry. Feature points are tracked across consecutive images and the established correspondences are used to compute the motion of the camera over time.
- An attitude and heading reference system (AHRS) uses inertial measurements to provide an orientation.
- Although we are aiming to be independent of GPS, it is currently used as absolute measurement of the vehicle's position.

All of the mentioned sources of the vehicle's pose are noisy and therefore their results are inconsistent. We resolve these inconsistencies by using an Extended Kalman Filter (EKF) that computes the most probable position and orientation of the vehicle at a given sampling rate, based on the input measurements and their noise described by their covariance matrices. The relative positions of the sensors, which are rigidly mounted on the vehicle to one other, are determined in an offline calibration process.

To support terrains of varying morphology we have chosen the 2.5D digital elevation model for map data representation. In addition to height information a cost map encoding the traversability of the terrain is also generated [3]. For instance grassland and a path are both quite flat, but the path should be preferred over grassland in a motion planning module. Another important feature of the representation is its ability to handle changes in the scene in order to correctly represent features such as moving obstacles.

Example visualizations of map data that have been generated by our system in real-time are shown in Figures 2 and 3.

This research has been funded by the Austrian Security Research Program KIRAS – an initiative of the Austrian Federal Ministry for Transport, Innovation and Technology (bmvit).

Link:
http://www.ait.ac.at/research-services/research-services-safety-security/3d-vision-and-modelling

References:
[1] Andreas Nüchter: "3D Robotic Mapping". Springer, Berlin 2009, ISBN 978-3540898832

[2] M. Humenberger et al.: "A fast stereo matching algorithm suitable for embedded real-time systems", J.CVIU, 114 (2010), 11

[3] T. Kadiofsky, J. Weichselbaum, C. Zinner: "Off-Road Terrain Mapping Based on Dense Hierarchical Real-Time Stereo Vision", Advances in Visual Computing, Springer LNCS, Vol. 7431 (2012).

Please contact:
Christian Zinner, Thomas Kadiofsky, Robert Rößler
AIT Austrian Institute of Technology GmbH, Austria (AARIT)
E- mail: {christian.zinner, thomas.kadiofsky, robert.roessler}@ait.ac.at

Automatic MRI Brain Tissue Classification

by Loredana Murino, Umberto Amato and Bruno Alfano

Involvement and morphological changes of brain structures both in aging processes and in neurodegenerative diseases can be analysed using Magnetic Resonance imaging. Our aim is to automate the procedure through supervised brain tissue classification.

Magnetic Resonance (MR) has proven to be one of the most promising diagnostic modalities for the study of the brain. This highly flexible technique makes it possible to examine the brain's interior structures as well as its activity at rest or under specific stimuli.

Automatic tissue recognition (also known as segmentation) is an important and challenging task in the processing and analysis of brain MR images, and during the last decade a large number of software packages have been developed to classify the three major brain tissues: grey matter, white matter and cerebro spinal fluid.

There is a growing interest in understanding the involvement and the possible morphological changes of minor brain structures both in normal aging processes and in the presence of neurodegenerative diseases. While the segmentation of major tissues has been intensively investigated, our target is tissue classification and segmentation of subcortical structures for which only a few tools are available.

Our study is carried out at the Istituto per le Applicazioni del Calcolo "Mauro Picone" (IAC-CNR, Naples) in collaboration with the Istituto di Biostrutture e Bioimmagini (IBB-CNR, Naples) under the project MEdical Research in ITaly (MERIT) funded by the Italian Ministry of University and Research. It focuses on the classification of multispectral MR brain images using supervised methods. This requires a training data set (gold standard), comprising a set of voxels and corresponding intensities for which the tissue is "certainly" known. Classifiers learn characteristics of all tissues from this data set and they are then able to classify voxels of new data sets.

As a gold standard we have considered a recently developed three-dimensional digital brain phantom (DBP) [1]. The

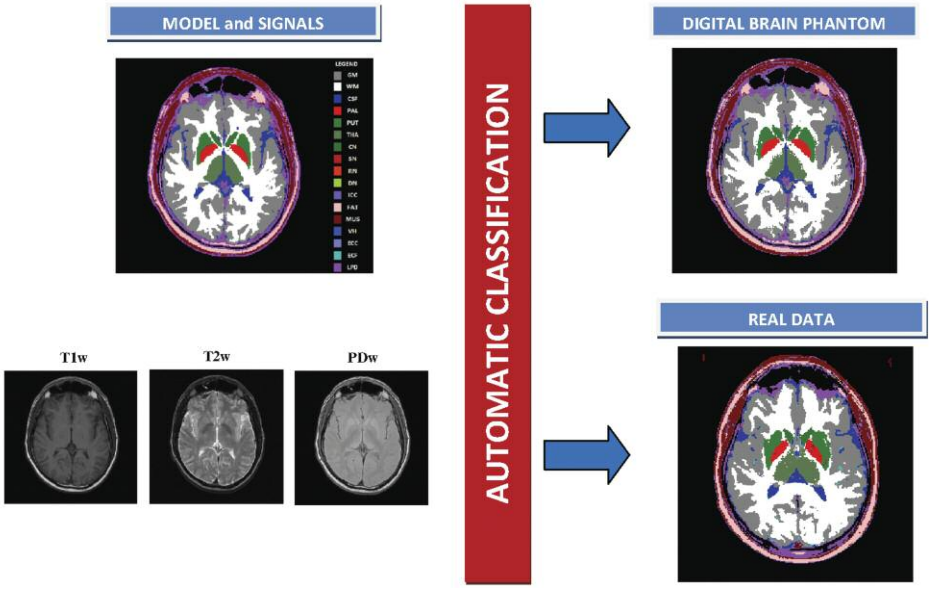


Figure 1: A single slice of the digital brain phantom and the corresponding MRI T_{1W}, T_{2W} and P_{DW} signals (left), automatic tissue classification obtained (right).

following 17 tissues are defined in the DBP: 11 intracranial tissues (gray matter, white matter, cerebro spinal fluid, pallidus, putamen, thalamus, caudate nucleus, red nucleus, dentate nucleus, substantia nigra, and intracranial connective), and five extracranial ones (fat, muscle, vitreous humor, extracranial connective, and extracranial fluid); a further class (LowPD) comprising intra and extra-cranial voxels characterized by a low proton density value is included. The model contains 256 x 256 x 150 near-isotropic 0.9375 x 0.9375 x 1 mm voxels; each voxel is labelled according to its assignment to a unique tissue. The corresponding simulated signals are provided in the form of T_{1W} (510/15ms TR/TE) and P_{DW} - T_{2W} (1867/15-90ms TR/TE) axial slices.

In addition to providing the gold standard, the DBP model also serves as a reference to evaluate the performance of the classification methods.

We compared several classification methods belonging to the family of Discriminant Analysis (DA) (Principal (PCDA), Independent (ICDA) and Oblique (OCDA) Component Discriminant Analysis [2]), with two popular methods from the machine learning literature, namely, K-Nearest Neighbour (KNN) and Support Vector Machine (SVM).

These methods do not use spatial correlation in images, since they classify images voxel by voxel independently. As a consequence, the accuracy of the classification degrades and artefacts are present in classified images (eg, "salt and pepper" noise). Therefore spatial information has been incorporated into the classification scheme and quantitatively assessed. This has been accomplished largely by introducing local prior probabilities in Discriminant Analysis that take into account both the tissues present in the neighbourhood of each voxel and a probabilistic atlas [3] of all the intracranial brain tissues. In a first phase of the project, the training step was made on a subset of the gold standard and classification was performed on the remaining voxels. In a second phase we have considered 20 real datasets from normal volunteers. Each study consists of T_{1W}, T_{2W} and P_{DW} axial slices covering the entire brain, acquired using a Philips Intera 1.5T MR scanner. The training step of classification is performed on full gold standard, whereas actual segmentation is performed on the real studies.

Figure 1 (left) shows a single slice of the digital brain phantom and the corresponding MRI T_{1W}, T_{2W} and P_{DW} signals, whilst on the right the automatic tissue classification obtained applying the OCDA method both to the corresponding slice of the DBP dataset and to a single slice of one selected real study are shown. Inspection of the classification images and analysis of the numerical results (success percentage greater than 90% for both major tissues and subcortical structures) confirm that the supervised methods (in particular OCDA) are able to correctly classify the real studies.

Future work will focus on the classification of pathological studies with the fundamental objective of locating and monitoring lesions in brain tissues.

Link:
http://lab.ibb.cnr.it/

References:
[1] Alfano, B et al.: "An MRI digital brain phantom for validation of segmentation methods", Med. Image Anal., 2011; 15(3):329–39.

[2] Selvan SE et al.: "Descent Algorithms on Oblique Manifold for Source-Adaptive ICA Contrast", IEEE Trans. Neural Netw. Learn. Syst., 2012; 23(12):1930-47

[3] Mazziotta J. et al.: "A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM)", Phil. Trans. R. Soc. Lond. B, 2001; 356(1412):1293–1322.

Please contact:
Loredana Murino, Umberto Amato
CNR -Istituto per le Applicazioni del Calcolo "Mauro Picone", Italy
E-mail: l.murino@na.iac.cnr.it, u.amato@iac.cnr.it

Bruno Alfano
CNR- Istituto di Biostrutture e Bioimmagini, Italy
E-mail: bruno.alfano@ibb.cnr.it

Connected Morphological Operators for Tensor Images

by Jos Roerdink

The processing, analysis, and visualization of tensor images has become very important in many application domains, such as brain imaging and seismology. In a tensor image the value at each pixel is not just a scalar (as in a grey scale image), but a matrix or tensor, hence the name. In the project COMOTI – Connected Morphological Operators for Tensor Images, funded by the Dutch National Science Foundation (NWO), we address the development of techniques for morphological filtering and visualization of tensor fields. Potentially, this could lead to new tools for the analysis of brain connectivity and diagnosis of connectivity-related disorders.

A prime motivation for our work is brain imaging, where diffusion tensor magnetic resonance imaging (DTI) enables the in-vivo exploration of nerve fibre bundles. This allows the determination and visualization of anatomical connections between brain regions. Whereas the theoretical and algorithmic foundations for morphological operators on scalar images are well developed, this is not the case for tensor images.

Our goals are the development of new mathematical methods and efficient algorithms for analysing and understanding tensor data, and evaluating these methods in selected application domains.

Our work is founded on the complete lattice framework of mathematical morphology. In this methodology, shapes of objects in binary (black and white) images are detected by moving small test shapes – structuring elements – of various forms and sizes over the image and recording the pixel locations where certain logical relations between the image and the structuring element are satisfied. In this way, it is possible to define a wealth of morphological operations that can extract various shapes and patterns from the image. Morphological image filters are nonlinear, which makes their mathematical analysis more complex.

This approach can be extended to grey scale images. A unified framework is provided by complete lattice theory, where the existence of a partial order on the space of images is essential. For the binary case this partial order is given by set inclusion, while in the grey scale case the (complete) order on the image values can be extended to a partial order on the image functions themselves. Also, it is very important to pay attention to invariance properties of the resulting operators, such as translation, rotation, or scale invariance.

Extending morphological filters to tensor data is difficult, since there is no obvious way to define a partial order on tensors. We started with the simpler case of colour images, and are now extending the results to tensor images.

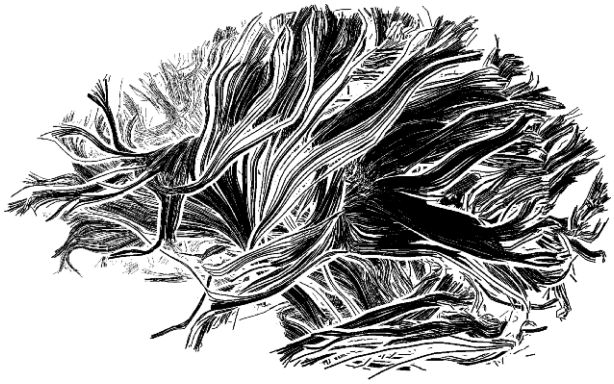


Figure 1: Illustrative visualization of DTI fiber tracts. Source: Scientific Visualization and Computer Graphics group, University of Groningen (Ref. [3]).

By using the frame (overcomplete basis) concept it is possible to construct operators that are invariant to a given group of transformations. For example, we obtained saturation- and rotation-invariant frames for colour images, leading to more intuitive and better quality results.

Of particular interest is the study of connectivity, which allows us to group pixels of the same grey value into connected components. Connectivity-preserving ("connected", for short) filters either remove or keep each connected component, but do not change its shape. However, the human observer sometimes interprets a single connected component as multiple visual entities, or vice versa. To deal with such cases hyperconnectivity and associated hyperconnected filters have been proposed. The axiomatization of hyperconnected filters is a very recent development, and extending this to tensor data will allow us to explore the full potential of such filters.

The processing of tensor images is computationally very demanding. Hence we use approximation techniques that achieve a compromise between accuracy and speed, and also foresee the use of parallel processing and special hardware such as graphical processing units (GPUs).

Although this project has a strong theoretical orientation, we also aim to test and validate our methods in various application domains, notably neuroscience, material science, and seismology.

Future work concerns the development of morphological operators for higher order tensors, which are relevant for High Angular Resolution Diffusion Imaging (HARDI), where data for a large number of gradient directions are measured in an MRI scanner.

The project started in 2010 and will run until 2014. Its research team consists of PhD student Jasper van de Gronde, Dr. Michael Wilkinson, and Prof. Dr. Jos Roerdink (PI) of the Johann Bernoulli Institute for Mathematics and Computer Science of the University of Groningen. In this project, we are collaborating with Dr. J. Angulo, Center for Mathematical Morphology (CMM), Fontainebleau, France; and with Dr. R. Duits, Department of Mathematics and Computer Science, Eindhoven University of Technology, the Netherlands.

Link: http://www.cs.rug.nl/svcg/SciVis/COMOTI

References:

[1] J.J. van de Gronde, J.B.T.M. Roerdink: "Group-invariant frames for colour morphology", in proc. ISMM 2013, LNCS 7883, 267-278
[2] J.B.T.M. Roerdink: "Group morphology", Pattern Recog. 33(6), 877-895, 2000
[3] M.H. Everts et al.: "Depth-Dependent Halos: Illustrative Rendering of Dense Line Data", IEEE TVCG 15(6), 1299–1306, 2009.

Please contact:

Jos Roerdink
Scientific Visualization and Computer Graphics, Johann Bernoulli Institute for Mathematics and Computer Science, University of Groningen
Tel. +31 50 363 3931
E-mail: j.b.t.m.roerdink@rug.nl

Person Re-identification

by Slawomir Bak and François Bremond

A retrieval tool that helps a human operator browse a network of cameras is being developed at Inria Sophia Antipolis. This tool addresses the problem of person re-identification: determining whether a particular individual has already appeared over a network of cameras.

Person re-identification (also known as multi-camera tracking) is defined as the process of determining whether a given individual has already appeared over a network of cameras (see Figure 1). This task can be considered on different levels depending on information cues which are currently available in video analytics systems. For instance, biometrics such as face, iris or gait can be used to identify people. However, in most video surveillance scenarios such detailed information is not available due to video low-resolution or difficult segmentation (crowded environments, such as airports and metro stations). Therefore a robust modelling of the global appearance of an individual (clothing) is necessary for re-identification. This is a particularly challenging problem owing to significant appearance changes caused by variations in view angle, lighting conditions and the individual's different positions.

Figure 1: Person re-identification task: the system should be able to match appearances of a person of interest extracted from non-overlapping cameras.

Owing to this complexity, current state of the art approaches have relatively low retrieval accuracy, thus a fully automated system is still unattainable. However, we propose a retrieval tool that helps a human operator to solve the re-identification task.

Our proposed tool (Figure 2) allows a human operator to browse images of people extracted from a network of cameras: to detect a person on one camera and to re-detect the same person few minutes later on another camera. The main stream is displayed on the left of the screen, while retrieval results are shown on the right. The results show lists of the most similar signatures extracted from each camera (green boxes indicate the correctly retrieved person). Below the main stream window a topology of the camera network is shown. Detection and single camera tracking (see the main stream) are fully automatic. The human operator only needs to select a person of interest, thus producing retrieval results (right screen). The operator can easily see a preview of the retrieval results and can go directly to the original video content.

Our tool is based on novel techniques, which outperform the best state of the

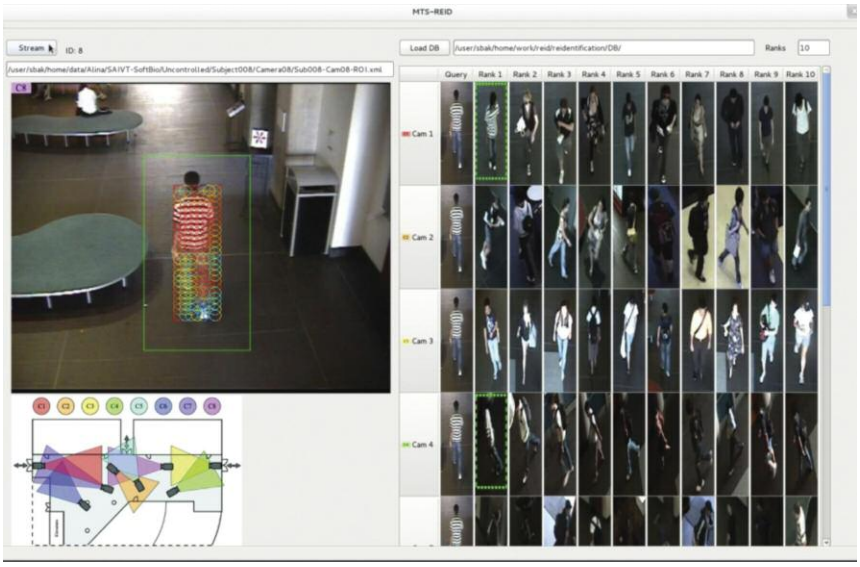


Figure 2: Re-identification tool

art solutions on the market. These techniques allow the identification of a visual signature (comprising a set of covariance matrices of reliable descriptors) characteristic of one individual while being camera invariant and the most discriminative compared to the other observed people [1][2].

We are working not only on re-identification libraries, which provide high recognition accuracy algorithms, but also on improving detection and single camera tracking algorithms. In addition to re-identification technology, we also focus on designing an intuitive graphical interface, an important feature for the human operator analysing retrieval results. Displaying retrieval results from a large camera network is still an issue, even after applying time-space constraints (the usage of topology of cameras).

Our framework, developed at Inria Sophia Antipolis, originated as the practical application of Bak's thesis [3] and has been developed under several European Projects (VIDEOID, PANORAMA, CENTAUR). Future activities will focus on both development of an intuitive graphical interface and improvement of recognition accuracy algorithms. It is the combination of these two aspects that will provide the solution for re-identification within a large camera network.

Link:
http://www-sop.inria.fr/members/Slawomir.Bak/research.html

References:
[1] S. Bak, et al.: "Multiple-shot human re-identification by mean riemannian covariance grid", AVSS 2011
[2] S. Bak, G. et al.: "Learning to Match Appearances by Correlations in a Covariance Metric Space", ECCV 2012
[3] S. Bak: "Human Re-identification Through a Video Camera Network", PhD Thesis, 2012

Please contact:
Slawomir Bak, Inria
Tel: +33 49715 5312
E-mail: slawomir.bak@inria.fr

Exploiting Computational Models of the Human Visual System

by Franco Alberto Cardillo, Giuseppe Amato and Richard Connor

Biological models of the human visual system can be exploited to improve the current state of the art in Content-Based Image Retrieval (CBIR) systems.

We focus on novel ways to include local image descriptors, such as SIFT (scale invariant feature transform) [1] and SURF (speeded up robust features), in Content-Based Image Retrieval (CBIR) systems. Since there can be thousands of SIFT keypoints in a single image, their number needs to be reduced in order to use them in a CBIR context. We use a computational model of the human visual attention system to compute image saliency and restrict the comparisons among images to their most salient regions. The computational model of bottom-up attention is an extension of a well-known model by Itti et al. [2].

Given an image as a query, a CBIR system returns a set of images ranked according to their visual similarity to the query image. Classical approaches to CBIR use global image features, ie features whose values depend on the overall appearance of the image. However, global features yield poor results on semantic queries where the user is not interested in the overall image but in certain sub-images containing particular objects.

The basic assumption behind our work is that a user selects a query image mainly by looking at its most salient areas, and thus only the most salient image areas should be compared by the similarity function. In order to compute the image saliency we implemented a biologically-inspired and biologically-plausible computational model of the human visual attention system. The model encodes the input image according to what is known about the human retina and the first stages of the higher visual system.

When we open our eyes we see a colourful and meaningful three-dimensional world. This visual experience starts in our retina where the light

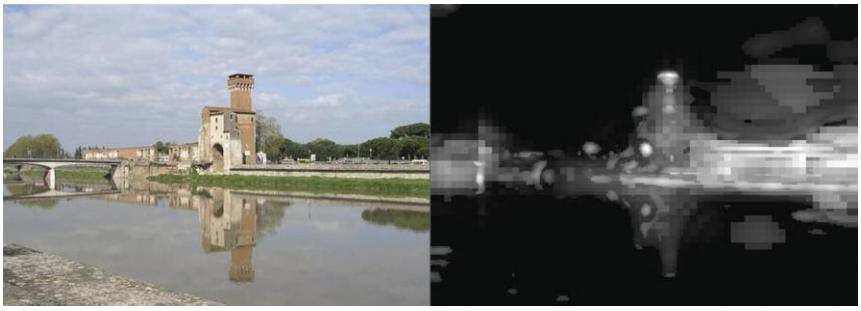


Figure 1: Left side: original image in the Pisa-Dataset. Right side: global saliency map computed by the visual attention model.

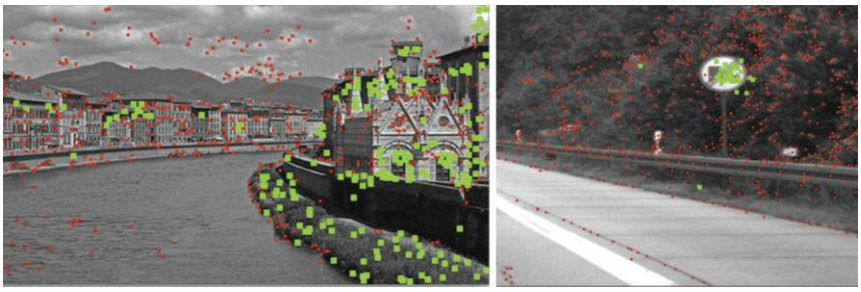


Figure 2: The image shows the SIFT keypoints retained by the visual attention model. The red points belong to non-salient regions and are thus not used in the matching.

strikes our photoreceptors, which are connected to ganglion cells through an intermediate layer of bipolar cells. Starting from the photoreceptors, the full visual experience is built up: features, such as intensity and colours, are extracted; lines with various orientations are detected; and the information flow is routed to the correct higher level neural circuits. The basic encoding performed at these early visual stages has a centre-surround organization: cells give a strong output when the centre of their receptive field contains their preferred stimulus, while the surrounding region does not (or vice-versa). Furthermore, the size of the receptive fields increases when moving from the centre to the periphery of the retina.

In our model, the input image is encoded in the Lab colour space, with the three channels L, a, and b encoding, respectively, the lightness, the red-green, and the blue-yellow opponent channels. The Lab values are then split into five different channels: intensity, red, green, blue, and yellow. These channels are used to build centre-surround (eg, red-on/green-off) feature maps, as found in the primate retina. Oriented features are extracted using Gabor filters with parameters set to values respecting the biological responses in the V1 cortical area. All the channels are encoded using Gaussian pyramids to mimic a space-variant retinal organization and introduce scale invariance into the model. The feature maps in the various dimensions (colour, intensity, oriented features) are merged into feature conspicuity maps encoding the contribution that the individual feature maps provide to the overall saliency. The feature conspicuity maps are then integrated into local saliency maps (one map per level of the Gaussian pyramid), which are finally merged into a global saliency map (see example in Figure 1).

We experimented with two different tasks using two well-known and publicly available datasets: the Pisa-Dataset for landmark recognition and the STIM dataset for object recognition. The two sets were split into a training set and a test set. We extracted both SIFT keypoints and the saliency maps from each image in the two sets. Image matching was performed by comparing the keypoint descriptors in two images and searching for matching pairs. The candidate pairs for the matches are verified to be consistent with a geometric transformation using the RANSAC algorithm [3]. The percentage of verified matches is used to determine whether or not the two images contain the same rigid object. For each test image, the best candidate among the images in the training set is selected and its label is assigned to the test image.

We compared the results with and without the saliency filtering. When using the saliency map to filter out the keypoints, we applied several thresholds in order to keep areas at various degrees of saliency. The results are very promising. On the two datasets, the accuracy improves or just slightly decreases while the processing time is drastically reduced. For example, on the Pisa-Dataset a single test image is processed in 7.2 seconds on average, and in under one second when either all the keypoints or only 30% of those belonging to image areas with high saliency are kept.

In the future, we plan to experiment with larger datasets. For this purpose, we have currently implemented a parallel version of the visual attention model and extracted the saliency maps from the MIRFLICKR dataset containing one million images (using the HECToR supercomputer at the University of Edinburgh - HPC-EUROPA2 project.

This research is conducted as part of a joint collaboration involving research groups at the Institute of Information Science and Technologies of the Italian National Research Council, and the Department of Computer and Information Sciences of the University of Strathclyde, UK.

References:

[1] Lowe, D.G., "Distinctive image features from scale-invariant keypoints", IJCV 3(1), 2004.
[2] Itti, L. et al., "A model of saliency-based visual attention for rapid scene analysis", PAMI 20(11), 1998.
[3] Fischler, M.A., Bolles, R.C., "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography", Commun. ACM, 24(6), 1981.

Please contact:

Franco Alberto Cardillo
ISTI-CNR, Italy
E-mail:
franco.alberto.cardillo@isti.cnr.it

Large Scale Image Retrieval Using Vectors of Locally Aggregated Descriptors

by Giuseppe Amato, Paolo Bolettieri, Fabrizio Falchi and Claudio Gennaro

We propose using vectors of locally aggregated descriptors (VLAD) to address the problem of image search on a very large scale. We expect that this technique will overcome the quantization error problem faced in Bag-of-Words (BoW) representations.

Conventional search engines use inverted index file indexing to speed up the solution of user queries. We are studying a methodology which will enable inverted files of standard text search engines to index vectors of locally aggregrated descriptors (VLAD) to deal with large-scale image search scenarios. To this end, we first encode VLAD features by means of the perspective-based space transformation developed in [1]. The idea underlying this technique is that when two descriptors are very similar, with respect to a given similarity function, they "see" the "world around" them in the same way. In a next step, the "world around" can be encoded as a surrogate text representation (STR), which can be managed with an inverted index using a standard text-based search. The conversion of visual descriptors into a textual form allows us to employ off-the-shelf indexing and searching functions with little implementation effort.

Our transformation process is shown in Figure 1: the blue points represent reference VLAD features; the other colours represent dataset VLAD features. The figure also shows the encoding of the data features in the transformed space and their representation in textual form (SRT). As can be seen intuitively, strings corresponding to VLAD features X and Y are more similar to those corresponding to X and Z. Therefore, the distance between strings can be interpreted as an approximation of the original VLAD distance *d*. Without going into the math, we leverage on the fact that a text-based search engine will generate a vector representation of STRs, containing the number of occurrences of words in texts. With simple mathematical manipulations, it is easy to see how applying the cosine similarity on the query vector and a vector in the database corresponding to the string representations will give us a degree of similarity that reflects the similarity order of reference descriptors around descriptors in the original space. Mathematical details of the technique are outlined in [2].

The idea described so far uses a textual representation of the descriptors and a matching measure based on a similarity offered by standard text search engines to order the descriptors in the dataset in decreasing similarity with respect to the query. The result set will increase in precision if we order it using the original distance function *d* used for comparing features. Suppose we are searching for the *k* most similar (nearest

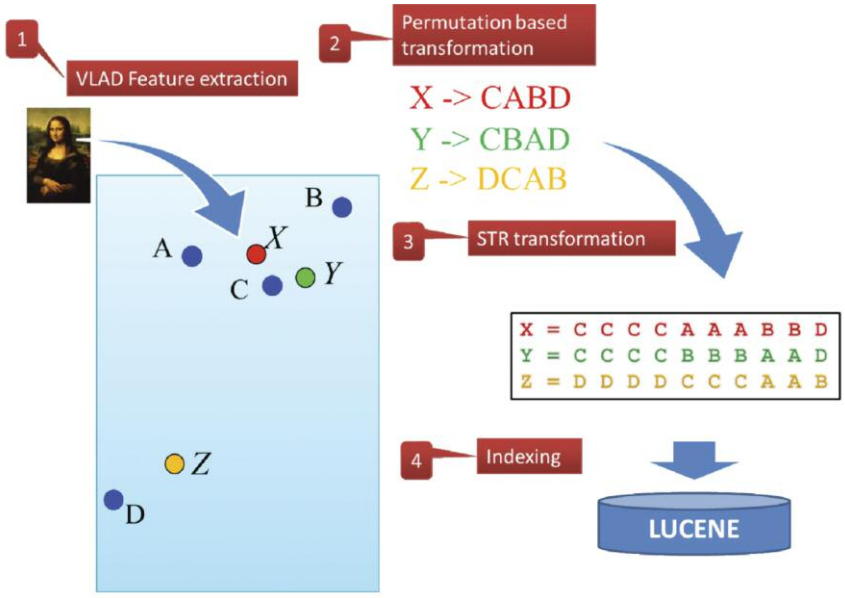


Figure 1: Example of perspective-based space transformation and surrogate text representation: 1) From the images we extract the VLAD features represented by points in a metric space. Blue points are reference features and colored points are data features, 2) The points are transformed into permutations of the references, 3) The permutations are transformed into text documents, 4) The text documents associated with the images are indexed.

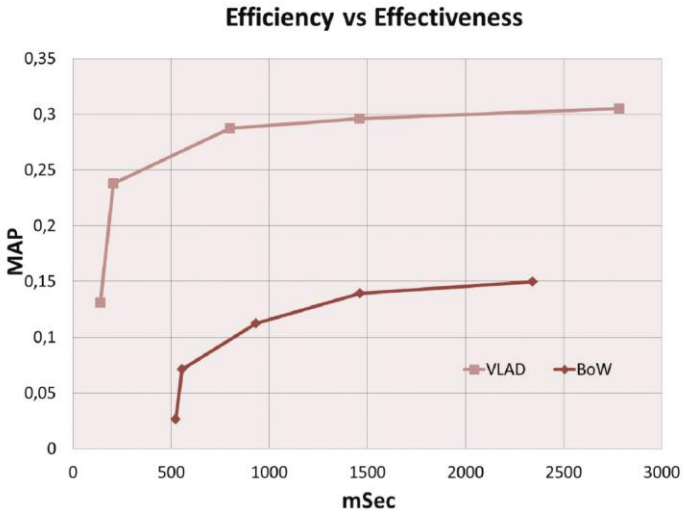


Figure 2: Effectiveness (MAP) with respect to efficiency (mSec per query) obtained by VLAD and BoW for various settings.

neighbours) descriptors to the query. We can improve the quality of the approximation by re-ranking, using the original distance function *d* and the first $c \cdot (c \geq k)$ descriptors from the approximate result set at the cost of more *c* distance computations. This technique significantly improves accuracy at a very low search cost.

We applied the STR technique to the VLAD method comparing both effectiveness and efficiency with the state-of-the-art BoW approach on the same hardware and software infrastructure using the publicly available and widely adopted 1M photos dataset. Given that the STR combination gives approximate results with respect to a complete sequential scan, we also compare the effectiveness of VLAD-STR with standard VLAD. Moreover, we considered balancing efficiency and effectiveness with both BoW and VLAD-STR approaches. For the VLAD-STR, a similar trade-off is obtained varying the number of results used for re-ordering. Thus, we do not only compare VLAD-STR and BoW on specific settings but we show efficiency vs effectiveness graphs for both. For the VLAD-STR, a trade-off is obtained varying the number of results used for re-ordering.

We show that the use of inverted files with VLAD significantly outperforms BoW in terms of efficiency and effectiveness on the same hardware and software infrastructure. In Figure 2, we plot mean average precision (MAP) with respect to the average query execution time for both BoW and VLAD. The graph underlines both the efficiency and effectiveness advantages of the VLAD technique with respect to the BoW approach. The efficiency vs effectiveness graph reveals that VLAD-STR obtains the same MAP values as BoW, for an order of magnitude less in response time. Moreover, for the same response time, VLAD-STR is able to obtain twice the MAP of BoW.

References:

[1] G. Amato, P. Savino: "Approximate similarity search in metric spaces using inverted files", in proc. of InfoScale '08, pages 28:1-28:10, ICST, Brussels, Belgium

[2] C. Gennaro, G. Amato, P. Bolettieri, P. Savino: "An approach to content-based image retrieval based on the Lucene search engine library", in proc. of ECDL 2010, Springer LNCS.

Please contact:
Giuseppe Amato
ISTI-CNR, Italy
Tel: +39 050 3152810
E-mail: giuseppe.amato@isti.cnr.it

Graph Based Keyword Spotting in Handwritten Historical Slavic Documents

by Kaspar Riesen and Darko Brodic

Many libraries globally have started digitizing their most valuable old handwritings in order to preserve the world's cultural heritage. To improve the accessibility of the large number of available handwritten document images they must be made amenable to searching and browsing. A recent research project aims at a novel graph based keyword spotting framework applicable to historical documents. For testing the novel framework, isolated word images from the Miroslav Gospels (one of the oldest surviving documents written in old church Slavonic) will be represented by graphs.

"Keyword spotting" refers to the process of retrieving all instances of a given keyword or key phrase from a document. A large variety of algorithms have been developed for keyword spotting over the last two decades. For instance, a common approach is to represent a word as a sequence of features, extracted via a sliding window. Subsequently, the sequences are compared using dynamic time warping (DTW) [1]. The major objective of the present project is to employ graph based pattern representation for the problem of keyword spotting. To the knowledge of the authors, graph based algorithms have never been used for this task. This is rather surprising as it turns out that the paradigm of graph matching, ie the comparison of pairs of graphs, perfectly meets the requirements of keyword spotting.

Various procedures for graph matching have been proposed in the literature. In [2] a novel graph matching procedure was proposed. This method is based on an (optimal) fast optimization procedure mapping nodes and their local structure of one graph to nodes and their local structure of another graph. This procedure allows the approximation of graph dissimilarities in polynomial time complexity. The distances found by this framework are equal to, or larger than, the exact graph distance. Empirical investigations show that relatively large distance values tend to be overestimated while small distances remain nearly unaffected by our approximation. Keyword spotting is a typical application where an overestimation of large distances is acceptable since we are only interested in document words that are very near to the keyword.

We plan to extract graphs from skeleton word images by means of a procedure similar to [3]. First, keypoints that include endpoints, intersections, and corner points of circular structures are extracted from the word skeleton image. Next, a node is added to the word graph for each skeleton keypoint, labeled with its position. After all keypoints have been included as nodes in the word graph, connection points are also added as nodes to the graph. Given a skeleton line connection between two keypoints in form of a pixel chain, connection points are inserted at regular distances along that chain. Finally, skeleton connections between two nodes are represented by undirected and unlabeled edges.

Through the representation of isolated words with graphs, the search and

retrieval of given words in documents can be interpreted as matching of an input graph (keyword) with a large set of graphs (document). More formally, in order to spot a certain keyword w_i, all p graph instances g_{i1}, … , g_{ip} of that word w_i occurring in the training set are matched against all graph words in each text line using our adapted graph matching procedure. That is, for a given word w_i and a specific text line s pairwise distances between all prototypical graphs g_{i1}, …,g_{ip} and the m word graphs $g_{'1}$, … ,$g_{'m}$ from text line s are obtained first. The minimum of these graph distances serves as a distance function $d(w_i,s)$ of the keyword's word class w_i to the text line s. If the distance $d(w_i,s)$ of a keyword to the text line is below a given threshold, the text line s and the word from s having the minimum distance is returned as a positive match to the keyword w_i.

A very important aspect of the whole project is the experimental evaluation of our novel graph based procedure for keyword spotting. We plan to carry out exhaustive experimental evaluations on the Miroslav Gospels. Miroslav Gospels is a 362-page illuminated manuscript Gospel Book on parchment with very rich decorations. It is one of the oldest surviving documents written in Old Church Slavonic and represents one of the most precious and significant documents in cultural heritage of Serbia (see Figure 1 for a detailed view of a page of the Miroslav Gospel).

Figure 1: The Miroslav Gospel (source: http://upload.wikimedia.org/wikipedia/commons/3/36/Miroslavs_Gospel.jpg)

The outlined project has been submitted as a proposal for joint research project to the SCOPES programme (Scientific co-operation between Eastern Europe and Switzerland). The research project will be mainly carried out by the Technical Faculty in Bor at the University of Belgrade (Serbia) and the Institute for Information Systems at the University of Applied Sciences and Arts Northwestern Switzerland. Further information about the SCOPES programme can be found at http://www.snf.ch/E/international/europe/scopes/Pages/default.aspx

Links:

Video lecture on our novel algorithmic framework for approximate graph distances:
http://videolectures.net/gbr07_riesen_bgm
The algorithmic framework for approximate graph distances:
http://www.fhnw.ch/wirtschaft/iwi/gmt
Miroslav Gospels:
http://en.wikipedia.org/wiki/Miroslav_Gospel
SCOPES programme:
http://www.snf.ch/E/international/europe/scopes

References:

[1] T. M. Rath, R. Manmatha: "Word spotting for historical documents", Int. Journal on Document Analysis and Recognition, p. 139-152, 2007
[2] K. Riesen, H. Bunke: "Approximate graph edit distance computation by means of bipartite graph matching", Image and Vision Computing, 27(4):950-959, 2009.
[3] A. Fischer, K. Riesen, H. Bunke: "Graph similarity features for HMM-based handwriting recognition in historical documents", in proc. Int. Conf. on Frontiers in Handwriting Recognition, pages 253–258, 2010.

Please contact:

Kaspar Riesen
University of Applied Science and Arts Northwestern Switzerland, Olten, Switzerland
Tel: +41 79 688 77 19
E-mail: kaspar.riesen@fhnw.ch

Highly Degraded Recto-verso Document Image Processing and Understanding

by Emanuele Salerno and Anna Tonazzini

The ITACA project (Innovative tools for cultural heritage archiving and restoration) is investigating new approaches to treat severe back-to-front interference in digital images of two-sided documents. This work is part of a vast research program on the study and preservation of historical documents, which, since 2004, has been supported in various forms by European funds.

Distinguishing the foreground pattern (eg text or graphics) from interference is a basic step in document understanding. When dealing with back-to-front interference, this is not always an easy task, as the front and rear patterns are mixed nonlinearly and the interference strength varies from point to point. When imaging a physical document, the transparency of the paper combined with the features of the capture device cause the content of each side to appear in the opposite side's image as "show-through interference". A similar effect (bleed-through) is caused by the ink bleeding from one side to the other.

These effects are intrinsically nonlinear. Moreover, the non-uniformity of the paper support, the varying strength of text strokes, and possible uncompensated variations in illumination often make these processes non-stationary. Based on a slight generalization of a model proposed in the literature for moderate show-through, we presented a

nonlinear model that can account for show-through and, partly, for bleed-through [1]. We then augmented the model by introducing a space varying interference strength [2]. This new model is not yet fully non-stationary, as the convolutive effects and the typical non-linearities are considered constant throughout the image, but the experimental results compare very favourably with previous approaches, and the possibilities of discriminating foreground and interference are broadened significantly.

Figure 1 shows our data model: two coupled nonlinear equations, written in terms of the front and rear optical densities. As is known, optical density is related logarithmically to the commonly used reflectance value. The appearance of each side of the document is a nonlinear combination of the foreground pattern and a blurred and attenuated version of the pattern on the opposite side. Given the front and rear-side appearances, inverting this model means estimating the show-through point spread functions (PSF's), the show-through gains, and the front and rear-side pure patterns. Strictly speaking, a non-stationary model

$$D_r^{obs} = D_r + q_v[h_v * (1 - e^{-D_v})]$$

$$D_v^{obs} = D_v + q_r[h_r * (1 - e^{-D_r})]$$

$D_{r,v}^{obs}$ = *recto* and *verso* observed optical densities

$D_{r,v}$ = *recto* and *verso* pure density patterns

$q_{r,v}$ = *recto* and *verso* (space variant) show-through gains

$h_{r,v}$ = *recto* and *verso* (space invariant) show-through PSF's

Figure 1: Non-linear and non-stationary model for back-to-front interference. The asterisk means convolution.

should include two space-variant kernels rather than two PSF's, and the show-through gains should depend on space. As a first attempt to treat non-stationary back-to-front interference, we use fixed PSF's all over the image pair while allowing the gains to depend on space. Approximated PSF's can simply be evaluated beforehand; once this is done, a straightforward formula allows us to estimate the gains in all the pixels where interference is present. An easy and fast constrained maximum likelihood scheme is then used to estimate the pure patterns. The results obtained are very promising, even though a general strategy is still to be found to cope with the saturation effects occurring where the foreground pattern occludes the show-through.

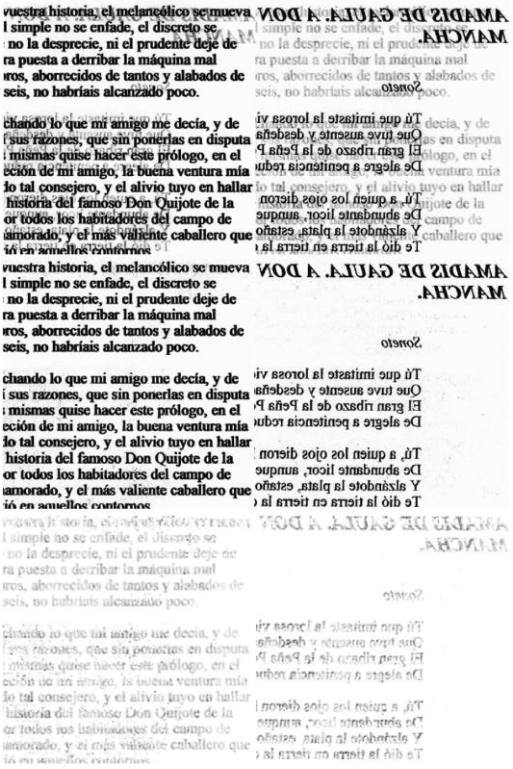


Figure 2: Example of non-stationary show-through removal. Top: observed image pair (appearance). Middle: restored image pair (estimated pure patterns). Bottom: show-through gain maps (linear grayscale, black=maximum).

Figure 2 shows a typical result obtained in an apparently non-stationary case: the removal of the back-to-front interferences is very effective, and the gain maps show us that the distortion is significantly non-stationary. A comparison between the new results and those obtained through stationary models confirms the potential advantages achievable by accounting for non-stationarity. The type of fixed non-linearity we adopt is the same as that introduced for weak show-through in the paper that inspired our work [3]. It is likely that a stronger show-through or other phenomena, such as bleed-through, would need different non-linearities. Finding a suitable model for documents showing different kinds of interference might even prove to be impossible. In such cases, introducing non-stationarity could produce an additional advantage: providing an accurate and comprehensive data model could become less important - we could even return to a linear model. This is the focus of our current research, and our first results are corroborating our conjecture.

The ITACA project (POR-FESR Calabria 2007-2013) is led by TEA Sas, Catanzaro, Italy, a firm that provides consulting and cultural heritage digitization, processing, and management. The activity described here is being conducted at ISTI-CNR, Pisa. The research teams at ISTI and TEA are both members of ERCIM's MUSCLE working group, on Multimedia Understanding through Semantics, Computation, and Learning.

Links:

http://www.isti.cnr.it/research/unit.php?unit=SI

http://www.teacz.com

References:

[1] E. Salerno et al: "Nonlinear model identification and see-through cancellation from recto-verso documents", Int. J. Docum. Anal. Recogn., 2013, DOI 10.1007/s10032-012-0183-y

[2] A. Tonazzini, et al: "Removal of non-stationary see-through interferences from recto-verso documents", in Machine Learning and Data Mining in Pattern Recognition, 2013, ISBN 978-3-942952-22-4, pp. 151-158),

[3] G Sharma: "Show-through cancellation in scans of duplex printed documents", IEEE Trans. Im. Proc., 2001, DOI 10.1109/83.918567)

Please contact:

Emanuele Salerno
ISTI-CNR, Italy
E-mail: emanuele.salerno@isti.cnr.it

Research activities and innovative developments in European research institutes

SECCRIT: Secure Cloud Computing for High Assurance Services

by Roland Bless, David Hutchison, Marcus Schöller, Paul Smith and Markus Tauber

Owing to concerns about privacy, security and resilience, critical infrastructure service providers are not yet making wide-scale use of cloud computing for their high-assurance ICT services. The EU-funded SECCRIT project aims to address this problem so the benefits of the cloud can be leveraged in this sector in a safe and secure manner.

Cloud computing is one of the most important and successful recent trends in consumer market service provisioning. According to the definition from NIST [1], it is an operational service model, wherein services are deployed on remote data centre infrastructures, which make use of virtualization technology. The use of virtualization enables rapid and flexible (elastic) service deployment. Attracted by the benefits of cloud computing, critical infrastructure providers, such as utility companies and government bodies, are considering deploying their high assurance ICT services in the cloud. However, privacy, security and resilience requirements are more stringent in this sector than in the general consumer market, thus hampering the wide scale use of cloud computing. In order to address these challenges, and enable the safe and secure usage of cloud computing for critical infrastructure ICT services, the EU-funded SECCRIT project has a number of objectives:

Propose critical infrastructure cloud legal frameworks and guidelines

The project will identify the relevant European legal framework and establish respective guidelines for the use of cloud services in the critical infrastructure sector. Without establishing such a legal framework, the use of cloud in this sector, in which there are often stringent regulatory and legal requirements, will continue to be severely limited. Furthermore, it is important to have clear guidelines on how to address liability issues in light of service failures.

Investigate and develop technologies for critical infrastructure cloud

Building on this legal framework, the project will investigate solutions for fault identification and localization using digital forensic methods [2] in cloud infrastructures. Attacks and failures are inevitable; therefore, it is important to develop approaches to understanding cloud behaviour in the face of challenges and attacks. In this area, we will investigate where are the appropriate points in the cloud to place monitoring and attack detection functionality. In order to contextualize the research in SECCRIT, we have derived an architectural model, shown in Figure 1, which represents our view that cloud computing needs to account for resilience, security and privacy concerns. A key task is to further develop this cloud architectural model to make it suitable for the critical infrastructure context, as current models are not well-

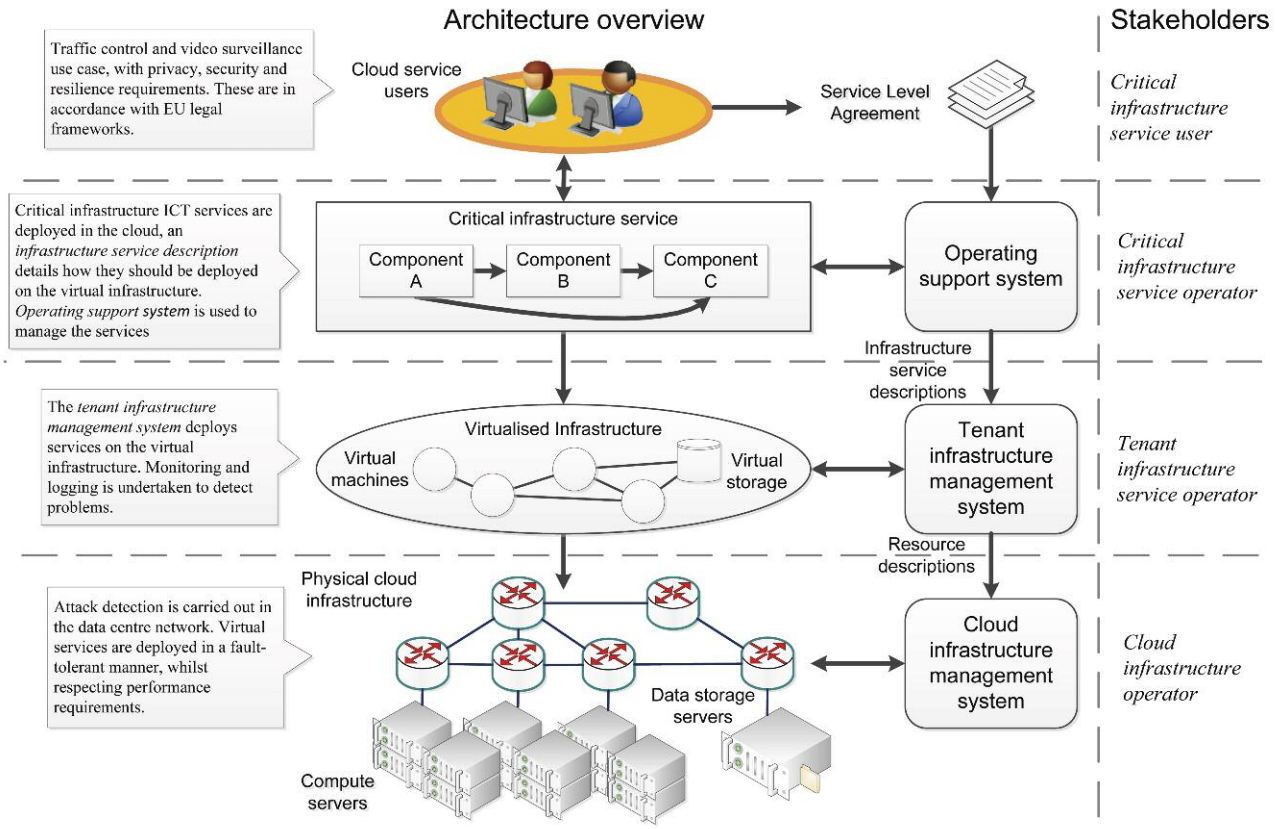


Figure 1: Overview of the SECCRIT architectural model

suited for use in this domain. For example, we have identified the need for a finer distinction between roles to reflect separation of administrative responsibilities and interfaces.

Understand risks and provide best-practice guidance

A key challenge faced by potential critical infrastructure cloud users is to understand the new cyber-security risks associated with this technology. Despite some work in this area, eg, on understanding cloud-specific vulnerabilities [3], there is a lack of suitable techniques and processes for understanding and managing risk associated with cloud environments, which involves numerous stakeholders. This project aims to address this shortcoming. Building on all the aforementioned activities, the project will establish a set of best-practice guidelines for secure cloud service implementations, which can be used by the various stakeholders in this area to ensure secure and resilient cloud services. For instance, the guidelines can be used to determine the appropriate cloud deployment model for a service, eg, public, private or hybrid community cloud.

Real-world evaluation and strong stakeholder engagement

In order to validate the project's outcomes, two demonstration deployments will be undertaken: (i) using the cloud to support a traffic management system in the city of Valencia; and (ii) implementing a video surveillance system that monitors critical infrastructures with the support of cloud-based services. Furthermore, the SECCRIT project has a growing user and advisory board, which provides requirements from various critical infrastructure domains and evaluates the project's outcomes. Members of the board receive privileged access to project results.

About the project

The SECCRIT project started in January 2013 and will run for three years. It is funded by the European Union under grant number 312758. The consortium, coordinated by AIT, Austrian Institute of Technology (AT), is drawn from across Europe, and includes Amaris (AT), ETRA I+D (ES), Ajuntamente De Valencia (ES), Fraunhofer IESE (DE), KIT, Karlsruhe Institute of Technology (DE), NEC Laboratories Europe (UK), Lancaster University (UK), Mirasys Ltd. (FI), and the Hellenic Telecommunications Organization S.A. (GR). News, details about how to join the project's advisory board, and our deliverables can be found on the project web site.

Link: https://www.seccrit.eu

References:

[1] P. Mell, T. Grance: "The NIST Definition of Cloud Computing, NIST Special Publication 800-145," September, 2011.
[2] T. V. Lillard: "Digital Forensics for Network, Internet, and Cloud Computing: A Forensic Evidence Guide for Moving Targets and Data," Syngress Publishing, 2010.
[3] B. Grobauer, T. Walloschek, E. Stocker: "Understanding cloud computing vulnerabilities," IEEE Security & Privacy, vol. 9, no. 2, pp. 50–57, March-April, 2011.

Please contact:

Markus Tauber
AIT Austrian Institute of Technology / AARIT, Austria
Tel: +43 664 8251011
E-mail: seccrit@ait.ac.at

Interdependencies of Genetic and Epigenetic Events in a Computational Model for Colon Cancer Dynamics

by Irina-Afrodita Roznovăț and Heather J. Ruskin

The aim of our current work is to investigate the interdependencies of genetic and epigenetic mechanisms leading to aberrations in cancer initiation and progression. The objectives are to develop a computational model for colon cancer dynamics, linking microscopic effects to macroscopic outcomes, and to analyse the impact of different risk factors on malignant tumour development.

Conducted at the Centre for Scientific Computing & Complex Systems Modelling, (SCI-SYM), School of Computing, Dublin City University, this research is being funded by the CIESCI ERA-Net Complexity Project, (EC/Irish Research Council and as an extended Embark project (IRC)). The project "Complexity of Interdependent Epigenetic Signals in Cancer Initiation" is a collaboration between three countries and laboratories: SCI-SYM, Dublin City University, Dublin, Ireland; the Bellvitge Institute for Biomedical Research, (IDIBELL), Barcelona, Spain, and the Bioinformatics, Algorithmics, and Data Mining Research Group (BIIT), Institute of Computer Science and Estonian Biocenter, University of Tartu, Estonia.

Defined as micro-molecular interactions that influence gene expression without altering the DNA sequence, epigenetic events have been detected (i) in the earliest stages of neoplastic disease initiation, (ii) in the ageing process and also (iii) in response to cellular stress. Understanding these phenomena, which are considered markers for tumour initiation, may increase the success of cancer therapy in affected patients. Given time and cost implications for laboratory experimentation and human genome studies, a range of computational models has been developed over recent decades, to help scientists and clinicians to better understand the impact of abnormal micromolecular modifications, observed to occur in neoplastic diseases. Different computational methods such as artificial neural networks, support vector machines and hidden Markov models, inter alia, have been applied to analyse and predict the patterns of genetic and epigenetic signals in malignant systems. Additionally, databases that focus on various aspects of tumour pathways are increasingly being developed and populated with specific data. Moreover, these sources capture information on epigenetic "signatures", such as DNA methylation, histone modification and changes in chromatin, together with genetic mutations of cancer-related genes.

The project aims to develop a multi-layered model to study colorectal cancer dynamics, incorporating a Bayesian network approach for estimation of incomplete information [1]. The model structure is given by the interdependencies between three main layers: micromolecular signals, gene relationships and cancer stage transitions. As a basis, the gene framework integrates empirical data on conditional relationships, between genetic and epigenetic events found in colon cancer development. The gene network is then allowed to evolve over time. After a number of iterations, which reflect cellular time scales, methylation level (for example) is checked for the entire gene network and provides a major input to the decision on cancer stage attained. A schematic outline of key influences and flows in the simple model is given in Figure 1.

The colon cancer model utilises an object-oriented framework (written in C++). Initially developed and tested in a single-threaded setting, preliminary results clearly indicated the need for speed-up, as well as inclusion of more complex computational flows. Parallelisation of the algorithm is thus the current focus, using the existing framework, (as outlined)

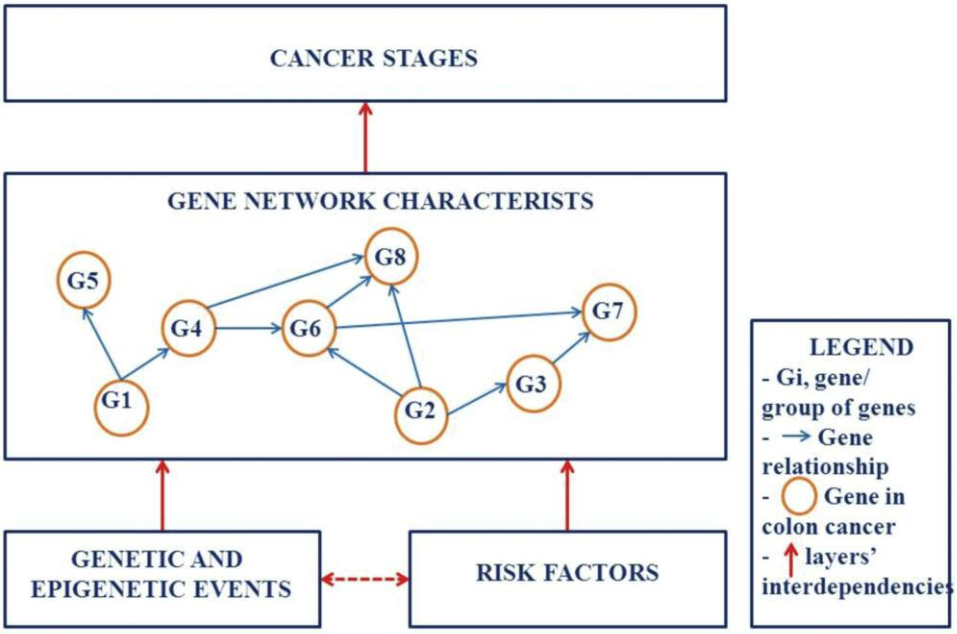


Figure 1: Interdependencies between the main layers of the Colorectal Cancer Model – Simplified illustration

and the SCI-SYM in-house computing cluster. Up-scaling to national facilities at the Irish Centre for High-End Computing, (ICHEC), is also anticipated.

Realistically, predictive model development requires integration of additional risk factors, (such as ageing, gender, heredity), and some way of quantifying the influence of these on disease progression. In consequence, the gene framework is being expanded to include and assess environmental and lifestyle impact, as well as extended mutation information. Moreover, identification of a commonly mutated/methylated gene base, offers the potential for the colorectal cancer model to be extended to other cancer types, such as stomach, lung, or liver.

SCI-SYM 's contribution to the CIESCI project complements its other work on computational modelling for molecular events in malignant diseases and related bioinformatics analyses. An agent-based model has been developed to determine the risk of gastric cancer, from aberrant DNA methylation levels induced in cells, following infection with Helicobacter pylori, (providing further insight on epigenetic events in abnormal conditions). Additionally, the StatEpigen database, [2], containing manually annotated and curated data from primary and secondary sources, has been built to provide specific information on genetic and epigenetic event determinants of gene relationships at different cancer stages.

Acknowledgements
IR is grateful for financial support from the CIESCI ERA-Net Complexity Project, (EC/IRCSET) for 2 years and final year top-up under Irish Research Council (IRC).

Links:
http://sci-sym.computing.dcu.ie/
http://computing.dcu.ie/
http://www.computing.dcu.ie/~msc/

Reference:
[1] I.A. Roznovat and H.J. Ruskin: "A Computational Model for Genetic and Epigenetic Signals in Colon Cancer,", in proc. IEEE BIBMW, Philadelphia, USA, pp. 188 – 195, 2012
[2] A. Barat and H.J. Ruskin: "A Manually Curated Novel Knowledge Management System for Genetic and Epigenetic Molecular Determinants of Colon Cancer", Open Colorectal Canc. J., 3, pp. 36 – 46, 2010.

Please contact:
Irina-Afrodita Roznovăț
Dublin City University, Ireland
Tel: +353 1 700 6747
E-mail: iroznovat@computing.dcu.ie

Heather J. Ruskin
Dublin City University, Ireland
E-mail: hruskin@computing.dcu.ie

Consensus in Computer and Communication Systems in a Stochastic Environment

by Natalia Amelina and Yuming Jiang

During the process of resource allocation in a computer system, it is vital that consensus is attained, ie: every node in the system must reach the same state with respect to certain chosen measures, such as having the same amount of steady-state remaining workload. Despite its importance, consensus is rarely considered in studies that model and evaluate resource allocation problems in computer and communication systems. Our research focuses on resource allocation from a consensus angle: we investigate the consensus properties of a distributed control strategy and apply this control strategy to various computer and communication systems.

Consensus, which has long been considered a crucial issue in distributed systems [1], requires that there are control strategies or protocols in place that can drive the states of all nodes in a system to the same steady-state values. Many computer and communication systems, such as distributed computing systems, sensor networks, and wireless mesh networks, also require distributed control and are run under a stochastic

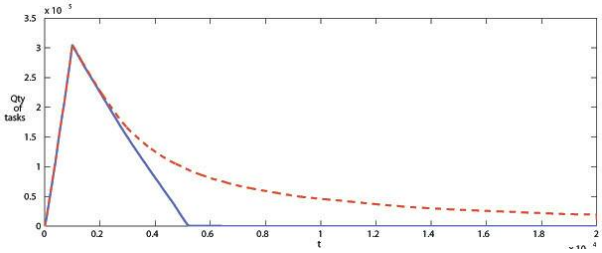


Figure 1: The number of jobs in queue v.s. time (solid line - with the adaptive strategy, dashed line – with no adaptation)

environment. To achieve optimal resource allocation in these systems, a consensus requirement is often implied. For example, a fundamental problem in distributed computing is how to distribute load to different servers. This problem is essentially one of consensus, hence consensus control strategies can represent a solution.

Specifically, in a distributed computing system, the load must be distributed among different computing nodes to shorten task completion time; in a sensor network, the battery usage must be balanced among sensor nodes to maximize the lifetime of the network; in a wireless mesh network, nodes need to be scheduled properly in using the wireless channel so that the overall network capacity is maximized. To address these needs, various resource allocation approaches have been proposed and evaluated. However, the literature has largely focused on the classic performance perspectives of the proposed approaches, such as throughput, delay, capacity, etc.,

and the consensus aspect is often overlooked, which can unfortunately sometimes lead to new issues [2].

Our research, which investigates computer and communication systems from a consensus perspective, assumes a more general and realistic stochastic environment. We show that many of the problems in the considered systems can be solved by achieving consensus among nodes. For example, in the load-balancing case, the completion time can be minimized by distributing the load in such a way that at any time, every node is given the same load weighted against its computing capability, ie when consensus in achieved in weighted load among nodes.

Specifically, we study and apply a control strategy that only requires local information and information exchange among neighbouring nodes, to various distributed computer and communication systems including distributed computing systems, sensor networks, and wireless mesh networks. We prove consensus properties of this control strategy for these systems, such as whether the strategy will lead to consensus, and how quickly consensus can be reached by what means. In addition, we investigate how well the system performs under this control strategy. A summary of early results can be found in [3]. Specifically, analytic conditions for achieving approximate consensus in a stochastic network with noise, delays and switched topology have been obtained. The results have been applied to a load balancing problem in the network. Simulation results show that the performance of an adaptive strategy, which has proven consensus properties, is significantly better than that of the strategy with no adaptation.

References:

[1] J. Turek, D. Shasha: "The Many Faces of Consensus in Distributed Systems", IEEE Computer, 25(6): 8-17, June 1992

[2] J. Cho, J-Y. Le Boudec, Y. Jiang: "On the Asymptotic Validity of the Decoupling Assumption for Analyzing 802.11 MAC Protocol", IEEE Transactions on Information Theory, 58(11): 6879-6893, 2012

[3] N. Amelina et al: "Approximate Consensus Multi-Agent Control Under Stochastic Environment with Application to Load Balancing", eprint arXiv:1306.3378, June 2013 (http://arxiv.org/abs/1306.3378).

Please contact:

Natalia Amelina (ERCIM Fellow)
Department of Telematics, NTNU, Norway
E-mail: natalia_amelina@mail.ru

Using the BonFIRE Testbed for Testing Scalability of the KOPI Service

by András Micsik, Péter Pallinger, László Kovács and András Benczúr

The KOPI Online Plagiarism Search Portal announced the detection of cross-language or translational plagiarism in 2011. This new feature requires more data intensive processing than the traditional monolingual plagiarism search (offered since 2004). As our service suffered from performance problems, we sought a platform for trying and testing automated scaling solutions applied on our service. Within the BonFIRE open call experiment we successfully implemented and analysed our elastic scaling solutions for KOPI.

BonFIRE is an EU-funded Future Internet project with the participation of several major European research clouds. BonFIRE offers a multi-site testbed with heterogeneous compute, storage and networking resources in a cloud federation for large-scale testing of applications, services and systems targeting the Internet of Services community. BonFIRE provides an API and a portal both supporting the uniform management of compute nodes, data blocks and network connections in the federated environment of seven clouds. Among the specific features of BonFIRE one can find network bandwidth control, Amazon integration, ubiquitous monitoring and other practical add-ons to regular cloud infrastructures.

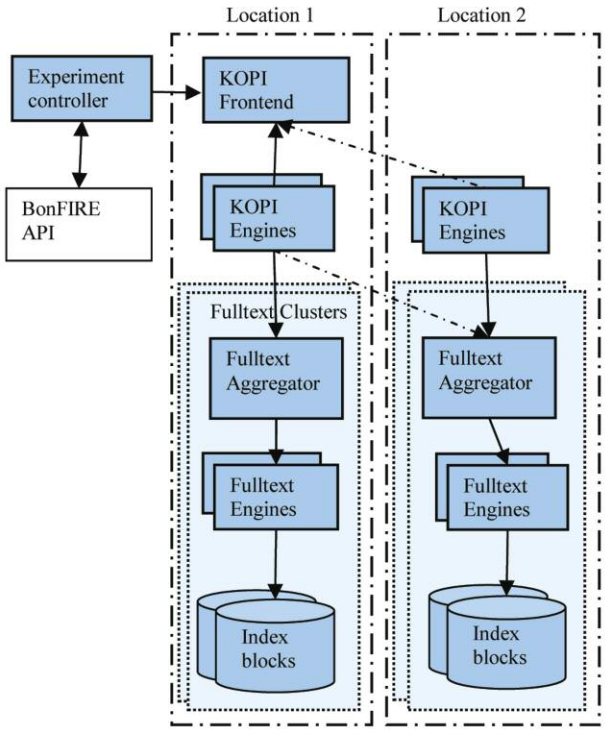


Figure 1: The KOPFire experiment setup

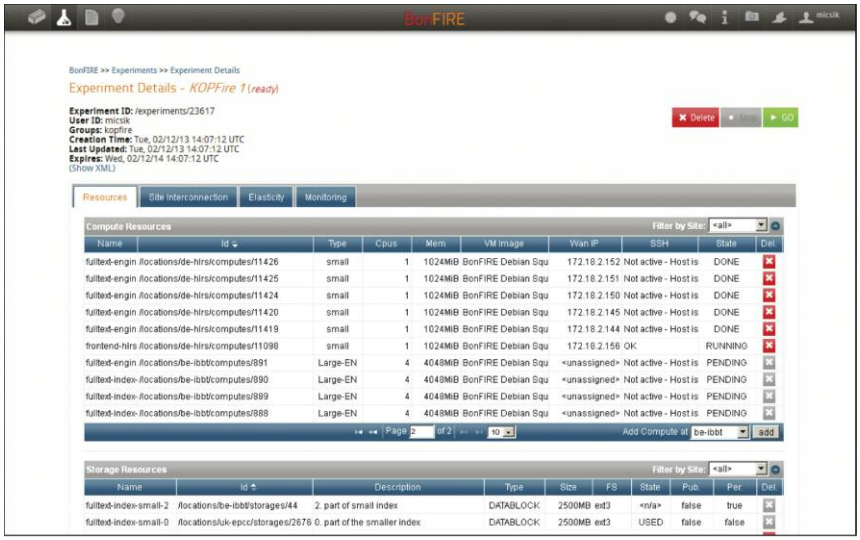


Figure 2: Managing cloud resources on the BonFIRE Portal

Figure 3: Monitoring resources in BonFIRE

The KOPI service works asynchronously: it accepts requests in the form of uploaded documents, which are checked for copied content over various databases. After this, a report is sent to the user containing the copied parts and their original sources. Processing of incoming user requests is based on a queue located on the KOPI Frontend (Figure 1), from which KOPI Engines take out requests and put back results after processing. KOPI Engines submit hundreds of search requests per document to a Fulltext Search Cluster. It typically takes 30-50 minutes to process a document.

Within the experiment called KOPFire we created a realistic test version of the KOPI service consisting of four different virtual machine (VM) types and suitable test data including test index, test requests and usage patterns. We partitioned the fulltext index and developed an aggregator of the search results. This search cluster typically used 6-11 VMs and 11-21 cores.

Our main measurement considers the effect of increasing the throughput of different service components. We collected possible atomic scaling actions together with the time needed to perform these actions. As a result, we could determine the optimal ratio of system components and the most suitable configuration options.

We measured performance by document characters processed per second (cps). Over a longer time period, the average cps can characterize the processing speed of the service. Therefore, the size of the document queue and the current processing speed can be used for triggering the scaling actions.

We created an automated scaling solution that can speed up or slow down document processing depending on the number and size of documents waiting in the queue. When we run out of capacity in one cloud, we can expand to other clouds in the federation provided by BonFIRE.

The scaling solution is based on a set of Ruby scripts using the BonFIRE API to manage cloud resources and BonFIRE Monitoring to collect measurements about virtual machines. Monitored data include both built-in metrics and our additional metrics. The scaling script detects situations when scaling is beneficial, selects and executes appropriate scaling actions and ensures that the new components are properly configured to work in cooperation with others.

Various algorithms can be plugged into the scaling script. We implemented and compared several algorithms using greedy, lazy or speed-oriented adaptation. Samples taken from the usage statistics of the real service were used to test and tune the scaling algorithms. Although we cannot say there is a single best algorithm for all usage patterns, most algorithms are good enough to raise the throughput of the service to an acceptable level and we could scale our performance in the 1:25 magnitude region during the experiment.

Furthermore, the scaling script was enhanced to continuously check the state of service components and to replace failed components. This means that as a side effect of elastic scaling we get improved fault-tolerance for the service.
These experiments helped us immensely to find the appropriate scaling solution, which at the end enables us to provide a faster and more reliable service to our growing user community.

Links:
Project home page: http://www.bonfire-project.eu/
KOPI home page: http://kopi.sztaki.hu/?lang=eng

Please contact:
András Micsik
SZTAKI, Hungary
Tel: +36 1 279 6248
E-mail: andras.micsik@sztaki.mta.hu

MIDAS: Automated SOA Testing on the Cloud

by Alberto De Francesco, Claudia Di Napoli, Marc-Florian Wendland and Fabio De Rosa

MIDAS is a research project funded by the European Commission under the Seventh Framework Programme. Its goal is to design and build an integrated framework for the automation of Service Oriented Architecture (SOA) testing able to cover the complete lifecycle of software testing (test generation, execution, evaluation, planning and scheduling). The framework will be made available as Testing as a Service (TaaS) on a public Cloud infrastructure, so as to accommodate, in an affordable way, the varying and sometimes unpredictable computational requirements typical of testing activities.

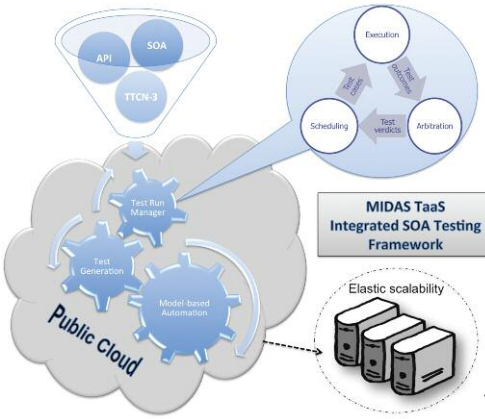


Figure 1: Sketch of the MIDAS TaaS framework

The MIDAS project is a three year European project, started in September 2012, that aims to realize a comprehensive TaaS framework able to support automation and intelligent management of Service Oriented Architecture (SOA) testing. The framework supports all testing cycle activities: test case planning, development and execution, reporting and result analysis, test campaign management and scheduling.

With the spread of Internet and Internet-related technologies, thousands of legacy and new applications, systems and devices are connected and collaborate, allowing the automation of business processes that support daily activities. The SOA design and implementation style is the most relevant technology allowing organizations to put into practice dynamic collaboration of loosely coupled systems in order to achieve flexible, dependable and secure business processes following the contract-based and model-driven paradigms.

In this context, SOA testing plays a crucial role in strengthening stakeholders' trust on the compliance of a SOA-based application with their business needs through rigorous, sound and open validation and verification processes. Nevertheless, SOA key characteristics such as lack of observable behaviour of the involved systems, lack of trust in the employed engineering methods, lack of direct control of the implementation lifecycles, make SOA testing a heavy, complex, challenging and expensive task. There are also additional challenges, including: late binding of systems, fundamental uncertainty of the test verdicts, organizational complexity, elastic demand of computational resources and increasing scale factor of the services architectures.

The MIDAS project will address these difficulties by providing:

- Design-time and run-time generation of test cases and oracles for the functional, interaction, security, usage-based and quality of service aspects of a service's architecture, based on black-box and grey-box testing methods, the only practicable verification methods for SOA,
- An environment for SOA testing allowing the automated configuration and initialization of test scenarios and the automated execution of test runs on a Service Architecture Under Test based upon a Testing and Control Notation version 3 (TTCN-3) distributed engine,
- Advanced methods and tools (probabilistic and symbolic inference-based ones) for the evaluation of test results and for the test campaign planning and scheduling, in order to help the tester to optimize the test campaign management on the basis of testing objectives.

To define detailed test specifications and test cases, the MIDAS platform will rely on TTCN-3 as a strongly typed test technology standardized and maintained by the European Telecommunication Standardization Institute (ETSI).

The MIDAS platform will itself be designed according to the SOA paradigm, and a prototype implementation will be delivered as a TaaS facility. The facility will be deployed on a public Cloud infrastructure providing elastic scalability of the testing environment, and allowing allocation of huge amounts of computation resources for relatively short test campaigns on very large services architectures. The adoption of the Cloud paradigms will also make the MIDAS platform accessible for SMEs that want to test their systems without making the commitment of a large investment. The MIDAS platform will be tested on two target pilots that are two real-world SOA-based applications, respectively in the domains of healthcare and supply chain management.

Moreover, the economic impact of inadequate infrastructures for SOA testing, and the improvements in terms of dependability and security of the "digital economy", coming from the availability of powerful and productive tools and infrastructures for SOA testing, will be evaluated to assess the MIDAS project outcomes.

Initiatives to activate relationships with standardization bodies, such as the ETSI TC-MTS (ETSI Methods for Testing and Specification Technical Committee), CEN (European Committee for Standardization), and the OMG (Object Management Group) will be undertaken with the purpose of assessing the impact and aligning the MIDAS approach and methodology with standardization efforts.

The MIDAS consortium is composed of universities, research centres, small and medium enterprises, and non-profit organizations with the necessary expertise to cover all the aspects outlined in the project.

Link:
http://www.midas-project.eu/

Please contact:
Riccardo Fontanelli, project coordinator,
Dedalus S.p.A., Italy
E-mail: riccardo.fontanelli@dedalus.eu

DRIVEN: Diagnostically Robust Ultrasound Video Transmission over Emerging Wireless Networks

by Andreas S. Panayides and Anthony G. Constantinides

The objective of the DRIVEN project is to develop a unifying framework for the diagnostically-resilient communication of medical video over emerging 3.5G and 4G wireless networks that is suitable for critically-needed clinical diagnosis. The goal is to develop new and integrate state-of-the-art methods in video coding, wireless networks, and medical video quality assessment into an adaptive, automated framework. It is anticipated that the integrated system will aid in establishing mobile-health (m-health) medical video communication systems in standard clinical practice.

Cardiovascular disease (CVD) causes nearly half of all deaths in Europe. The Medical Standing Committee of the European Science Foundation has recognized the development of new technologies for cardiovascular diseases as a target priority.

The Diagnostically Robust Ultrasound Video Transmission over Emerging Wireless Networks (DRIVEN) framework, which can support any medical video modality, is currently validated for remote diagnosis of two life threatening cardiovascular conditions: symptomatic atherosclerotic plaque (associated with critical stroke events) and abdominal aortic aneurysm (AAA), which can cause life-threatening internal bleeding. In both cases, timely intervention, especially for the elderly and patients with mobility problems will be supported by the proposed m-health ultrasound video transmission system.

The system also has the potential to assist with early diagnosis of patients in remote locations and mass population screening, especially in developing countries, to reduce patients' risks of more general cardiovascular diseases. Other application scenarios range from emergency telematics, which provide a critical time advantage for the patient's survival, to second opinion provision and medical education. The objective is to establish how specific diagnostic decisions can be safely made based on the transmitted medical video.

Figure 1 depicts the basic system components required for real-time transmission of medical ultrasound video. This procedure comprises four steps: raw medical video is first pre-processed so that it is suitable for encoding. This step typically involves video resolution and frame rate adjustments, as well as video de-noising where applicable. Video coding standards are then used to compress the video. Given the best available wireless transmission medium, the medical video is transmitted to the remote location where the reverse procedure is followed for decoding, post-processing and error recovery.

Over the last decade, there has been an impressive growth in the development of m-health systems and services [1]. In spite of this remarkable growth, there has been limited penetration into standard clinical practice. This is in part due to the inability of wireless networks of the previous decade to support data transfer rates that would allow medical video communications at the acquired video resolution and frame rate. The clinical capacity of the communicated medical video is compromised when the transmitted video is of limited resolution, and the same holds for clinical motion when

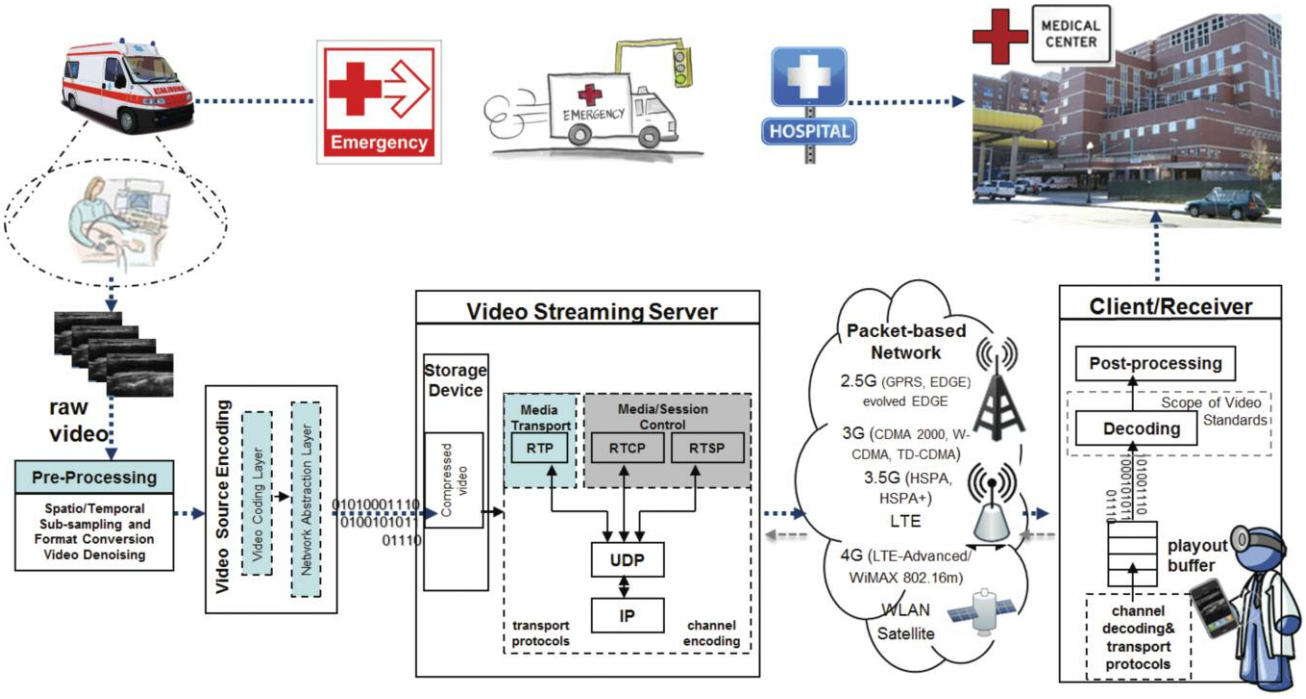


Figure 1: m-health Medical Video Communication System Architecture.

restricting the frame rate. Moreover, the compression ratio cannot be tolerated to fall below diagnostically acceptable levels. The efficiency of medical video transmission systems is measured in terms of their ability to provide for a reliable and dependable diagnosis. Degradation of clinically-sensitive regions can lead to unacceptable deterioration of the system's objective of remote diagnosis.

High-resolution low-delay and high frame-rate medical video communication over 3.5G mobile WIMAX and HSPA wireless networks using diagnostically driven techniques is demonstrated by the DRIVEN platform in [2], [3]. Diagnostically driven methods adapt the encoding, transmission, and evaluation process to the underlying medical video modality, aiming to maximize the diagnostic capacity of the transmitted medical video. Clinically-sensitive video regions form diagnostic regions-of-interest (d-ROI) which are in turn encoded in higher quality than non-diagnostically important regions, protected more strongly during transmission, and assigned higher-importance weights during video quality assessment.

On-going work includes integrating the new High Efficiency Video Coding (HEVC) standard in the DRIVEN framework, which together with 4G wireless networks deployment, is expected to play a decisive role in communicating medical video that can rival the standards of in-hospital examinations. The use of medical video at the clinically acquired resolution and frame rate that can be robustly transmitted in low-delay without compromising clinical quality will aid in wider adoption of m-health systems and services in standard clinical practice.

The DRIVEN project (12/2012-11/2014) is coordinated by the Communication and Signal Processing Group at Imperial College London and is funded by Marie-Curie Intra-European Fellowships scheme. Collaborators are the Electronic Health laboratory at the University of Cyprus (UCY), while medical ultrasound video acquisition, evaluation, and remote diagnosis are performed by medical experts at Imperial College, the Vascular Non-invasive Diagnostic Centre (VNDC) in London, UK, and the Cyprus Institute of Neurology and Genetics (CING) in Nicosia, Cyprus.

References:

[1] E. Kyriacou et al.: "m-Health e-Emergency Systems: Current Status and Future Directions", IEEE Antennas Propag. Mag., vol. 49(1), pp. 216-231, Feb. 2007.
[2] A. Panayides et al.: "High-Resolution, Low-delay, and Error-resilient Medical Ultrasound Video Communication Using H.264/AVC Over Mobile WiMAX Networks," IEEE J-BHI vol.17(3), pp.619-628, May 2013.
[3] A. Panayides, I. Eleftheriou, M. Pantziaris, "Open-Source Telemedicine Platform for Wireless Medical Video Communication," Int J Telemed Appl, Article ID 457491, 2013.

Please contact:

Andreas S. Panayides,
Imperial College London, UK.
E-mail: a.panagidis@imperial.ac.uk

An e-Science Collaboration Platform for Effective Multimedia Research

by Péter Mátételki, László Havasi and András Micsik

As part of the CrossMedia project, researchers from MTA SZTAKI have created a new testbed to support media retrieval research activities. The platform consists of a portal providing domain-specific functionality, collaborative features and a multimodal user interface, and is supported by a robust and scalable backend system.

Researchers within the field of media retrieval have always been in a tough situation as they need to deal with huge amounts of test data while evaluating the results of their scientific activities. Since processing visual and audio media requires large amounts of memory and computing power, there is an infrastructural need to efficiently index, store, and retrieve information from multimedia databases. Sharing achievements or performing joint activities is a hard task in such a heavy-duty environment.

Our new e-science platform supports collaborative research communities by providing a simple solution to develop semantic and media search algorithms on common datasets. The project was funded by MTA SZTAKI – Hungarian Academy of Sciences, Institute for Computer Science and Control, and was executed by the Department of Distributed Systems and the Distributed Events Analysis Research Laboratory.

The platform's architecture is shown in Figure 1. The system's functionality can be used through the portal served by a distributed backend system organized in a loosely coupled service-oriented architecture:

- The Media Store (MS) is responsible for safekeeping all searchable multimedia elements.
- The Media Indexer and Search Subsystem (MISS) is responsible for generating index trees for a specific algorithm on a specific media set in the MS and it is also capable of executing similarity-based search queries.
- The Semantic Indexer and Search Subsystem (SISS) is responsible for creating semantic databases and indices and executing semantic search queries.
- The Search Fusion Subsystem (SFS) is responsible for combining the results of the MISS and SISS in case of multi-input multimodal search expressions.
- The Search User Interface (SUI) enables users to easily create complex multimodal search expressions and to evaluate results.
- The E-Science Community Portal (ECP) is responsible for integrating and providing all the functionality through a Web2.0 interface enabling users to perform collaborative research.

We separated the community management (ECP) and multimedia management (MS) functionality into loosely coupled components. This separation detaches storage functionality (millions of test data for the content based search) from the

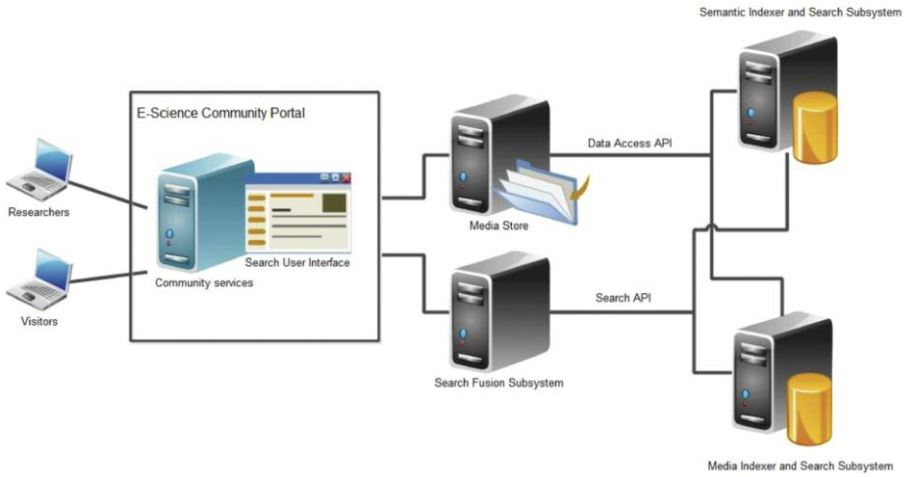


Figure 1: Architecture of the CrossMedia e-science platform.

community portal's permission control and keeps these management tasks independent, resulting in a flexible, scalable and responsive ecosystem overall.

Researchers can upload their media indexer algorithms through the portal. For evaluation, media collections can be created to build index trees for an algorithm. An image index is built using one indexer algorithm and one or more media collections. Defining an index on the portal interface launches a series of asynchronous automated operations, while the portal is regularly updated with status information. Once ready, the generated index becomes available for testing via the user interface.

An index can either be content-based or annotation-based (semantic). Semantic and free text annotations can be attached to any media item, enabling the SISS to perform semantic search. The MISS overcomes the performance problems of content-based indices using a hybrid indexing structure (RAM-SSD-HDD combination) in a locally distributed computational framework. The applications are not limited to well-formed feature descriptors; indices can receive arbitrary binary data as a feature with the relating distance definition.

The applied LHI-tree is similar to M-index where base points are chosen randomly to reduce the high-complexity space. The LHI-tree uses base points to compute reference distances and to calculate hash codes for every input vector from the quantized distances. To assign a disk partition to a part of the feature space, we used a hashing function of quantized distances.

For the visual content indexing we built a descriptor composed of four different information representations: edge histogram, entropy histogram, pattern histogram, dominant colour characteristics. The descriptor's dimensionality is 52. We experimentally proved that a good choice for similarity measure is the weighted Euclidean distance where the fusion of the different features is carried out by tuning the weighting scalars.

Sample semantic indices were built for the CoPhiR database by matching photo metadata - such as tags and titles - to DBpedia nodes using entity extraction. Various semantic queries were then built, eg to find photos taken at a given time of day (sunset, morning), containing a given plant or animal, or taken at a given place. Our semantic reasoner exploits the transitivity of the semantic relations, therefore we find photos tagged with a narrower search term than the original one (eg trains -> Shinkansen).

The available functionality not only satisfies the domain-specific needs of an individual researcher but also offers community-based collaboration facilities as users can be engaged in research groups. Group members can work in the group's private space and share content with other groups or with the public.

The portal supports two ways for the fusion of the output of different indices: a general approach, where the lists are internally aggregated and re-ranked using the fitness value. The second option suits the non-compatible modalities (eg semantic-visual); in this case the visual search directly reduces the search space of the semantic search engine.

The search interface allows image descriptors and semantic indices to be tested. Users can assemble complex multipart search queries, where each query consists of a media item and an index. Different indices may be combined in a multimodal query. Results for multipart queries are unified by the SFS using internal weighting mechanisms, which can be fine-tuned by the user moving an item forward or backward in the result list.

The CrossMedia e-Science Community Portal enables researchers to work in groups and to collaborate with other research communities. The infrastructure ensures scalable and fast manipulation of indices, while the user interface provides testing and evaluation facilities. To enable the portal to be utilized to its full potential, we plan to adopt the SZTAKI Cloud infrastructure, which will ensure maximum processing speed and availability.

Link: http://dsd.sztaki.hu/projects/CrossMedia/en/

References:

[1] P. Mátételki et al. CrossMedia: Supporting Collaborative Research of Media Retrieval. Procedia - Social and Behavioral Sciences, Volume 73, 27 February 2013, ISSN 1877-0428, http://dx.doi.org/10.1016/j.sbspro.2013.02.083

[2] L. Havasi et al. Search in WikiImages using mobile phone. International Workshop on Content-Based Multimedia Indexing, CBMI 2013, Veszprém, Hungary.

Please contact:

Péter Mátételki, SZTAKI, Hungary
Tel: +36 1 279 6271
E-mail: peter.matetelki@sztaki.mta.hu

Control Systems and Technologies for Cyber-Physical Systems

by Francoise Lamnabhi-Lagarrigue

A joint workshop of the BALCON project (The Gateway to ICT Monitoring & Control Research in the Western Balkans) and HYCON2 network of excellence (Highly-complex and networked control systems), held on 2-3 July 2013 in Belgrade, highlighted recent research efforts and challenges in cyber-physical systems (CPS).

"Cyber-physical systems (CPS) are physical, biological and engineered systems whose operations are monitored, coordinated, controlled and integrated by a computing and communication core". This definition was adopted at the Workshop on Control of Cyber-Physical Systems held at the University of Notre Dame London Centre, 20-21 October 2012, organized by Panos Antsaklis, Vijay Gupta and Karl Henrik Johansson [1] Desired characteristics of well-designed and engineered CPS include: coordinated, distributed, connected, heterogeneous, robust and responsive, providing new capability, adaptability, resilience, safety, security, and usability.

This intimate coupling between the cyber and physical will be manifested across a broad range of length scales, from the nano-world to large-scale wide-area systems of systems. Correspondingly, CPS will often exhibit dynamics at a wide range of time-scales from the discrete clock scale for some computational aspects, to multi-day or even year-long time scales for system-wide properties and evolution. Applications with enormous societal impact and economic benefit will be created and cyber-physical systems will transform how we interact with the physical world just as the Internet transformed how we interact with one another.

Control is a major CPS challenge, inherent in many, if not most, cyber-physical systems since it exists at the intersection of the cyber and physical components, making use of each and mediating their interaction. Very difficult challenges are posed for control of CPS, however, due to a variety of factors such as very broad time and length scales, the presence of network communication and delays, coordination of many components (with an associated increased risk of component failure as the number of components grows to be very large), model reduction, tractability, etc.

In order to highlight recent research efforts and bring together European and West Balkans researchers to share results and discuss the major research challenges in CPS, a joint workshop of the BALCON and HYCON2 projects was organized on 2-3 July 2013 in Belgrade. The main topics included:

- Taming system complexity and big data: control hierarchies and tropical algebras; mining from the medical databases; mining from the sensor networks; integrated model-based support for the design of complex controlled systems.
- Stability and estimation tools: adversary control strategy for cyber-physical networked systems; advances in the stability analysis of interconnected nonlinear systems by ISS-related concepts; robustness of stochastic discrete-time switched linear systems with application to control with shared resources; global exponential sampled-data observers for nonlinear systems with delayed measurements; non-asymptotic estimation for online systems: finite-time algorithms and applications.
- Control under strong communication constraints: robust self-triggered coordinated control; wireless control networks and controllability with switched delays; fault tolerant control of multi-hop control networks; passification based control, estimation and synchronization in CPS.
- Large-scale systems: a critical bisimulation approach to safety analysis of large scale complex systems; a Markov decision framework for online learning in residential demand response; adaptive scheduling in LTE networks for SmartGrid applications; plug-and-play control for complex, real-life large-scale systems.
- Embedded systems: soft real–time scheduling for embedded control systems; real-time operating systems for eHealth wearable devices; a motion planner for mobile vehicles on rough terrains.

Links:
BALCON Project:
http://www.balcon-project.eu/
HYCON2 Network of Exellence:
http://www.hycon2.eu
Workshop presentations:
http://www.hycon2.eu/?page=14&id=47

Reference:
[1]: Report from the Workshop on Control of Cyber-Physical Systems, London, 20-21 October 2012, http://controls.ame.nd.edu/mediawiki/images/c/c4/Workshopreport.pdf

Please contact:
Francoise Lamnabhi-Lagarrigue
CNRS, Laboratoire des Signaux et Systemes & European Embedded Control Institute (EECI), SUPELEC, France
E-mail: Francoise.Lamnabhi-Lagarrigue@lss.supelec.fr

HCI International 2013

by Constantine Stephanidis

HCI International 2013, the 15th International Conference on Human-Computer Interaction, was held in Las Vegas, Nevada, USA, 21-26 July 2013.

HCI International 2013 and the affiliated conferences explored a wide variety of new hot topics which reflect and contribute to a paradigm shift towards ubiquitous interaction, intelligent environments and interactive technologies supporting virtually any aspect of human life and activities in a global and social perspective.

The 29-volume conference proceedings are published by Springer. Papers appear in volumes of the Lecture Notes in Computer Science (LNCS) and Lecture Notes in Artificial Intelligence (LNAI) series. Posters are published in the Communications in Computer and Information Science (CCIS) series. All volumes are available on-line through the SpringerLink Digital Library, readily accessible by all subscribing libraries around the world.

An impressive number of more than 2300 people from 62 countries participated in this truly international in scope event, where the work of the world's foremost leaders in the field was presented. In 264 sessions, 1666 papers were presented, and 306 posters were displayed during specific sessions. The keynote speaker was Prof. Hiroshi Ishii (Massachusetts Institute of Technology, USA) and the title of his address was "Defy Gravity: The Art of Tangible Bits". Furthermore, during the opening session, for the third time in the history of the HCI International Conference series, fourteen awards were conferred. Twelve awards were conferred to the best papers in each affiliated conference / thematic area. Among these twelve best papers, one paper was selected as Best HCI International 2013 Conference paper. Finally, the Best Poster also received an award.

Becoming annual - HCI International 2014

The 16th International Conference on Human-Computer Interaction will be held jointly with the affiliated Conferences in Creta Maris Resort and Convention Centre, Heraklion, Crete, Greece, 22 - 27 June 2014.

HCI International 2014 will also include two new affiliated conferences: the 1st International Conference on HCI in Business, and the 1st International Conference on Learning and Collaboration Technologies.

Please visit the conference website for the full list of the thematic areas, including the topics for each thematic area.

Figure 1: Hiroshi Ishii, Keynote Speaker

The program will feature, among others: pre-conference half-day and full-day tutorials, parallel sessions, poster presentations, an opening session with a keynote address, and an exhibition.

The proceedings, including both papers and posters, will be published by Springer, and will be indexed by a number of services including EI and ISI CPCI-S.

Links:
HCI International 2013,
http://www.hcii2013.org/
HCI International 2014,
http://www.hcii2014.org/
HCI International Conference Series,
http://www.hci-international.org/

Please contact:
Constantine Stephanidis
General Chair, HCI International Conference
E-mail: cs@ics.forth.gr

Figure 2: HCII 2013 Award winners

CLEF 2013 and Beyond: Evolution of the CLEF Initiative

by Nicola Ferro and Paolo Rosso

The CLEF Initiative is structured in two main parts: a series of Evaluation Labs, to conduct evaluation of information access systems and a peer-reviewed Conference on a broad range of issues on evaluation. The annual CLEF events have been partially supported by the EU FP7 PROMISE project (contract n. 258191) and by the ELIAS RNP network.

CLEF 2013: Information Access Evaluation meets Multilinguality, Multimodality, and Visualization

The annual meeting of the CLEF Initiative was hosted this year by the Universitat Politècnica de València, Spain, 23-26 September as a 4 days event, attended by more than 200 participants, in which conference presentations, laboratory meetings, workshops, and community sessions were smoothly interleaved to provide a continuous stream of discussions on the different facets of experimental evaluation.

22 papers were accepted for the Conference and published by Springer in their Lectures Notes for Computer Science (LNCS) series. Two keynote speakers highlighted important developments in the field of evaluation. Evangelos Kanoulas, Google Zurich, Switzerland, focused on how to model variance in topics, collections, system and measures when evaluating systems. Rada Mihalcea, University of Michigan, USA, presented several scenarios where multilinguality can be pivoted to improve natural language processing performances instead of being a barrier to be crossed.

Nine benchmarking activities ran as evaluation labs in CLEF 2013. The results were presented and discussed in Valencia in dedicated sessions:

- CHiC (Cultural Heritage in CLEF): investigating systematic and large-scale evaluation of cultural heritage digital libraries and information access systems;
- CLEFeHealth (new): aiming at developing processing methods and

resources to enrich difficult-to-understand health text as well as their evaluation setting;
- CLEF-IP: studying IR techniques in the patent domain;
- ImageCLEF: proposing experimental evaluation of image classification and retrieval, with a focus on the combination of textual and visual evidence;
- INEX: evaluating XML retrieval;
- PAN: uncovering plagiarism, authorship and social software misuse;
- QA4MRE: evaluating machine reading systems through question answering and reading comprehension tests;
- QALD-3 (Question Answering over Linked Data, new): a benchmarking activity on question answering over linked data;
- RepLab: evaluating reputation management technologies.

There was also an exploratory workshop CLEF-ER: a workshop on multilingual annotation of named entities and terminology resources acquisition.

CLEF 2014 and Beyond

During CLEF 2013 considerable steps were taken to further reshape and improve the overall organization of CLEF. In particular, it has been decided to establish a legal entity to support the organization of the CLEF initiative.

CLEF 2014 will be hosted by the University of Sheffield, UK, 15-19 September 2014 while CLEF 2015 will be hosted by the Institut de Recherche en Informatique de Toulouse (IRIT), France, in early September 2015.

Finally, bids for hosting CLEF 2016 are now open and will close on 7th April 2014. Proposals can be sent to the CLEF Steering Committee Chair at chair@clef-initiative.eu.

Links.
CLEF: http://www.clef-initiative.eu/
CLEF 2013 Web site:
http://www.clef2013.org/
CLEF 2014 Web site:
http://clef2014.clef-initiative.eu/
PROMISE: http://www.promise-noe.eu/
ELIAS: http://www.elias-network.eu/

Please contact.
Nicola Ferro
CLEF Steering Committee Chair
University of Padua, Italy
E-mail: ferro@dei.unipd.it

Jay Parikh, the Vice-President of Infrastructure Engineering at Facebook

VLDB 2013 Conference Supported by ERCIM

by Yannis Velegrakis

Very Large Data Bases (VLDB) is the premier annual international conference on database technologies, attracting more than 700 researchers, vendors, practitioners, application developers, and users from the top universities and the most important industrial players from the entire world. The conference features research talks, tutorials, demonstrations, and workshops, and covers current issues in data management, and database and information systems research. Data management and databases are among the main technological cornerstones of emerging applications of the twenty-first century.

This edition of VLDB took place in Riva del Garda (Italy) from 26-30 August 2013, and was organized by Prof. Themis Palpanas and Prof. Yannis Velegrakis, co-directors of the Data and Information Management group at the University of Trento. The conference attracted 730 participants from 40 countries - the highest number of participants in the last 27 years (and the second highest in the history of VLDB). The participants represented all the top international universities, as well as all the major international companies and industrial research labs, including Google, Oracle, Facebook, IBM, Microsoft, SAP, Twitter, Yahoo!, NEC, Hewlett-Packard, Greenplum, HTC, and EMC2.

A central theme of this year's conference was Big Data and its related issues. The technical program included three keynote talks: the first by Jay Parikh, the Vice-President of Infrastructure Engineering at Facebook, on how to build data infrastructures for web-scale problems; the second by Sam Madden, professor at MIT and leader of the MIT big data group, on a collaborative data analytics platform for large-scale data; and the third by Cynthia Dwork from Microsoft Research, on the problem of performing large-scale data analytics whilst preserving privacy. It also included two panel discussions on the information hidden in big data, and on experiences from startup companies, as well as four demos, six tutorials, four industrial and 16 research sessions. The videos of the keynotes and panels are available on the VLDB 2013 web site, alongside all the papers presented at the conference. The papers are published by the VLDB endowment and will also be available on the VLDB web site and in the ACM Digital Library.

ERCIM's support of the conference played a significant role in its success, helping to attract high-profile speakers offering talks on state-of-the-art topics. Furthermore, the numerous ERCIM leaflets, articles and magazines that were supplied and distributed during the conference acted as excellent advertisements for European research.

More information:
http://www.vldb.org/2013/
http://www.vldb.org/

The First Tangible Interaction Studio

by Nadine Couture

"Tangible interaction", a research field addressing the interface between the physical and the digital worlds, provides a new tool to facilitate the study of digital data. It overcomes the limitations of the traditional mouse, facilitating interaction and manipulation of 3D digital objects using real objects.

Tangible interaction research has a vast and diverse array of applications, for example: in music with the "reactable" used by the singer Björk, in cultural mediation to explain the stages of evolution, in natural history areas such as cladistics, in archaeology to reassemble archaeological fragments, in manufacturing for assembling aircraft or car engines, and in energy to validate geological models of subsoil.

For many years, researchers, designers, practitioners and artists around the world have been establishing the theoretical and empirical foundations of this field: designing technical frameworks, prototypes and interfaces. There is still great potential for future work, and the European community could be more active at an international level. Researchers from ESTIA (France), the Bauhaus University (Germany) and the HES-SO (Switzerland) organized the first Tangible Interaction Studio at ESTIA, in the village of Bidart, in the heart of Basque Country, France, from 26 to 30 of August 2013. The aim of this meeting was to connect young European researchers in the field of Tangible Interaction with internationally renowned researchers, and to establish networks for their academic future. In doing so we hope to ensure a strong future for the discipline.

The TANGINT/FR group, supported by AFIHM (the Francophone Association in Human Computer Interaction) is behind the Tangible Interaction Studio. This French speaking working group was created in 2011 on the initiative of Prof. Dr. Nadine Couture and Dr. Guillaume Rivière from ESTIA. This event follows a first workshop held in June 2012, which brought together

Workshop participants

Nadine Couture signing the joint expression of interest

French and German researchers. This year, on the occasion of the inaugural Studio, Swiss, German, English, Italian, Austrian, Dutch, Chinese, Spanish and French researchers joined the working group.

At the closing of the Tangible Interaction Studio, on 30 August, the Institute of Advanced Industrial Technologies ESTIA (France), the University of Applied Sciences of Western Switzerland – HES-SO (Switzerland), the Bauhaus-Universität Weimar (Germany), the Université de Toulouse (France), the Ludwig-Maximilians-Universität München (Germany), and the University of Arts and Industrial Design Linz (Austria) signed a Joint Expression of Interest. The purpose of the agreement is to facilitate student and staff exchanges and to increase the involvement of the organization in summer-school and studios. The different groups have thereby agreed to devote a joint effort towards research on a European scale in the areas related to human-computer interaction, tangible interaction and smart environments. This convention represents a first step towards the constitution of a network that could lead to the creation of a European Network of Excellence (NoE) on Tangible Interaction.

The studio's program spanned a five day period and was led by both internationally recognized senior researchers and PhD students. Day one of the program focused on laying the theoretical foundation thanks to Eva Hornecker, Michel Beaudoin-Lafon and Emmanuel Dubois. Nadine Couture and Brygg Ullmer gave an opening presentation (the latter via video message). A "panel session" moderated by Catherine Letondal concluded the first day. On Day two, we focused on applications and concepts thanks to Allan Dix and Martin Kaltenbrunner who gave a talk and a hands-on demonstration. The end of days two and three were devoted to student presentations by young researchers: Marie Schacht, Leonardo Angelini, Marvin Schneider, Nassrin Hajinejad, Renaud Gervais and Simon Stusak. On Days four and five, four workshops were led by Leonardo Angelini, Jochen "Jeff" Rick, Arthur Vanpoucke and Marie Schacht. Exhibitions were submitted by Patrick Reuter, Simon Stusak, Stéphane Kreckelbergh and Guillaume Rivière. We decided that the Tangible Interaction Studio would take place in Weimar, Germany, in 2014, and in Fribourg, Switzerland, in 2015.

As a conclusion we asked participants to write down their "one word summary" of their experience at the studio, and the summaries included: design, atmosphere, food, inspired, exchange, tangible, friend, motivation, thanks, fun, interaction, knowledge, instructive, movement, creativity, arduino, interconnect, work, wonderful, perfect, rich.

Links:
http://www.fgtis.estia.fr
http://www.tangint.org/fr
http://be-greifbar.de/
http://www.estia.fr

Please contact:
Nadine Couture, ESTIA, France
Tel : +33 5 5943 8467
E-mail : n.couture@estia.fr

EvAAL Evaluation Workshop

by Francesco Potortì

For the third consecutive year, the EvAAL competition gathered teams from all around the world to compare their systems in the participating living labs. EvAAL, a project embedded into the AALOA association, was led by the Institute of Information and Technologies (ISTI) of CNR (Pisa, IT) and was mostly financed by the universAAL FP7 project.

EvAAL 2013 awards ceremony at the AAL Forum

EvAAL, which stands for Evaluating AAL Systems through Competitive Benchmarking, aims at establishing benchmarks and evaluation metrics to compare Ambient Assisted Living solutions and assess advances in the field.

During the competitions, data were collected in realistic environments, the living labs: the data sets, which are publicly available, can be used by researcher communities as an aid for the simulation and test of their solutions.

This year, the competition was composed of two tracks, a demo, and a final workshop:

- Competition on Indoor Localization and Tracking for AAL held on 1-5 July 2013 at the Living Lab of the Polytechnic University of Madrid
- Competition on Activity Recognition for AAL, held on 8-12 July 2013 at the CIAmI Living Lab in Valencia
- Demo on Companion Robots for AAL, held on 30 July 2013 at DomoCasa Lab in Pisa
- EvAAL workshop, held as a side event of the AAL Forum in September 2013, in Norrköping.

More information:
http://evaal.aaloa.org

Joint ERCIM, ARTEMIS, Euromirco Workshops

by Erwin Schoitsch

SEAA 2013

A special workshop session on Teaching, Education and Training for Dependable Embedded and Cyber-physical Systems was held in cooperation with the ERCIM Dependable Embedded Systems Working Group at the 39th Euromicro Conference on Software Engineering and Advanced Applications (SEAA 2013) Santander, Spain September 4-6, 2013. The workshop session was part of the regular session scheme of SEAA 2013. It included four presentations:

- "Efficient embedded systems education by adopting component based software development paradigm" by Sasikumar Punnekkat, Mälardalen University, Sweden
- "Reuse in Safety Critical Systems: Educational Use Case" by Miren Illarramendi Rezabal, Xabier Elkorobarrutia Letona and Leire Etxeberria, University of Mondragon, Spain
- "Teaching and Training Formal Methods for Safety Critical Systems" by Michael Lipaczewski and Frank Ortmeier, Otto von Guericke University, Magdeburg, Germany
- "European Perspectives on Teaching, Education and Training for Dependable Embedded and Cyber-physical Systems" by Erwin Schoitsch, AIT, Vienna, and Amund Skavhaug, NTNU, Norway.

The papers will be published in the conference proceedings published by IEEE.

SAFECOMP 2013

An ERCIM/EWICS/ARTEMIS workshop on Dependable Embedded and Cyber-physical Systems (DECS) attracted 26 participants at the 32nd International Conference on Computer Safety, Reliability and Security (SAFECOMP 2013), Toulouse, France, 24 September 2013

After an introduction on "ERCIM, EWICS, ARTEMIS: Embedded Systems Safety, Security and European Strategy" followed a session with four papers on "Dependable Embedded Systems Applications". One part of workshop was then dedicated to Robotics and Autonomous Systems, mainly reporting from the ARTEMIS project "R3-COP" (Resilient Reasoning Robotic Co-operating Systems), with two sessions (four papers each). Another part of the workshop covered "Systems Safety Analysis and Fault Tolerance" with the presentation of three papers. Proceedings will be available electronically in the HAL Open Archives (http://kwz.me/5q).

Links:
http://seaa2013.ii.metu.edu.tr/
http://kwz.me/0F

Please contact:
Erwin Schoitsch
ERCIM DES WG Chair
AIT Austrian Institute of Technology GmbH/AARIT
E-mail: Erwin.Schoitsch@ait.ac.at

Call for Participation

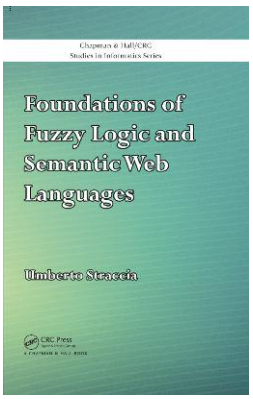
IEEE International Symposium on Signal Processing and Information Technology

Athens, 12-15 December 2013

The IEEE ISSPIT 2013 is the thirteenth in a series of international symposia aiming at the coverage of key aspects in the fields of signal processing and information technology.
Topics include:

- Signal processing theory and methods
- Signal processing for communications
- DSP architectures and implementation
- DSP for space applications/highly available architectures
- Multimedia signal processing
- Image and multidimensional signal processing
- Audio and electro acoustics
- Sensor array and multi-channel
- Speech processing
- Radar signal processing
- Neural networks
- Internet software architectures
- Multimedia and image-based systems
- Mobile computing and applications
- E-Commerce and pricing
- etc.

More information:
http://www.isspit.org/isspit/2013/



Umberto Straccia

Foundations of Fuzzy Logic and Semantic Web Languages

Managing vagueness/fuzziness is starting to play an important role in Semantic Web research, with a large number of research efforts underway. Foundations of Fuzzy Logic and Semantic Web Languages provides a rigorous and succinct account of the mathematical methods and tools used for representing and reasoning with fuzzy information within Semantic Web languages. The book focuses on the three main streams of Semantic Web languages: Triple languages RDF and RDFS; Conceptual languages OWL and OWL 2, and their profiles OWL EL, OWL QL, and OWL RL; and Rule-based languages, such as SWRL and RIF.

Written by a prominent researcher in this area, the book is the first to combine coverage of fuzzy logic and Semantic Web languages. The first part of the book covers all the theoretical and logical aspects of classical (two-valued) Semantic Web languages. The second part explains how to generalize these languages to cope with fuzzy set theory and fuzzy logic. With an extensive bibliography, this book provides in-depth insight into fuzzy Semantic Web languages for non-fuzzy set theory and fuzzy logic experts. It also helps researchers of non-Semantic Web languages get a better understanding of the theoretical fundamentals of Semantic Web languages.

Chapman & Hall/CRC Studies in Informatics Series, 2013, hardcover, 386 pages, ISBN: 978-1439853474.

Matteo Mio Wins Ackermann Award

Matteo Mio, ERCIM postdoctoral fellow currently hosted by CWI and previously by Inria (and Ecole Polytechnique), received the 2013 Ackermann Award for his PhD thesis "Game Semantics for Probabilistic mu-Calculi" obtained at the University of Edinburgh in 2012 under the supervision of Prof. Alex Simpson. The Ackermann Award is the "Outstanding Dissertation Award for Logic in Computer Science" of the European Association for Computer Science Logic (EACSL). The award was presented in a ceremony during the 22nd EACSL Annual Conference on Computer Science Logic, Torino, Italy, 2-5 September 2013 (CSL'13).
http://www.eacsl.org/award.html

Fabio Martinelli Appointed Chair of the WG3 Working Group on "Secure ICT Research and Innovation"

Fabio Martinelli, senior researcher of the Institute of Informatics and Telematics of the Italian National Research Council, Pisa, and founder of the ERCIM Working Group on Security Trust and Management, has been appointed Chair of the WG3 working group on "Secure ICT Research and Innovation". The group is part of the Network and Information Security (NIS) public private Platform. It will provide input on how to improve cybersecurity risk management and information sharing to the European Commission and will contribute to devising the European secure ICT research and Innovation agenda. The kickoff meeting took place in Brussels in late September and attracted participants from the public sector (national authorities, research agencies) and the private sector (companies working in ICT, finance, post, transport, healthcare, defence and energy).
https://ec.europa.eu/digital-agenda/en/news/nis-platform-kick-meeting-working-groups

ERCIM Fellowship Programme

ERCIM offers fellowships for PhD holders from all over the world.

Next deadlines for applications: 31 October 2013 and 30 April 2014 (expected).

Topics cover most disciplines in Computer Science, Information Technology, and Applied Mathematics. Fellowships are of 12 month duration.

Conditions:
Applicants must:
- have obtained a PhD degree during the last 8 years (prior to the application deadline) or be in the last year of the thesis work with an outstanding academic record.
- have completed the PhD before starting the grant (a proof will be requested).
- be fluent in English.
- be discharged or get deferment from military service.

More information: http://www.ercim.eu/activity/cor-baayen-award

ERCIM
European Research Consortium for Informatics and Mathematics

ERCIM – the European Research Consortium for Informatics and Mathematics is an organisation dedicated to the advancement of European research and development, in information technology and applied mathematics. Its member institutions aim to foster collaborative work within the European research community and to increase co-operation with European industry.

ERCIM is the European Host of the World Wide Web Consortium.

Austrian Association for Research in IT
c/o Österreichische Computer Gesellschaft
Wollzeile 1-3, A-1010 Wien, Austria
http://www.aarit.at/

Portuguese ERCIM Grouping
c/o INESC Porto, Campus da FEUP,
Rua Dr. Roberto Frias, nº 378,
4200-465 Porto, Portugal

Consiglio Nazionale delle Ricerche, ISTI-CNR
Area della Ricerca CNR di Pisa,
Via G. Moruzzi 1, 56124 Pisa, Italy
http://www.isti.cnr.it/

Science and Technology Facilities Council
Rutherford Appleton Laboratory
Chilton, Didcot, Oxfordshire OX11 0QX, United Kingdom
http://www.scitech.ac.uk/

Czech Research Consortium
for Informatics and Mathematics
FI MU, Botanicka 68a, CZ-602 00 Brno, Czech Republic
http://www.utia.cas.cz/CRCIM/home.html

Spanish Research Consortium for Informatics and Mathematics
D3301, Facultad de Informática, Universidad Politécnica de Madrid
28660 Boadilla del Monte, Madrid, Spain,
http://www.sparcim.es/

Centrum Wiskunde & Informatica
Science Park 123,
NL-1098 XG Amsterdam, The Netherlands
http://www.cwi.nl/

SICS Swedish ICT
Box 1263,
SE-164 29 Kista, Sweden
http://www.sics.se/

Fonds National de la Recherche
6, rue Antoine de Saint-Exupéry, B.P. 1777
L-1017 Luxembourg-Kirchberg
http://www.fnr.lu/

Swiss Informatics Research Association
c/o Professor Abraham Bernstein, Department of Informatics,
University of Zurich, Binzmühlestrasse 14, CH-8050 Zürich
http://www.sira.s-i.ch/

FWO
Egmontstraat 5
B-1000 Brussels, Belgium
http://www.fwo.be/

F.R.S.-FNRS
rue d'Egmont 5
B-1000 Brussels, Belgium
http://www.fnrs.be/

Magyar Tudományos Akadémia
Számítástechnikai és Automatizálási Kutató Intézet
P.O. Box 63, H-1518 Budapest, Hungary
http://www.sztaki.hu/

Foundation for Research and Technology – Hellas
Institute of Computer Science
P.O. Box 1385, GR-71110 Heraklion, Crete, Greece
http://www.ics.forth.gr/

University of Cyprus
P.O. Box 20537
1678 Nicosia, Cyprus
http://www.cs.ucy.ac.cy/

Fraunhofer ICT Group
Anna-Louisa-Karsch-Str. 2
10178 Berlin, Germany
http://www.iuk.fraunhofer.de/

University of Southampton
University Road
Southampton SO17 1BJ, United Kingdom
http://www.southampton.ac.uk/

Institut National de Recherche en Informatique
et en Automatique
B.P. 105, F-78153 Le Chesnay, France
http://www.inria.fr/

Universty of Warsaw
Faculty of Mathematics, Informatics and Mechanics
Banacha 2, 02-097 Warsaw, Poland
http://www.mimuw.edu.pl/

Norwegian University of Science and Technology
Faculty of Information Technology, Mathematics and Electrical Engineering, N 7491 Trondheim, Norway
http://www.ntnu.no/

Universty of Wroclaw
Institute of Computer Science
Joliot-Curie 15, 50–383 Wroclaw, Poland
http://www.ii.uni.wroc.pl/

I.S.I. – Industrial Systems Institute
Patras Science Park building
Platani, Patras, Greece, GR-26504
http://www.isi.gr/

Technical Research Centre of Finland
PO Box 1000
FIN-02044 VTT, Finland
http://www.vtt.fi/