ANNUAL SCIENTIFIC REPORT 2012-2013
2013 was particularly busy with, in addition to our research and technology transfer activities, the review of the CEA LIST Institute research activities by the AERES committee. Therefore, we decided to exceptionally skip the publication of our scientific report for the year 2012 but to consolidate both years 2012 and 2013 in a single report.

Following the positive evaluation of CEA LIST, I am particularly proud to share with you this 2012-2013 scientific report. Despite its volume, we lack space to present all our research results and I invite you to visit our website (www-list.cea.fr) to learn more about our research activities, but also about our collaborations and the companies that trust us and with which we conduct our R&D work.

The CEA LIST Carnot Institute develops, within the Research and Technology Direction of CEA, R&D activities focused on new technologies for smart digital systems bringing beyond state of the art research results to efficient innovation transferred to industry.

With more than 300 publications per year in major scientific journals and international conferences, 65 patents applications per year and 76 PhD thesis defended during the period, CEA LIST continues to invest in high quality research in order to create key enabling technologies which will support innovative solutions for future industrial needs.

CEA LIST, as part of the Paris-Saclay Campus ecosystem, pursues a voluntary strategy to establish privileged partnership through the Digiteo alliance and the Nano-INNOV integration center.

Overall, in 2013, CEA LIST continued its successful path, with new major industrial partnerships backed by Carnot partnerships, technology platforms that structure our activity, a scientific assessment that places us at the highest international level: In this context, and once again, the CEA LIST 2012-2013 scientific report translates this balance between high level R&D and technology transfer by presenting an anthology of our best results obtained during the year.
SUMMARY

AMBIENT INTELLIGENCE ........................................................................................................ 6

DATA ANALYSIS ................................................................................................................... 7

Multi-Agent Systems (MAS) modelling and management .................................................... 7
Understanding ageing of electric vehicle through statistical data analysis .......................... 8
Uncalibrated non-Lambertian photometric stereo Application to microbial colonies surface reconstruction .................................................................................. 9
Predicting meteorological variables with neural networks ................................................... 10
From sparse coding to predictive and representation learning: a spike-based approach ....... 11
A fuzzy spatio-temporal approach for activity recognition .................................................. 12

DIAMOND MATERIAL & SENSORS ..................................................................................... 13
Single crystal diamond dosimeter for dose measurement in small beams used in stereotactic radiotherapy .............................................................. 13
Surface transfer doping can mediate both colloidal stability and self-assembly of nanodiamonds ................................................................................. 14
Boosting the electrochemical properties of diamond electrodes using carbon nanotube scaffolds ..................................................................... 15
Grafting odorant binding proteins on diamond bio-MEMS ................................................... 16
Optical analysis of p-type surface conductivity in diamond with slotted photonic crystals ........................................................................... 17

INTERNET OF THINGS ......................................................................................................... 18
Belt Mounted IMU With Enhanced Distance Estimation For Pedestrian Indoor Positioning .................................................................................. 18
Influence of interfacial adhesion on the mechanical response of magneto-rheological elastomers at high strain ....................................................... 19
Design and control of high performance actuators for human-machine interfaces .............. 20
TactPiEd: Easy Prototyping of Tactile Patterns ................................................................. 21
Providing Localized Tactile Feedback on a Transparent Surface Through Time-Reversal Focusing ................................................................. 21
Ultrasonic piezoelectric actuators for compact spherical motors .......................................... 23
Server Assisted Key Establishment Protocol for WSN .......................................................... 24
Delay-Tolerant Space-Time Coding and Decoding for Femto-Assisted Cellular Networks ........................................................................ 25
Reliability for emergency applications in Internet of Things ............................................. 26
VII Six: VIN-based IPv6 Provider Independent Addressing for Future Vehicular Internet Communications ...................................................... 27
3D camera modeling and calibration .................................................................................. 28
Dense Image Matching: a General Framework Combining Direct and Feature-based Costs .............................................................................. 29
Human action segmentation using Deeply Optimized Hough Transform ............................ 30
Improving person detection performance using context ...................................................... 31
People reidentification in non overlapping cameras ........................................................... 32
Hierarchical scene model for Spatial-color Mixture Of Gaussians ....................................... 33

KNOWLEDGE ENGINEERING .............................................................................................. 34
Locality-constrained and Spatially Regularized Coding for Scene Categorization ............... 34
Identifying Bad Semantic Neighbors for Improving Distributional Thesauri ......................... 35
Space-Time Robust Video Representation for Action Recognition ..................................... 36
Building Specialized Bilingual Lexicons Using Large-Scale Background Knowledge ........ 37
CEA LIST’s Participation at MediaEval 2013 Placing Task .................................................. 38

EMBEDDED SYSTEMS ........................................................................................................ 39

COMPUTING ARCHITECTURES ............................................................................................ 40
Automatic deployment on embedded parallel systems .......................................................... 40
SNet, a flexible, Scalable NETwork paradigm for manycore architectures ......................... 41
Enhancing Cache Coherent Architectures with Access Patterns for Embedded Manycore Systems ........................................................................ 42

DESIGN & ANALYSIS .......................................................................................................... 43
An Optimization Approach for the Synthesis of AUTOSAR Architectures .......................... 43
Code generation for component-based applications .............................................................. 44
Design Refinement toward Implementation Methodology ...................................................... 45
A step-forward in formalizing UML: Precise Semantics of UML Composite Structures (PSCS) .............................................................................. 46
MARTE, a standard modeling language dedicated to CPS ..................................................... 47
Model-Driven Safety Assessment of Robotic Systems ............................................................ 48
Moka: an open and flexible framework for model execution in Papyrus ................................ 49
A new approach for real-time tasks synthesis and their placement on distributed architectures .............................................................................. 50
Sophia Framework for Model-Based Safety Analysis ............................................................ 51
An Implementation Relation and Test Framework for Timed Distributed Systems ............... 52
Guaranteed Integration of Ordinary Differential Equations using Affine Arithmetics ........... 53
A Homotopical Completion Procedure with Applications to Coherence of Monoids .......... 54
SECURED HW SW COMPONENTS

Time Reversal Reflectometry for Cable Aging Characterization .................................................................55
On a Useful Tool to Localize Jacks in Wiring Network ..................................................................................56

SENSORS INTEGRATION

Spiking Neural Network for Embedded image sensors ..........................................................57
Gesture Recognition On Smart Cameras ...............................................................................................58
A New Approach of Smart Vision Sensors .........................................................................................59

VALIDATION & VERIFICATION

Formal Verification of Software Important to Safety .........................................................................60
E-ACSL: An Executable Specification Language ......................................................................................61
Towards Verified Cloud Computing Environments .................................................................................62
Frame-C from a Software Analysis Perspective .......................................................................................64
Binary-Level Testing of Embedded Programs ........................................................................................65
GATE: a V&V platform for SCADE Models ...............................................................................................66
Static Analysis of Numerical Programs and Systems ...............................................................................67
Program Transformation for Information Flow Security ...........................................................................68
Results for compositional timed testing .............................................................................................69
Homomorphic Encryption in Cloud Computing .......................................................................................70
Time- and angle-triggered real-time kernel and its use for Powertrain applications ................................71
A Heuristic Algorithm for Stochastic Partitioning of Process Networks ...............................................72
Adapting Just-In-Time Compilation to Embedded Systems ..................................................................73
Dynamic code generation: Large Spectrum, Many Applications ..............................................................74

ADVANCED MANUFACTURING

Fast and Automatic City-Scale Environment Modeling for an Accurate 6DOF Vehicle Localization ..........76
An Interactive Augmented Reality System: a Prototype for Industrial Maintenance Training Applications .77

COLLABORATIVE ROBOTICS

Dry friction modeling in dynamic identification for robot manipulators: theory and experiments ................78
Assistance Tools for Generic Definition of ITER Maintenance Tasks and Scenarios in Advanced Supervisory Control Systems ...............................................................................................................79
Optimal design of compliant joint, grippers and actuator for miniature and portable robotic devices ....80
A Framework for the Classification of Dexterous Haptic Interfaces Based on the Identification of the Most Frequently Used Contact Areas .........................................................................................81
DOF-Decoupled Active Force Sensing (D-DAFS): A Human-Inspired Approach to Touch-based Localisation Tasks .............................................................................................................................82
High level functions for the intuitive use of an assistive robot ................................................................83
Flexible Robotics - Applications to Multiscale Manipulations ..................................................................85
Design of a Novel Long-Range Inflatable Robotic Arm: Manufacturing, Analytical and Numerical Evaluation ........................................................................................................................................86

INSTRUMENTATION

An adaptive smoother for counting measurement ...................................................................................88
ZnO nanowires as an effective luminescent sensing material for nitroaromatic derivatives ..........89
Data acquisition for underwater explosive detection ..............................................................................90
Active neutron interrogation based on a LINAC for nuclear waste packages characterization ............91
GAMPIX: A second-generation gamma camera .....................................................................................92
High temperature resistant Fiber Bragg GRATings for Sodium-cooled Fast Reactors continuous monitoring ........................................................................................................................................93

METROLOGY OF RADIOACTIVITY

Analytical calculation of atomic effects in allowed6- decays .................................................................94
Evaluation and publication of radionuclide decay data ...........................................................................95
Digital pulse processing and optimization of the front-end electronics for nuclear instrumentation ....96
Beta spectrometry with metallic magnetic calorimeters ..........................................................................97
Simulations of the LNHB manganese bath facility ...............................................................................98
NON-DESTRUCTIVE TESTING

Adaptive mesh image representation for X-ray tomographic reconstruction ................................................................. 99
Metal artefact suppression in CT ........................................................................................................................................... 100
Structural Noise Filtering In Ultrasonic Imaging By Decomposition Of The Time Reversal Operator .................................. 101
Ultrasonic Field Radiated by EMAT into Ferromagnetic Media .......................................................................................... 102
Simulation of Acoustic Emission Testing ........................................................................................................................... 103
Experimental study for the validation of CIVA ultrasonic testing (UT) predictions ............................................................... 104
Simulation of guided waves inspections based on a coupling of Modal an Finite Elements methods ...................................... 105
Maxwell’s equations in covariant form for simulating the inspection of a conductive slab of complex shape ............................ 106
The triple singularity: a semi-numerical model for the near-critical angle scattering ........................................................... 107
Head Waves Ray Tracing on Irregular Surfaces for TOFD ultrasonic inspection ................................................................. 108
Simulation of ultrasonic inspections of welds using a paraxial ray tracing method ............................................................... 109
Generic simulation of ultrasonic echoes from planar cracks by combining the Kirchhoff and GTD models ............................ 110
How to determine POD curves from simulation results with confidence? .............................................................................. 111
A Bayesian approach for the determination of POD curves from empirical data merged with simulation results ...................... 112
Surface Integral Equations for Electromagnetic Testing ...................................................................................................... 113
Modelling of ferrite cored coils with complex shape for electromagnetic testing ................................................................. 114

VIRTUAL REALITY

Dynamic Digital Human Model for ergonomic assessment based on human-like Behavior ...................................................... 115
Posture Optimization for Manipulation Tasks ......................................................................................................................... 116
Passive Hierarchical Control based on Wrench Bounds ......................................................................................................... 117
Humanoid Push Recovery Control in case of Multiple non-Coplanar Contacts .................................................................. 118

CONTROL OF RADIATIONS FOR HEALTH .......................................................................................................................... 120

METROLOGY OF DOSE

Comparison of absorbed-dose-to-water units in different beam sizes and beam energies between LNHB and PTB ................... 121
Absorbed dose to water standards established by water calorimetry at the LNE-LNHB for medium energy X-rays ...................... 122
Comparison of alanine dosimetry systems for absorbed-dose-to-water measurements in high-energy photon beams ................ 123
Characterisation of the X-ray tube spectra using semi-conductor detectors ........................................................................ 124

MODELING & SIMULATION FOR RADIOThERAPY & IMAGING ................................................................................................ 125

PenSSaRT, a new Monte Carlo system for quality control in radiotherapy ........................................................................... 125
Monte Carlo PENRADIO software for dose calculation in medical imaging ................................................................. 126
High-resolution portal image prediction for radiotherapy treatment verification ................................................................. 127
OVERVIEW

SMART DIGITAL SYSTEMS

Based in Paris region, CEA LIST Institute aims at developing research activity on smart digital systems. Building long-lasting relationships with its industrial partners, based on innovation and technology transfer, the institute maintains a high level upstream research through academic partnerships. The scientific independent agency AERES assessment conducted in 2013 has positively highlighted CEA LIST Institute qualities and specificities, tagging it as “a singular scientific institution at Germany’s Fraunhofer’s scope”.

AN ASSET FOR INDUSTRIAL COMPETITIVENESS

CEA LIST Institute is intended to support industrial competitiveness through innovation and technology transfers. The institute’s ability to create a perennial relationship with its industrial partners requires that its researchers understand and respond to their needs. Often initiated within R&D projects, mainly in business competitiveness clusters as Systematic, these links take shape in strategic partnerships. To maintain the necessary high level of research and innovation, CEA LIST has also forged a number of academic partnerships, particularly within RTRA Digiteo and the Carnot Institutes network. Currently, CEA LIST proactively participates to the construction of the new Paris-Saclay University. Such an environment enables CEA LIST to address future societal needs, including more user-friendly human-machine interfaces, more secure and reliable digital systems, intelligent energy management systems and tools for industrial eco-design processes that take into account environmental issues.

INTEGRATED PROJECTS: FROM IDEA TO INNOVATION

CEA LIST’s research covers several major sectors of activity surrounding our daily life, among which information and communication technologies (ICT), energy, transport, security, health and manufacturing. The Institute’s leading edge is based on its ability to integrate hardware and software using its know-how and technological skills. Designing quality comprehensive systems is made possible thanks to the institute’s strong culture on project building and to the excellence achieved by its 700 researchers, engineers and technicians. The CEA LIST research covers the entire innovation process, from the publication of original results to development of demonstrators and transfers of technology. Advanced manufacturing, ambient intelligence, embedded systems and radiotherapy technologies are the targeted areas the Institute gives priority.

AERES: Governmental Evaluation Agency for Research and Higher Education
AMBIENT INTELLIGENCE

SMART MANIPULATION
MULTI-AGENT SYSTEMS (MAS)
MODELING AND MANAGEMENT

RESEARCH TOPICS: MULTI-AGENT SYSTEMS FOR ENERGY SYSTEMS MANAGEMENT
B. LACROIX, C. PANIAH, J. GIL-QUIJANO, C. PAULUS (CEA LITEN), D. MERCIER
PARTNERSHIP: CEA-LITEN, SUPELEC, ALSTOM GRIDS

In the current context of energy transition towards distributed and renewable generation a radical change in the way in which energy (and particularly electricity) is generated, distributed and consumed is mandatory. Addressing this challenge requires future energy systems (such as the smart grid) to be capable of autonomously and intelligently configuring themselves to make the most efficient use of available resources, to be robust to the cascading failures that plague current networks, and to be extendable and adaptable in the face of rapidly changing technologies and requirements.

The distributed nature of these systems, and the autonomous behavior expected of them, naturally lend themselves to a multi-agent methodology. Since 5 years CEA LIST develops methodologies and tools for modeling, simulating and managing distributed energy systems based on multi-agent systems [1, 2]. Multi-agent systems properties such as flexibility, robustness to failures and distributed management have been proven. Former work [1] provides an adapted and generic multi-agent architecture for energy management (see Fig. 1). This architecture has been evaluated in the case of thermal management in buildings. The performances (costs of energy, comfort maintain) of the multi-agent management system are equivalent to best state of the art systems.

That former work is currently being extended to medium and large scale systems (eco-districts and wind farms coalitions) by including a multi-actor approach. This approach allows taking into account the diversity of individual constraints and objectives of the different actors (producers, consumers, system operators, etc.) in order to provide optimal decentralized management. As that type of management is based on distributed decision making processes, compared to centralized traditional management approaches, the information flows are reduced, the confidentiality of critical data and the autonomy of the actors are guaranteed.

The multi-actor approach has been initially implemented for management of coalitions of autonomous wind farms [2]. The approach is based on the concept of Cooperative Virtual Power Plants (CVPP). The objective is to gather and coordinate numerous independent producers and perform an efficient market participation of such coalition. A CVPP allows combining non-controllable sources (e.g. wind farms) with controllable sources (mass storage) as a single virtual aggregated producer. That aggregated producer is subject to the traditional rules of the spot market; we thus consider that there is no electricity purchasing obligation. In our initial work [2] we provide the multi-actor management architecture (see Fig. 2). To perform an efficient market participation of the CVPP, we are currently modelling its behavior as a Markov Decision Process. Our model takes into account renewable generation prior uncertainty, market constraints and optimizes sequentially the utilization of available resources. This current work shows very promising performances results of the management strategy.

References:
UNDERSTANDING AGEING OF ELECTRIC VEHICLE THROUGH STATISTICAL DATA ANALYSIS

A. Barré, Z. Younès, F. Suard

The main concern of electric vehicles for real use resides in the low energy autonomy. This limitation is explained by the technology of energy storage which can not currently store and provide enough electricity for long travels. On top of that, this range autonomy is decreasing over years due to the battery ageing.

Battery ageing can be characterized by two phenomena: raise of internal resistance and decrease of storage capacity [1]. These phenomena have been widely explained by physical modeling that could justify the interaction between ageing and a specific factor like the ambient temperature. However, during real use of a vehicle many factors intervene in a non-linear way, so that such physical approaches are limited to analyze and explain all interactions. We propose here to deal with real data coming from battery and vehicle measurement. Our approach consists in applying statistical tools to analyze all data and quantify the main factors that impact battery ageing. All results are obtained from a large dataset containing around three years of real use.

The impacting factors could be classified into three categories: environmental conditions, self-technical wear and user actions. These two first categories have been considered by analyzing measurement of battery only [2], in order to quantify all the variables impacting directly the battery ageing.

Thanks to statistical tools one can identify the variables that are linked with capacity decrease like the ambient temperature or storage duration whereas the resistance raising is generally linked with energy solicitation [3]. A specificity of this work is to provide not only numerical analysis of relevant variables, but also a visual tool to display the result. Such tool, like PCA (Principal Component Analysis) represented in Fig. 1 is able to clearly show to non-expert what are the main factors in an intelligible way. Since the battery ageing can be justified in a statistical way, the next step is to learn a model based only on data examples. Some algorithms, like linear regression or Relevance Vector Machine, have then been applied to predict the ageing of the battery by considering the novel values of all variables. The results have shown that the obtained precision complies with an objective of vehicle monitoring and management.

The third category is related to driver’s behavior. Data-based approaches seem to be more appropriate here, since no prior knowledge could be set up to describe and quantify the interaction of driver with vehicle ageing. Statistical tools have then been developed to extract features which describe the aggressiveness level of car driving considering only onboard measurement like speed or energy [4]. At first, the results have shown the relevance of main variables impacting the vehicle ageing or energy consumption. Then, these variables have been applied into a classifier that is able to define in real time the type of driving regarding the most recent historic of data. Figure 2 shows an example of online analysis of a vehicle driving that discriminates the ways considered as an aggressive driving.

References:
Photometric Stereo is an existing technique for object surface reconstruction from images based on analysis of the interaction of light on surfaces. In this work we aim at applying such techniques to microbial colonies on Petri dishes in a real-world configuration. To this aim we need to tackle the two main limitations of classical Photometric Stereo technique that are: need for prior knowledge on lightings that we often not have in real-life and hypothesis of Lambertian, that is diffuse, surfaces which is most of the time not true for real objects and particularly microbial colonies.

We proposed an original technique of uncalibrated (no prior on lightings) Photometric stereo for non-Lambertian objects. The proposed approach consists in two phases (see Fig. 1): first we correct images of the input sequence from specularities in order to obtain images of pseudo-Lambertian surfaces, and then realize Lambertian photometric stereo reconstruction. We apply the novel processing to Petri dish images for microbial colonies surface reconstruction.

Specularity is an optical effect of a complex physical nature which is useful for human 3D objects perception but affects automated image processing. In order to be able to apply the Lambertian photometric stereo model, specularities should be removed from the input images. We propose an original algorithm, which is able to detect specularities as abnormally elevated pixel intensity values in an image of the input sequence and to correct the found zones using information from all other images of the sequence and a specific continuous correcting function [1]. This method allows removing specularities while still preserving all other particularities of shading important for the further surface reconstruction.

We then propose an original stereo photometric method for Lambertian surface reconstruction with no prior on illuminations [2]. The implemented photometric stereo model consists of four components, two of them (albedo and normals) describe surface properties and the others (light sources intensities and directions) describe illumination. The proposed algorithm uses the alternating optimization principle. Each model component is found iteratively fixing all variables but one and applying value and quality constraints for the optimization function. The original scheme of resolution allows separating the different information types included in input images. Thanks to such matrix factorization, the surface reconstruction is made with no prior information on lighting directions and the reconstructed objects properties.

The proposed method was demonstrated on microbial colonies as shown in Fig. 2 but could also be applied to any object with at least 3 images with a fix camera position and different lightings with unknown position. It has been patented [3, 4].
PREDICTING METEOROLOGICAL VARIABLES WITH NEURAL NETWORKS

RESEARCH TOPICS: TIME SERIES PREDICTION, SOLAR ENERGY, MULTI-LAYER PERCEPTRON
L. BOUDET, P. BRUNEAU
PARTNERSHIP: CEA LITEN

This work proposes to study different neural architectures for the prediction of meteorological time series [1] and to select the most interesting inputs for predicting [2]. In the context of energy management system in buildings using solar energy, 24 hours ahead prediction of global solar irradiation and air temperature are performed using data from the last 24 hour measurements of 4 meteorological variables: air pressure, hygrometry and the two to be predicted. Pre-processing comprises downsampling data to tri-hourly data and normalization.

Multi-Layer Perceptrons (MLP’s) are very popular to predict meteorological time series. The standard approach consists in learning only one MLP which corresponds to a relative daily architecture given by the combination of options a) and c) in Fig. 1. Nevertheless, the daily cyclic nature of data is not well handled by this approach: either the complexity of the network learned is very high, or the output space has to be reduced.

Alternatively, we propose to decompose the prediction problem. MLP models are specialized to a current time slot (option b in Fig. 1) or a predicting horizon (option d in Fig. 1) and become building blocks of architecture of models for time series prediction. The number of neurons on the hidden layer is determined by cross-validation for each MLP during the learning stage.

Fig. 2 shows that neural architectures outperform the two reference models. Absolute architectures exhibit a slight gain in performance for short term prediction. They are composed of many more (8 to 64) but much simpler MLPs than the relative daily approach. Computational costs for learning and prediction are reduced by a factor of 2 to 8 (Fig. 3). The best architecture found is the absolute daily one according to this criterion.

Furthermore, a Bayesian variable selection procedure has been applied in order to select the most salient input variables [2]: 34% to 73% of inputs have been selected that gather at least 95% of total saliency. Prediction quality loss is very small: 1.3% on average (5% in the worst case). Computational costs have been reduced after Bayesian selection even if the number of neurons retained on the hidden layer has increased. For each of the two predicted variables, air pressure is the most important variable and the variable itself is the second one for its own prediction. Hygrometry shows only few utility (0 to 2 slots retained).

References:
FROM SPARSE CODING TO PREDICTIVE AND REPRESENTATION LEARNING: A SPIKE-BASED APPROACH

RESEARCH TOPICS: MULTIVARIATE SPARSE DECOMPOSITIONS, SPIKE-BASED METRIC, CLASSIFICATION, CLUSTERING

ANTHONY MOURAUD, QUENTIN BARTHÉLEMY, CÉDRIC GOUY-PAILLER, AURÉLIEN MAYOUE, ANTHONY LARUE, HÉLÈNE PAUGAM-MOISY (LIRIS, CNRS, UNIV. LYON 2)
PARTNERSHIP: LRI, UNIV. PARIS XI

Sparse decompositions of multivariate signals are keys to efficient compression, storage and denoising, but they lack appropriate methods to exploit the sparsity for automated predictive and representation learning. For example, experiments on multivariate input signals from 2D natural handwriting data show that most of the information about the input signal is contained in the label and pulse-timing of the sparse coded signals. Figure 1 represents two-dimensional velocity signals recorded while a user is handwriting on a tablet. Sparse decompositions amount to successively identifying inside the raw signals occurrences of kernels from a pre-defined or learnt dictionary. They are graphically depicted through spikegrams, which encode the appearance of kernels versus time to account for raw signals. The importance of each kernel at a certain time is graphically coded using a colour scale (Fig. 1).

Multineuronal spike trains are actively studied by computational neuroscientists who search to understand neural coding through convenient similarity measures and comparisons of recorded spike trains. The present work addresses a promising approach to multivariate signals classification through the coupling of sparse coding processes and neuro-computational spike trains analysis methods. The proposed methodology consists of three steps. First, a sparse coding step converts multivariate signals into multi-source pulse trains. This step entails the use of alternating algorithms to learn adapted representations of the signals (Multivariate Dictionary Learning) and optimal projections on the resulting dictionary (Multivariate Orthogonal Matching Pursuit).

Resulting pulse trains stand for the parsimonious representation of input signals. Second, similarities between sparse pulse trains are computed with a metric-space analysis, considering only the time of occurrence and source labels. Third, a stimulus based clustering is computed from distances between pulse trains in order to assess the efficiency of the coding method and the ability of the metric to retrieve original classes. The multivariate spike-based distance is assessing proximity between two spike trains by quantitatively evaluating spikes swap, translation and relabelling to transform one spike train into one another (see Fig. 2).

Based on this spike-based metric, various applications have been considered to address prediction and representation learning from sparse decompositions. In mass spectrometry, Support Vector Machine has been plugged in the process with a custom kernel built from the spike-based distance between samples. In automatic handwriting recognition, such metrics have been involved in a representation learning task (clustering) to assess separability of handwritten letters from their sparse representations (Fig. 3).

References:
A FUZZY SPATIO-TEMPORAL APPROACH FOR ACTIVITY RECOGNITION

RESEARCH TOPICS: FUZZY EXPERT SYSTEM, SPATIO-TEMPORAL OPERATORS
LAURENCE BOUDET, JEAN-PHILIPPE POLI, JEAN-MARIE LE YAOUANC
PARTNERSHIP: EGIDIUM

Over the last decade, there has been a significant deployment of systems dedicated to surveillance. These systems make use of real-time sensors that generate continuous streams of data. Despite their success in many cases, the increased number of sensors leads to a cognitive overload for the operator in charge of their analysis. However, the context and the application require an ability to react in real time. The research presented in this paper introduces a spatio-temporal-based approach that aims to provide a qualitative interpretation of the behavior of an entity (e.g., a human or vehicle).

In order to consider the uncertainties of the activities’ structures, the semantic interpretation of the entity’s trajectory is supported by a fuzzy-based approach. It is designed to be as generic as possible, and considers objects with rather bona fide or fiat boundaries. The former are objects with physical discontinuities (e.g. a mountain or a valley), the latter gets boundaries induced through human demarcation (e.g. a building or an administrative region). The development of formal models of topological relations has received much attention in the literature of GIS, computer vision and image understanding. In this page, we describe some of the spatio-temporal operators we have developed.

The spatio-temporal relation IsMoving characterizes the moving of a mobile entity in a non-constraint space. Its evaluation takes into account the positions of the considered entity during a past time interval. It is based on the assumption that the value at time \( t_1 \) may not only be based on the past move between \( t_{i-1} \) and \( t_i \), but on their recording in the past. Consequently, if the entity \( e \) is not moving between times \( t_{i-1} \) and \( t_i \), the value of the relation IsMoving is pondered by its previous moving during a given past time interval. In other words, if a pedestrian stops walking at time \( t \) because he is looking for his keys, the value of the relation IsMoving will decrease in time if he stops during a significant time.

The spatio-temporal relation IsComingCloseTo characterizes the approach of a spatial object \( o \) by a mobile entity \( e \) in a non-constraint space. For instance, this relation may be useful for characterizing a boat that is coming close to a navigational buoy. Intuitively, the closer the entity to \( o \) in a way that minimizes the distance to reach the object, the higher the fuzzy value of IsComingCloseTo. Two spatial configurations are defined, with respect to the geometry of \( o \) (either a point or an open polyline), as shown in Fig. 1. The evaluation of this operator takes into account:

- the evolution of the location of \( e \) between times \( t_i \) and \( t_{i-1} \);
- the orientation \( \alpha(t) \) of \( e \) at time \( t_i \); thus, the more \( \cos(\alpha(t)) \) tends to 1, the higher the value of IsComingCloseTo;
- the recording of orientations \( \alpha(t) \) during a temporal interval.

In the case we want to apply the IsComingCloseTo operator to an object which is now a closed object, the formulation is quite different, as shown in Fig. 2. The evaluation takes into account:

- the evolution of the location of \( e \) between times \( t_i \) and \( t_{i-1} \);
- the location of \( e \) relatively to \( o \);
- the orientations \( \alpha(t) \) and \( \beta(t) \) of \( e \) at time \( t_i \); thus, the more \( \cos(\max(\alpha(t),\beta(t))) \) tends to 1, the higher the value of IsComingCloseTo;
- the recording of orientations \( \alpha(t) \) during a temporal interval.

We have designed several spatio-temporal relations that relate an entity to an object of the environment. Further theoretical work concerns an extension of the ontological background of the approach and the development of complementary spatio-temporal relations, e.g., IsGoingThrough, IsEntering, IsGoingOut and IsFollowingARoute.

References:
ANNUAL SCIENTIFIC REPORT 2012-2013

SINGLE CRYSTAL DIAMOND DOSIMETER FOR DOSE MEASUREMENT IN SMALL BEAMS USED IN STEREOTACTIC RADIOTHERAPY

RESEARCH TOPICS: DIAMOND DOSIMETER, STEREOTACTIC RADIOTHERAPY


SPONSORSHIP: ANR TECSAN DIADOMI

Recent developments of new therapy techniques using small photon beams, such as stereotactic radiotherapy, require detectors to determine the delivered dose with a high accuracy. Lack of lateral electronic equilibrium in small beam causes errors in the measured dose with inaccurate device, particularly for lateral dose profile, depth dose curve and output factor (OF) measurements. The main goal of this project was to develop a diamond dosimeter which is as close as possible to tissue equivalence and exhibits a small detection volume (0.15 mm³) compared to the size of the irradiation field.

Simulation using PENELOPE Monte Carlo Code was performed in order to optimize materials used to fabricate the dosimeter. Aluminium or DLC (Diamond Like Carbon) electrodes were selected as electrode material as well as P8nMA and PMMA material for encapsulation. Graphite resin was used to take electrical contact on the diamond surface [1].

Following this first optimisation, final geometry of the dosimeter was defined. Comparison of simulation results and clinical measurements was performed to give evidence of the influence of each parameter (diamond thickness and size, electrode geometry, height of encapsulation material) on the dose measurement. The resulting optimized diamond dosimeter is shown in Fig. 1. A diamond 150 µm thick was used with 1 mm × 1 mm lateral size. Completely covering electrodes were deposited on each face. A minimized 1.1 mm height of P8nMA was used to encapsulate the upper part of the diamond.

Single Crystal Diamond Dosimeter (SCDDo) was tested in a 10×10 cm² beam produced by a Varian Clinac 2100 C accelerator at La Pitié Salpêtrière Hospital. SCDDo exhibits a high sensitivity of 44.5 nC.Gy⁻¹ and an excellent repeatability of 0.06%. The dose linearity of the SCDDo response was verified with 6 MV photon beam for a large dose range and a low dose rate dependence less than 1% was observed. Finally, a low energy dependence of 1.2% was observed between 6 MV and 18 MV beam quality [2].

Output factors measured with the SCDDo for field sizes smaller than 1.8×1.8 cm² are higher than those measured with the PTW 31014 PinPoint ionization chamber which is well known to underestimate the OF values in small beams. Further the difference with µLif (here considered as our passive reference dosimeter) is only 3.1% showing the real interest of this type of device for OF measurement with small beams. The SCDDo could also find a real use in high photon flux as also in proton beams.

References:
SURFACE TRANSFER DOPING CAN MEDIATE BOTH COLLOIDAL STABILITY AND SELF-ASSEMBLY OF NANODIAMONDS

RESEARCH TOPICS: NANODIAMONDS, PLASMA HYDROGENATION, SELF-ASSEMBLY
T. PETIT, H. A. GIRARD, A. TROUVÉ, I. BATONNEAU-GENER (IC2M, POITIERS), P. BERGONZO, J. C. ARNAULT

Hydrogenated detonation nanodiamonds (H-NDs) were prepared by microwave-enhanced plasma treatments. Plasma hydrogenation has been shown to be the most adapted method to efficiently passivate the ND surface with hydrogen atoms while etching non-diamond carbon and oxygen species. In a previous study, we demonstrated that vanishing of the oxygenated terminations and the presence of C–H bonds after such plasma treatments occur, evidenced by FTIR and XPS analysis [2]. Furthermore, hydrogenation of NDs was also confirmed by the validation on our NDs-H of grafting routes specific to hydrogen terminated diamond surfaces [3].

Hydrogenated NDs exhibit a diamond core surrounded by stable C-H terminations. A high affinity of H-NDs toward water molecules was demonstrated by adsorption isotherms (BET) with more hydrophilic sites compared to NDs-COOH (Fig. 1).

As a consequence, stable H-NDs suspensions were obtained in water exhibiting a positive Zeta potential (ZP), +45 mV at pH = 7.4 (Fig. 2). The iso-electrical point (IEP), i.e. the point of zero charge, is located at pH = 12.4. The origin of the positive Zeta potential was related to a transfer doping occurring onto 5 nm diamond nanoparticles based on the diamond semi-conductive behavior. Indeed, a similar surface phenomenon was observed on hydrogenated bulk diamond material occurring after air exposure or in water.

Aggregates formed by HCl addition in H-NDs suspended in water were observed by Scanning Electron Microscopy (Fig. 3A–C). At low concentrations, aggregates have spherical shapes (Fig. 3A). Surprisingly, cubic structures with smooth surfaces appear at higher concentrations (Fig. 3B-C), with a high reproducibility on several samples. Organized structures with sizes ranging up to the millimeter scale resulting from NDs-H/Cl- ion coupling were even observed in solution at a HCl concentration of 80 mM (Fig. 3D). Aggregation is most probably related to electrostatic interactions between Cl- counterions and NDs-H.

References:
**BOOSTING THE ELECTROCHEMICAL PROPERTIES OF DIAMOND ELECTRODES USING CARBON NANOTUBE SCAFFOLDS**

**RESEARCH TOPICS:** NANO DIAMONDS, PLASMA HYDROGENATION, SELF-ASSEMBLY

**CLEMENT HÉBERT, LIONEL ROUSSEAU¹, JEAN-PAUL MAZELLIER², EMMANUEL SCORSONE, MICHEL MERMOUX³, PHILIPPE BERGONZO**

**SPONSORSHIP:** FP7 NEUROCARE (2010-2013)

**PARTNERSHIP:** ¹ ESIEE GROUP PARIS, ² THALES TRT, ³ LEPMI GRENOBLE, INSERM

Boron doped diamond (BDD) electrodes exhibit a number of superior electrochemical properties over more conventional electrodes: they feature a low background current resulting from an intrinsically low double layer capacitance, thus enabling high signal to noise electrochemical measurements. Furthermore, their wide potential window in aqueous media, typically above 3 V, offers the possibility to detect a broad range of analytes that could not be addressed with other types of electrodes. Finally, in addition to their antifouling capabilities and long term stability, Kiran et al. have demonstrated that they can be electrochemically cleaned or “reactivated” in situ in biological fluids thus opening the way to continuous monitoring in situ [1]. Due to the combination of those exceptional properties, BDD based electrochemical sensors have been reported e.g. for the detection of glucose, catecholamine, norepinephrine etc.

More recently, BDD electrodes have also been identified as a promising material for implantable stimulation electrodes. Such electrodes are used to generate electric field gradients that produce a compensating flux of charges toward the neurons and tissues surrounding the electrodes. Several studies have indicated that diamond is both biocompatible and chemically inert, which are two prerequisites for long-term tissue implantation. Nevertheless, even though the electrode must be capable of injecting enough charges to fire neurons for decades without degradation of the tissue nor of the electrode, and excess voltage limits raise the pH and degrade the tissues. BDD electrodes suffer in this case from a low double layer capacitance and a high impedance which limits the amount of injectable charges.

We propose an alternative approach where the diamond surface is structured in order to increase its specific surface area. We take advantage of the use of a diamond coated carbon nanotube bundles as a new electrode material combining low impedance, high capacitance, high stability and fairly wide potential window in water. The material structure starts from a very dense vertically aligned multiwalled carbon nanotubes (VACNTs) array, on which diamond is grown using optimised conditions to obtain a uniform coating enabling diamond electrochemical performances to be obtained, and taking advantage of the VACNTs large area (Fig. 1).

As shown in Fig. 2, the performances of these structures are astonishingly superior to flat BDD materials: in optimised geometrical conditions, the capacitance is increased by a factor 116, and the impedance reduced by a factor 100 [2].

We believe with such advances towards a nano-textured material at the hundreds of nanometre scale, such progress opens up the route for diamond to match the performances of other materials, while it surpasses them with its extreme stability and biocompatibility [3].

![Figure 1. Cross-section of 3D-shaped diamond electrodes.](image1)

![Figure 2. Improvement of 3D BDD electrodes in terms of capacitance (top) and impedance (bottom).](image2)

**References:**


ODORANT BINDING PROTEINS ON DIAMOND BIO-MEMS

RESEARCH TOPICS: MICROSENSORS, ODORANT BINDING PROTEINS, SYNTHETIC DIAMOND


SPONSORSHIP: EU-FP7 SNIFFER PROJECT

Odorant binding proteins (OBPs) are small soluble proteins found in olfactory systems that are capable of binding several types of odorant molecules. Beside, cantilevers based on polycrystalline diamond surfaces are very promising as chemical transducers [1,2]. Thus two methods were investigated here for chemically grafting porcine OBPs onto polycrystalline diamond microcantilevers for the development of biosensors [3]. The first method results in random orientation of the immobilized proteins over the surface (Fig. 1a). By contrast, the second approach based on chelating a histidine-tag located on the protein with nickel allows control of the proteins orientation (Fig. 1b).

These grafting methods were validated using electrochemical impedance spectroscopy, fluorescence imaging and X-ray photoelectron spectroscopy analysis. For instance Fig. 2 reveals the presence of patterned OBP immobilized on a diamond surface when the proteins are binding to a fluorescent tag.

The sensing performances of the two resulting sensors were compared for IBMP (Fig. 3a), a reference compound commonly used for the characterization of OBPs affinity constants, as well as for 2,4-dinitrotoluene (2,4-DNT; Fig. 3b), used as a surrogate for the explosive trinitrotoluene (TNT).

The second grafting method led to typically 20 % more sensitive sensors, as a result of better access of ligands to the proteins active sites and also perhaps a better yield of protein immobilization. This new grafting method appears as highly promising for further investigation of the ligand binding properties of OBPs in general and for the development of non-specific biosensors arrays for artificial olfaction applications.

Figure 2. Summary of OBP immobilization on hydrogenated diamond surfaces using two strategies.

Figure 3. Calibration curves of cantilever grafted with OBP with the 2 methods in response (a) to IBMP (inset: typical response kinetic to a given conc. of IBMP), and (b) to 2,4-DNT.

References:
OPTICAL ANALYSIS OF P-TYPE SURFACE CONDUCTIVITY IN DIAMOND WITH SLOTTED PHOTONIC CRYSTALS

RESEARCH TOPICS: PHOTONIC CRYSTALS, HYDROGENATED DIAMOND
C. BLIN, X. CHECOURY1, H.A. GIRARD, C. GESSET, S. SAADA, P. BOUCAUD1, P. BERGONZO 1: IEF, CNRS UNIV. PARIS SUD
SPONSORSHIP: CHAIRE MIXTE CEA-UNIVERSITÉ PARIS SUD
PARTNERSHIP: IEF, UNIVERSITÉ PARIS SUD

Using integrated photonic technologies, it is possible to fabricate very compact, high performing and low-cost chemical and biochemical sensors. A novel class of nanoscale biosensors which have been increasingly studied in recent years is photonic crystals (PhCs). Such devices can offer unique and high sensitivity for the detection of specific chemical and biological agents using small active volumes, thus making them promising structures for their use in medical or environmental applications. Due to the tight confinement of the optical mode, PhCs are highly sensitive to local changes in their environment, such as the refractive index, which affect the resonance wavelength and the quality factor of the resonator. This shift of the resonance wavelength thus confers a sensing capacity to the cavity. Specificity in detection is brought by functionalizing the surface of the resonant cavity or waveguide, which can then be used to selectively capture a specific target molecule, in gas or liquid phase.

We report on the fabrication of two-dimensional slotted diamond-based photonic crystals (see Fig. 1 and Fig. 2) with Q factors up to 6500 and their optical characterization at 1550 nm in order to probe surface molecular modifications. In this study [2], we intentionally focus on the simplest surface modifications that can modify the diamond PhC optical properties, namely, hydrogenation and oxidation. We observed that, depending on the chemical surface termination, these diamond PhCs exhibit a strong modification of their spectral features. When the surface is tuned from oxidized to hydrogenated, a resonance wavelength shift of the cavity occurs and is accompanied by a decrease of the Q factor. Moreover, we give experimental evidence that this phenomenon is reversible as the initial value of the Q factor is recovered when the surface is re-oxidized. We attribute this result to the sub-surface conductive layer that is due to transfer doping in hydrogenated diamond and which is absent from oxidized diamond. Thanks to 3D-FDTD simulations, we give an estimate of the effective refractive index of the surface conductive layer at 1.5 µm as a function of its thickness. This result highlights the high sensitivity of slotted diamond PhC and the importance of surface control for biosensing with diamond.

Figure 1. Scanning electron microscope (SEM) views of the slotted photonic crystal in polycrystalline diamond with the 135-nm wide slot and its access waveguide.

Figure 2. Optical image of the cavity mode resonating at 1621.8 nm. Scale bar: 1 µm (false colors).

References:
BELT MOUNTED IMU WITH ENHANCED DISTANCE ESTIMATION FOR PEDESTRIAN INDOOR POSITIONING

RESEARCH TOPICS: MOBILE POSITIONING IN CHALLENGING ENVIRONMENTS WITH HYBRID INERTIAL AND RADIO MEASUREMENTS.
A.PATAROT, M.BOUKALLEL, S. LAMY-PERBAL, A.VERVISCH-PICOIS, N SAMAMA
PARTNERSHIP: INSTITUT MINES-TELECOM (FRANCE)

We are experiencing an increasing interest for indoor pedestrian localization, which has proven to be a significant source of information for applications in the fields of protection, navigation and assistance. The rapid growth of location based services is motivated by recent technological progress in mobile device design and by their added value for urban mobility. Indoor pedestrian localization has been initially addressed in the literature with pre-installed infrastructure solutions such as GNSS, UWB, WiFi Bluetooth or RFID. Even if these solutions are smartphone-oriented with potential sub-meter accuracies, they require infrastructures to be deployed and maintained in the environment with a reliable calibration process to ensure accurate positioning. The cost and constraints introduced by these approaches may limit usability for some applications. Recent research efforts have been put into the development of infrastructure-less solutions including inertia, camera and magnetometry. The inertial navigation systems have been extensively studied lately thanks to their potential compacity, their commercial availability, and their achieved localization accuracy [1].

Inertial solutions require angular rates from gyroimeters and specific forces from accelerometers, in order to compute a position relative to a known starting location and attitude. Without effort to mitigate the biases of micro electromechanical sensors, the estimation of position is strongly inaccurate over time. Therefore, a periodic update of speed and attitude can be introduced, when the inertial sensor is assumed to be motionless, to reduce the drift of position over time. The motionless periods or step detections are more accurate with thresholds over a wide dynamic range when the sensors are foot-mounted [2]. However, the limitation of a device on the shoe and the consideration of thresholds are only suitable for a given set of sensors and pedestrians. Our belt-mounted approach attempts to address issues of both the placement and calibration without step detection or map information, while maintaining the mitigation of drifts in distance and orientation. Our development carefully deals with the mobility and the computation cost. The resulting position is computed and displayed to the user in real-time on a smartphone or on a geographical information system without any specific application installed.

The technique, involving the norm of the acceleration jerk and a regression on attitude, is developed with the analytical material necessary to design a relevant experimental methodology adapted to sensor specifications [3]. Dozens of repeatable experiments in real situations support the potential of this approach with a median error on distance of 3% of the traveled distance about 200m, which is similar to a performance of 5% for foot-mounted systems without their integration and tuning constraints. Future works will involve the hybridization with radio positioning systems to facilitate the calibration of this infrastructure-less inertial belt-mounted system.

References:
INFLUENCE OF INTERFACIAL ADHESION ON THE MECHANICAL RESPONSE OF MAGNETO-RHEOLOGICAL ELASTOMERS AT HIGH STRAIN

RESEARCH TOPICS: MAGNETO-RHEOLOGICAL ELASTOMER, INTERFACIAL ADHESION, TACTILE SURFACE

T. POSSINGER, C. BOLZMACHER, L. BODELOT, N. TRIANTAFYLLIDIS

LABORATORY: CEA LIST SENSORIAL AND AMBIENT INTERFACES LABORATORY

PARTNERSHIP: ÉCOLE POLYTECHNIQUE (PALAISEAU), LABORATOIRE DE MÉCANIQUE DES SOLIDES

Materials whose rheological properties can be varied by the application of magnetic fields belong to the specific class of smart materials because they can respond rapidly and reversibly to changes in their environment. As a branch of this kind of materials, so-called magneto-rheological elastomers (MREs) are typically composed of micron-sized magnetic particles dispersed in an elastomeric matrix. When a magnetic field is applied to the elastomer composite during crosslinking, particle chain structures form and remain locked in place upon final curing, thus imparting anisotropic properties to a viscoelastic material (Carlson and Jolly 2000; Kallio 2005).

Not only do MREs alter their viscoelastic properties in response to an external magnetic field but they also can undergo very high deformation states. While the former effect is mainly exploited in variable and controllable stiffness devices (Li and Zhang 2008), the latter may be of interest in haptic devices such as a tactile interface (Vidal-Verdu and Hafez 2007). Indeed, a large out-of-plane deformation can be introduced in response to a spatially localized magnetic field. Using a matrix of solenoids or permanent magnets placed underneath a soft MRE surface, patterns can be displayed on the surface (see Fig. 1) as already achieved with magneto-rheological fluids (Bolzmacher et al. 2011).

In the perspective of developing a persistent tactile MRE surface with reversible and large out-of-plane deformation, this work investigates the interfacial adhesion between the iron fillers and the silicone matrix at high strain. For this purpose, different MRE samples between 3.5% and 30% particle volume fraction of aligned carbonyl iron powder (CIP) were prepared with and without modification of the surface of the iron particles by the application of a silane primer following two different procedures.

Scanning electron microscope (SEM) images of simulated debonding at high strain showed that a more elaborated particle-matrix interface could be obtained with the primer additive. The primer treatment adds another challenging process step to the fabrication. The application of the primer by stirring the particles in the primer dilution lead to an enhanced stiffness. However, it has the disadvantage of forming large particle agglomerates entrapped in the matrix, which is a situation that should be avoided.

In the case of MRE, the best procedure to apply the primer is to spray the primer dilution in a thin layer on the particles. By this treatment, the primer could be applied homogeneously on the particles while avoiding particle aggregates among the aligned chains.

Remarkable stiffening could be observed past a critical strain threshold that increases as the volume fraction of the particles decreases (> 80% deformation in the case of the 30% volume fraction MREs, see Fig. 2). Beyond that critical strain threshold, strong bonds due to the primer application start to stiffen the composite material and seem to prevent debonding of the particles from the matrix.

References:
Haptic devices have been proven to be an effective contribution to improve human interaction with various types of virtual or tele-operated environments. Usually, such devices provide force feedback to the user using electric motors, which are known to cover only a limited range of haptic sensations. For instance, the maximum simulated stiffness by a motor is limited in the controllers in order to avoid instability. Although several methods exist to deal with these issues, new actuation techniques are needed in order to enhance the quality of the haptic rendering.

Within the context of human-machine interfaces, we have developed a series of novel actuators with improved haptic feedback quality (transparency), high torque, and low power consumption in a reduced physical size. Our approach relies on the combination of brakes and motors, from the design to optimization, integration and control. While brakes can provide high passive torque, the motor can restore energy to the operator allowing the actuators to cover a wide range of haptic feedback rendering.

The brakes, shown in Fig. 1 (felt), are based on a magnetorheological (MR) fluid. MR fluids are composed of a suspension of ferromagnetic micron-sized particles distributed in a non-magnetic carrier liquid. The fluid is placed between several shearing cylinders. Under the application of a controllable magnetic field across these cylinders by a coil, the fluid particles form chain-like structure or aggregates. This phenomenon is macroscopically manifested as an alteration of the fluid’s apparent viscosity. Consequently, the resistance against the relative angular displacement of the cylinders can be controlled through the current of the coil in order to provide a specific torque.

A new approach to design the brakes has been developed [1-2]. The magnetic circuit is designed and optimized to provide and focus a specific magnetic flux over the fluid gaps across the above mentioned shearing cylinders. Thus, inherent saturations and nonlinearities introduced by the fluid are avoided, and the power consumption is minimized. The brakes have 60 mm diameter, 40 mm width and can apply 5.3 Nm at 20 Watts [2]. This represents 23% more torque in a volume 76% smaller compared to a commercial MR brake (Lord Corporation RD2078).

Another novelty introduced by our work is the dual unidirectional brake approach. Each brake is connected to an overrunning clutch, thus it can impose a torque only in one defined rotational direction, allowing the controller to be implemented without any measurement of the use’s behavior [3].

Fig. 1 (right) shows the proposed hybrid actuator which is composed of two identical brakes shown on the left. The brakes are connected to a DC motor (Maxon RE40) through a common shaft. Since brakes are intrinsically stable in haptic interactions, the proposed brake-motor configuration combined with its specific control laws [4] has been shown to allow for the simulation of stiff virtual environments (up to 1.7 kN/rad) without the instability issues observed when using the motor only, and with higher transparency. Compared to the motor, for instance, the hybrid actuator has a torque density 8 times higher. When compared to a torque-equivalent motor-captor transmission system, our actuator presents 240 times less inertia and generates 5 times less friction.

We have also developed the associated control laws for stable interaction, which are able to calculate the gains of each actuator in real time. The gains sent to the motor are bounded in order to respect a specific stability criterion and the brake compensates for the difference in haptic interaction forces [4].

The proposed actuation system is completely independent of the virtual environment model and of the application, allowing the implementation of the actuator in any existent haptic device.

References:
TACTIPED: EASY PROTOTYPING OF TACTILE PATTERNS

RESEARCH TOPICS: TOOL FOR THE FACILITATION OF TACTILE PATTERN AUTHORING; CROSS-DEVICE METAPHOR DESIGN
S. PANÉELS, M. ANASTASSOVA, L. BRUNET
SPONSORSHIP: ENTRANCE PROJECT?
PARTNERSHIP: CEA LIST

The development of wearable tactile devices, in the forms of bracelets, vests or belts, has been increasing, in particular for mobile and wearable computing. By providing an alternative channel for communication, these devices enable the transmission of information in an eyes-free, ear-free and discrete manner, useful not only in contexts where other modalities are not available but also when they are heavily used, all the while letting the user focus on the surrounding environment. The haptic modality has successfully been used for example for providing navigation cues, status information or for enhancing the education and learning, among many other existing applications.

However, a lot of effort is still spent on designing appropriate, intuitive and discriminable vibrotactile cues, through numerous development iterations and user evaluations, often specific to one device and one application, and usually requiring specific device and programming knowledge. Therefore few research prototypes reach the market or the industry. In fact, the general public is only confronted to basic haptic interaction (e.g. simple vibrations from mobile phones) whereas interface design has been giving increasing importance to user experience. Solutions to integrate user experience in design frameworks are currently being investigated. One such solution lies in making the prototyping of interactions accessible to all, thus enabling a greater involvement of the user, for example using participatory design. This could potentially foster the creativity and intuitiveness.

Therefore, to ease the design and testing, and to help promote tactile research and make it more widely accessible, an interface is needed that enables the rapid prototyping and the easy authoring or tuning of patterns, not only by developers but also non-developers, such as ergonomists, designers and users, without the need for specific signal and computing knowledge. Such an interface should also support several types of devices including devices with several actuators. Few such interfaces are currently available and widely used. In order to fill this gap, we have developed TactiPED, a novel tactile editor fulfilling these requirements (see Figure 1). It is based on the graphical metaphor of the shape of the device, which is used for tuning the main tactile characteristics that are common knowledge, namely activation, amplitude, frequency and duration. The editor includes file systems functionalities using the XML format along with playing and recording functionalities.

The editor was thoroughly evaluated. First, it has successfully been used with a number of custom-developed devices, including a prototype bracelet with 6 electromagnetic actuators and a gamepad with 8, and a bracelet with 8 coin motors. It also enabled to reproduce patterns described in the literature. The interface can potentially be extended to support any device by providing the corresponding communication protocol and device template. Second, it was also successfully used by 9 participants (3 haptic ergonomists, 3 haptic engineers and 3 non-experts) during a usability evaluation that resulted in positive feedback about satisfaction and ease of use, all the while highlighting several areas for improvements (e.g. enabling comparison of patterns). Therefore, TactiPED was easily used with little training and enabled users to design patterns in little time (less than 5 min) no matter their level of expertise. Details about the results can be found in [1]. The resulting patterns shared common characteristics across the devices for a given metaphor, thus providing positive initial insights about the possibility of designing patterns that are transferrable from one device to another (see Figure 2).

Figure 1: Screenshot of the interface. In the central area, each line represents a sequence where the amplitude of each actuator can be set directly on the shape of the device, whereas the duration is set through a timeline.

Figure 2: Examples of the most representative patterns for each metaphor and each device. The grey blocks indicate a pause whereas the black lines indicate a change of actuator with the same amplitude. The colours correspond to the actual values in the interface.

References:
PROVIDING LOCALIZED TACTILE FEEDBACK ON A TRANSPARENT SURFACE THROUGH TIME-REVERSAL FOCUSING

RESEARCH TOPICS: TACTILE DISPLAY, SURFACE HAPTICS, WAVE FOCUSING
C. HUDIN, J. LOZADA, V. HAYWARD (ISIR)

The sense of touch is an effective communication channel through which we sense simultaneously with multiple fingers the temporal and spatial variations of a surface properties. Hence, to fully exploit the possibilities offered by the sense of touch in human computer interaction, a tactile feedback device must be able to produce a stimulation varying in time and space independently for each finger in contact with the surface.

The standard approach consists in using matrix of actuators, each dedicated to the stimulation in one point of space. These devices are however complex, cumbersome and opaque, due to the presence of actuators at each stimulation position. They do not lend themselves to the superposition on top of a screen.

To preserve the optical transparency of the tactile feedback interface, another approach consists in using remote actuators and take advantage of the transportation of vibration energy by mechanical waves. This principle is used in most consumer electronic devices that employ dc or piezoelectric motors to provide a tactile feedback on a screen. More elaborate implementations of this approach make use of ultrasonic vibrations to create a squeeze film or surface acoustic waves and thus modify the apparent friction coefficient between the finger and a surface.

Because of the uniform distribution of vibrations over the surface, information displayed with these approaches only vary in time. Although correlating the overall vibration of the surface with the finger position might give the illusion of a spatial variations, this illusion breaks as soon as multiple finger explore the surface.

After Gavrilov demonstrated that focused ultrasound propagating in water could evoke tactile sensations, Shinoda and other teams exploited the interference between ultrasonic acoustic waves emanating from an array of transducers to produce a tactile stimulation localized spatially in a volume of air.

The purpose of the work presented here is to transpose the idea of focused ultrasound waves to the case of a transparent surface, and thus to achieve localized tactile stimulation on a surface that can be stacked over a screen. Among the many types of mechanical waves that propagate in a solid volume, we consider the case of flexural waves in thin plates since they induce displacements out of the surface plane susceptible to be tactually perceived.

To the difference of acoustic waves propagating in free space, a thin plate is a bounded medium where waves reverberate. In such media, focusing cannot be achieved by driving the transducers with delayed harmonic signals. The dispersive nature of the flexural waves propagating in thin plates is another difference with the propagation in air. This property induces a frequency dependent propagation velocity and leads to the spreading of wave packets during propagation. The simple delayed emission of pulsed flexural waves would thus not yield the focusing in a single spot. The time-reversal technique, enables the focusing of any kind of waves in complex media and allows one to overcome these two difficulties.

Thus, we explored in this work the possibility to use the time-reversal technique to focus flexural waves propagating in a thin transparent plate. We investigated the outcomes of design choices on the performance of a multitouch and transparent tactile feedback interface. A mock-up device was built using 32 actuators bounded to the sides of a 0.5 mm thick glass plate. Focusing was successfully achieved with this device and produced a 5 mm wide spot with amplitude over 7 µm for a power consumption of 45 mW. Focusing simultaneously in multiple points was also achieved and demonstrates the feasibility of displaying multitouch tactile stimulation. The mechanism underlying the tactile detection of the peak displacement produced at focus point was finally explored.

With this demonstrator we achieved focusing with an amplitude greater than the tactile sensibility threshold within an area of the plate approximately the size of finger contact area, while spending an amount of energy compatible with the perspective of a portable device. We also achieved the focusing in four points simultaneously, thus showing the possibility to provide a multitouch tactile stimulation.

Figure 1: Superposition of the glass plate displacement measured at focus time with the mock up drawing.

References:
ULTRASONIC PIEZOELECTRIC ACTUATORS FOR COMPACT SPHERICAL MOTORS

RESEARCH TOPICS: ACTUATORS, PIEZOELECTRICITY, MULTIDEGREE-OF-FREEDOM

COMPACT MOTORS

E. LEROY, J. LOZADA, M. HAFEZ

Multi-degree-of-freedom (MDOF) angular motors are used in numerous applications, such as robot articulations, omnidirectional wheels or the stabilization of optical components. Although they are conventionally used for large size applications they could find important applications for compact and miniature systems such as microcamera stabilization.

The classical method used to build angular MDOF motors is the serial or parallel association of electromagnetic motors using gimbal mechanical assemblies. Because of numerous mechanical parts and high inertia, this method limits the miniaturization and dynamic capabilities of these devices. The use of a single spherical rotor does not require gimbal systems; therefore it facilitates the realization of compact spherical motor with high dynamic capabilities.

In this work, we propose a piezoelectric ultrasonic actuator adapted to the fabrication of compact spherical motors. The novelty of the developed actuator is an adapted contact surface that allows for an extended contact area and thus the complete guidance of a spherical rotor. The actuation along the three orientations is ensured by three actuators placed around a spherical rotor.

The actuator is designed using a piezoelectric material that deforms when it is electrically excited. The shape of the actuator is optimized using finite element analysis so that two vibration modes are at the same frequency (see figure 1). The excitation of these two modes at resonance with a phase difference causes elliptical motion at the edge of the actuator. When this edge is placed against a rotor, it is possible to set it in rotation. The direction of rotation is controlled by changing the phase difference between the two exciting phases.

Based on finite element analysis results (figure 3), an actuator has been built. A picosecond laser is used for cutting the PZT material and engraving the silver electrodes. Two independent electrode sets are used to excite each vibration mode independently. The actuator is placed inside a printed-circuit board support that is used both for holding the piezo-element and ensuring electric connections. A layer of epoxy resin is deposited on the edge of the actuator to avoid damaging the ceramic.

The actuator principle has been demonstrated on a single degree-of-freedom set-up. Admittance measurements (Figure 4) show good correlation with FEM predictions with resonance frequencies of 105.3 kHz and 105.6 kHz. It is possible to actuate a cylindrical rotor in the two directions with a speed up to 120 rpm at 40Vpp. Perspectives include validation on a MDOF set-up and development of integrated electronics.

Figure 1: Spherical motor obtained by placing 4 actuators in parallel around a spherical rotor.

Figure 2: Vibration modes combined to obtain elliptical motion at the edge of the actuator.

Figure 3: Laser-cut actuator in a printed-circuit support.

Figure 4: Admittance response of the two set of electrodes.

References:


SERVER ASSISTED
KEY ESTABLISHMENT PROTOCOL FOR WSN

RESEARCH TOPICS: MACHINE TO MACHINE SECURITY
AYMEN BOUDGUIGA, ALEXIS OLIVEREAU, NOUHA OUALHA

Wireless Sensor Networks (WSNs) in industrial environment involve constrained devices that need to exchange sensitive information in dynamic topologies (e.g. monitoring and control of an assembly line), with limited existing knowledge about one another. A dedicated key establishment protocol is needed for that purpose. Nevertheless, state of the art candidates are not suitable for this role.

The Multimedia Internet KEYing (MIKEY) protocol, specified in the RFC 3830 [1], defines a simple key management scheme adapted to real time applications such as streams sharing between a server and a group of users. MIKEY assumes that the two communicating parties (an Initiator I, and a Responder R) have pre-shared credentials i.e., long term master keys or certificates. However, MIKEY relies on very large messages and involves heavy computations at peer sides. MIKEY–Ticket (RFC 6043) [2] partly solves this problem by leveraging on a Key Management Server (KMS i.e. the trusted third party), which assists the key establishment operation between I and R. Two of MIKEY–Ticket key establishment modes [2] have been inspired from two well-known server-assisted key distribution protocols of the legacy Internet, namely, Kerberos and Otway–Rees. These two protocols are however asymmetric by design, and are therefore not convenient for generic peer-to-peer key establishment. A third mode, the general MIKEY–Ticket key establishment mode, is new. Yet it suffers from an important risk of DoS attack against the responder and/or the KMS via the responder [2]. We proposed therefore an improvement to this mode in order to remove that DoS attack.

Our new protocol, named SAKE (Server Assisted Key Exchange) was proposed as a novel MIKEY-Ticket mode [3]. Obviously, specific emphasis was paid during its design to the fulfillment of security requirements:

- Replay attack: To avoid Replay attacks 128-bit nonces are used. Thus, the freshness of all exchanged messages can be ensured.
- DoS attack: As each message of SAKE contains a MAC computed with a key shared with the receiver, the DoS attack risk is reduced.
- Man In The Middle (MITM) attack: As the four first exchanges of SAKE pass through the KMS, an attacker cannot realize a MITM attack. Indeed, the attacker would need to know the keys shared between I and KMS or R and KMS in order to realize a MITM attack.
- Key escrow attack: The KMS is the only node capable of realizing a key escrow attack as it is in charge of generating I and R shared key. However, the KMS is assumed to be a trusted party. That is, the KMS will not attempt to impersonate as a legitimate node.

Eventually, SAKE was proved secure using the ProVerif formal validation tool, whose output is displayed in Figure 1.

In addition to its improved security level, SAKE also represents an improvement in terms of protocol efficiency, since it can achieve the same goal as MIKEY-Ticket in only 5 messages (instead of 6). Beyond that, SAKE was more closely adapted to WSNs needs. Its syntax was made lighter, in order to suit sensors capabilities and reduce the need for fragmentation during the protocol exchange. Meanwhile, counters were proposed as an alternative to nonces, in order to save precious sensor memory space.

SAKE was implemented on ARM and AVR (32kB RAM, 256 kB flash) platforms as part of the FP7 TWISNet (Trustworthy Wireless Industrial Sensor Networks) project.

![Figure 1: Formal Security Analysis Result Listing.](image)

References:
DELAY-TOLERANT SPACE-TIME CODING AND DECODING FOR FEMTO-ASSISTED CELLULAR NETWORKS

RESEARCH TOPICS: FEMTOCELL NETWORKS, COOPERATIVE COMMUNICATIONS

MOHAMED KAMOUN

SPONSORSHIP: EUROPEAN COMMISSION UNDER CONTRACT NUMBER ICT-248891

OF FREEDOM PROJECT

Nowadays, the concept of Femtocell networks is considered as a promising solution to increase cellular systems capacity and coverage while improving the macrocells reliability. In these networks, Femtocell Access Points (FAPs) are autonomous cheap entities that can operate either as stand-alone base stations or as relays assisting Macrocell Base Stations (MBSs) to transmit data to the users and operating on the same spectrum as the MBSs. Considering the latter scenario, the FAPs are connected to the MBSs through IP-backhaul links that can experience random delays generated by the variations of the IP-backhaul quality.

This issue motivates us to investigate the Delay-Tolerant Space-Time Codes (DT-STCs), constructed in [1], in the femto-based cooperative network. Such distributed M×T STC can be used by the M different MBSs and FAP antennas where each antenna sends a row of the space-time code over T time slots. The delay-tolerant codes of [1] are optimal in synchronous transmission with full-rate, full-diversity and non-vanishing determinants. Moreover, they preserve their full-diversity for arbitrary delays in asynchronous transmission. However, these codes were designed for short dimensions and no interference was assumed between subsequent codewords when transmission delays are introduced. This can be achieved by adding guard intervals between the codewords, from the MBS and the FAP. These extra-delays induce significant rate loss especially for short codes.

In order to improve the rate and to be able to handle larger delays, we propose in this paper [2] to use DT-STCs with larger code length by allowing continuous codewords transmission. Therefore, the new space-time code will be a concatenation of subsequent DT-STCs. Then, when delays are incurred by the MBS-FAP backhaul link, the transmission will result in overlapped sub-codewords. In this case, optimal decoders such as Maximum Likelihood and Sphere Decoder (SD) become very complex with a complexity increasing exponentially with the code length. Therefore, we are interested in sub-optimal decoders that can achieve performance-complexity tradeoffs. First, we study the Enhanced Maximum A Posteriori Decision Feedback Equalizer (E-MAP-DFE) that is based on an iterative interference Cancellation (IC) technique. Then, in order to reduce further the complexity, we propose a new sub-optimal decoder called Successive Sub-Codewords IC (SSC-IC) decoding. It consists in partitioning the whole codewords frame into windows of sub-codewords and applying a successive decoding on the subsequent sub-codewords to decode the symbols. In this way, the received vector at each decoding window can be decomposed into three terms: the previous detected sub-codewords, the target sub-codeword to be detected, and the next sub-codewords that are acting as interference added to the Gaussian noise.

In Figure 2, we evaluate the performance of the concatenated DT-STCs with the described optimal SD and sub-optimal decoders, E-MAP-DFE and proposed SSC-IC decoder with two variants. We consider the cooperative system of Figure 1 with an MBS using one antenna and one FAP equipped with one antenna, then M = 2, and the destination has 2 receive antennas. We assume that the new DT-STC used by MBS-FAP is a concatenation of 3 subsequent DT 2×2 codes of [1], and a delay of 1 symbol is introduced by the IP-backhaul link. The performances are represented in terms of Bit Error Rate (BER) versus signal-to-noise ratio Eb/N0 per receive antenna.

While reducing significantly the SD complexity, we can observe that SSC-IC SD is only 1 dB far from SD at a BER=10−4. However, the decoder variant with MMSE DFE preprocessing loses the full-diversity order due to erroneous estimations and error propagations. On the other hand, at the same BER, the SSC-IC E-MAP-DFE loses 1.6 dB and 2.6 dB compared to the E-MAP-DFE and SD decoders, respectively, but gains in lower complexity.

References:
RELIABILITY FOR EMERGENCY APPLICATIONS IN INTERNET OF THINGS

RESEARCH TOPICS: WIRELESS SENSOR NETWORKS, IOT, ROUTING PROTOCOL, RELIABILITY

NOURHENE MAALEL, ENRICO NATALIZIO, ABDELMADJID BOUABDALLAH, PIERRE ROUX, MOUNIR KELLIL

SPONSORSHIP: UTC. PARTNERSHIP: DCOSS 2013

The proliferation of a variety of connected devices, including smart devices, through the Internet and the need for making them communicate efficiently despite their heterogeneity, has led to the emergence of a new Internet paradigm called Internet of Things (IoT) [1]. IoT application covers different domains like transport, agriculture, defense, and smart environments like homes and buildings. However, “Things” may face various technical challenges. A network of things has to be self-adaptive and resilient to communication errors by providing efficient mechanisms for information distribution especially in the multi-hop scenario. These requirements have to be satisfied in an architecture that can be constrained by limited processing capabilities, scarce energy resources and unreliable communication channels. In particular, in harsh environments, the radio signal is often affected by interference; medium access conflicts, multipath fading, shadowing etc. These problems may result in significant packet losses. Also, the success of any application (particularly mission-critical ones like life-care data and alarms) requires the delivery of high-priority events to sinks without any traffic loss.

To solve these problems, we define a reliable and energy-efficient data transmission protocol, called AJIA (Adaptive Joint protocol based on Implicit ACK) [2]. To ensure reliable data transmission, we use a basic routing scheme for data transmission/routing while exploiting the overhearing characteristic of wireless communications for any further data re-transmission needed to recover lost packets.

In the basic routing strategy we have considered, the data packet is transmitted level-by-level towards the sink. This is made possible by assigning to each node a level corresponding to its hop-distance from the sink (cf. Figure 1).

In addition, the overhearing characteristic of wireless communications is exemplified in Figure 2.

Concretely, when a node transmits a packet, nodes in its neighborhood overhear the packet transmission even if there are not the intended recipients, due to the broadcast nature of the wireless channel. The overhearing mechanism enables to reduce the communication overhead for data retransmission compared to the usage of conventional ACK/NACK messages [3], which, by essence, are bandwidth-consuming, and thus are not adapted for highly constrained and error prone environments.

Furthermore, in order to avoid error accumulation (induced by traditional end-to-end error recovery mechanisms) we propose a hop-by-hop packet loss recovery mechanism, in which intermediate nodes also take responsibility for loss detection and recovery.

The retransmission of lost packets is carried out on the most reliable link (the highest AJIA metric) between the node that sent the (lost) packet and its one-hop neighbors. The said AJIA metric is calculated with the following formula:

$$AJIA(A, B) = K \times e^{\text{age}} \times \frac{1}{\text{Phist} AB} + E$$

Where $LQI_{AB}$ denotes the link state indicator between nodes A and B and age corresponds to the delay since the LQI value has been recorded. The Phist represents the probability of link success between nodes A and B. K is a constant used to weight the equation. E is related to the resource available at the device.

This metric is calculated for all the candidate nodes closer to the sink.

We have evaluated the performance of AJIA via a simulation tool. The performance results will be published in a future work.

References:
VIN6: VIN-BASED IPV6 PROVIDER INDEPENDENT ADDRESSING FOR FUTURE VEHICULAR INTERNET COMMUNICATIONS

RESEARCH TOPICS: IPV6 VEHICLE-TO-INTERNET COMMUNICATIONS

SOFIANE IMADALI, VÉRONIQUE VÈQUE, ALEXANDRU PETRESCU, MICHAEL BOC
PARTNERSHIP: SUPELEC

Recent years witnessed growing concerns regarding scalability of today's Internet architecture. Universality of the single IP numbering space is one cause of unbound growth of entropy and inter-domain routing tables in the core network. Recently, the Internet Architecture Board proposed to review the Internet addressing architecture to separate name and location spaces. Recent Future Internet proposals from industry and academia considered Locator/Identifier split realm as a fundamental building block. Locator/ID Separation Protocol (LISP) defines two addressing elements: Routing locators (RLOCs) and Endpoint Identifiers (EIDs) on the IP numbering space. The translation of EIDs to RLOCs is achieved through efficient and scalable network-based mapping system [1].

We focus in our research [2] on the integration of Vehicle-to-Internet communications to LISP-based (and compatible) Future Internet architectures with the definition of a new Vehicle-specific namespace: Vehicle Identification Numbers (VIN) (Figure 1). In particular, as illustrated in Figure 2, we define a scalable network-based architecture with Provider Independent subset of the IPv6 space that bears the semantics of the VIN hierarchy.

The VIN is an ISO 3779, ISO 3780 standardized 17 alphanumeric characters hierarchical code that uniquely identifies a vehicle worldwide. The code contains 3 sections: WMI, VDS, and VIS. The WMI is 3 digits long and uniquely designates the manufacturer’s continent, country, and the unique national identifier. The VDS is 6 characters long and describes the vehicle: weight, model, engine type, body style or the engine power. The VIS is 8 digits long. Combined with VDS, they uniquely identify a vehicle within a car manufacturer for 30 years. Combined with WMI, they uniquely identify a vehicle worldwide. The VIN validity is maintained by the Society of Automotive Engineers (SAE) along with national automotive authorities. By hierarchical design of our VIN numbering space, it will consist in the tuple (WMI, VIS). This tuple is then converted to a topological IPv6 address and a Provider Independent VIN-based IPv6 prefix.

Figure 2 illustrates the use of the obtained address and prefix. Regardless of the current location of the vehicle, a correspondent node at an arbitrary location in the Internet, can issue a packet using the Provider Independent VIN-based addressing towards the vehicle. This data packet is then routed to the the IP Location and Forwarding Service located in the Manufacturer Domain. In LISP terminology, this domain performs the mapping between the RLOCs (current actual vehicle location) and the EID (VIN in our case). It is noteworthy to mention here that the correspondent node may be a vehicle as well, placing the manufacturer domain in the V2V communication. Our work is not the first that considers VIN as valuable information for IPv6 addresses. Authors of [3], use some VIN fields to map them into decimal numbers separately, before binary conversion to Extended Unique Identifiers (EUI-64). One IPv6 address is generated. This method is inefficient in terms of compression and uniqueness of the EUI-64 and is also debatable (some included values are from VDS which is non-standard). Our compression method in comparison is 15.687% more efficient (16 bits). Finally, one generated IPv6 address is not enough for the architecture design that we described above.

Authors of [4] provide a method to determine the IPv6 address of a component inside the vehicle. This proposal uses only the VIS section of the VIN as part of the generated prefix along with a non-standard method to set a global prefixes. In order to assure global scope for generated prefixes, an administrative check must be performed.

References:
3D CAMERA MODELING AND CALIBRATION

RESEARCH TOPICS: COMPUTER VISION, SENSOR CALIBRATION
A. BELHEDI, P. SAYD, V. GAY-BELLILE, S. BOURGEOS, M. DHOME
PARTNERSHIP: INSTITUT PASCAL (UNIVERSITÉ BLAISE PASCAL – CLERMONT-FERRAND)

Compared to conventional camera, 3D cameras provide an image for which each pixel value represents the depth of the scene. Therefore, such a sensor provides a real-time (25 frames per second) 3D reconstruction of the scene observed by the camera.

While such data can be extremely useful in many field of application, such as robotic, human-computer interaction, video-surveillance, as-built reconstruction, etc., their use was limited due to their cost (several thousands of $), their low resolution (less than 100x100 pixels), their electrical consumption and their weight. However, the recent apparition of low-cost sensor with higher resolution, such as Microsoft Kinect, and low consumption embedded camera, such as Primesense Capri, has boosted the exploitation of this technology.

Until now, most of the applications relying on this technology ignore the measurement error affecting the depth estimation while a modelling and calibration of this error would improve the accuracy and robustness.

The objective of this work was to provide calibration process that was both easy-to-use and accurate.

PROPOSED APPROACH

This work focused on two aspect of the calibration process. First, we worked on the modeling of the sensor error. Second, we worked on the development of an easy-to-use calibration process.

Concerning the error modeling, we proposed to model both the noise and the systematic error that affect the depth measure of time-of-flight 3D camera.

First, we demonstrated experimentally that the noise affecting the depth measure can be correctly approximated with a Gaussian distribution whose standard deviation depends both on the observation distance and the position of the pixel in the image. To model these variations, we proposed to use 3D Thin-Plate-Spline model. Our evaluations demonstrated that such a model allow to reach a high accuracy with a limited number of parameters.

Second, to model the systematic error affecting the depth measure, we proposed to replace the usual spline-based model with a non-parametric model [1]. Indeed, spline-based model rely on strong assumption whose validity is, until now, not demonstrated. Experimental evaluation demonstrated that our solution allow to reach a similar accuracy with only weak assumption on the nature of the systematic error. Therefore, compared to state of the art approach, our solution is able to handle a wider variety of distortions, and, consequently, camera sensors.

Concerning the calibration process, until now, the state of the art approaches relied on a complex calibration process. Indeed, state-of-the-art process requires both to observe a surface from a distance from 1 meter to 7 meters, and to provide an error-free estimation of the depth measure for each pixel of the camera. In practice, such calibration process is extremely time-consuming, requires a large calibration room, and should be achieved by an expert.

To reach an easy-to-use process, we introduced an algorithm that simultaneously estimates the camera calibration and the ground-truth[2]. Our solution relies on an alternate optimization of the ground-truth and the camera calibration. The ground truth estimation relies on an assumption of a planar-by-piece environment: each calibration image is segmented into planar region, and a 3D plane is fitted onto each region. The resulting 3D planes are then considered as ground-truth and used to estimate the camera calibration. These two steps are repeated until convergence of the process.

Figure 1. Camboard Pico XS camera from PMD
Figure 2. Distortion of the depth measure before (left) and after (right) the calibration process

References:
DENSE IMAGE MATCHING:
A GENERAL FRAMEWORK COMBINING
DIRECT AND FEATURE-BASED COSTS

RESEARCH TOPICS: COMPUTER VISION, DEPTH MAP, OPTICAL FLOW, OPTIMIZATION TGV2
JIM BRAUX-ZIN, ROMAIN DUPONT, ADRIEN BARTOLI
PARTNERSHIP: ISIT LAB (UNIVERSITÉ D’AUVERGNE)

ABSTRACT
Dense motion field estimation (typically optical flow, stereo disparity and surface registration) is a key computer vision problem. Many solutions have been proposed to compute small or large displacements [1] [3], narrow or wide baseline stereo disparity, but a unified methodology is still lacking. We here introduce a general framework that robustly combines direct and feature-based matching. The feature-based cost is built around a novel robust distance function that handles keypoints and “weak” features such as segments. It allows us to use putative feature matches, which may contain mismatches to guide dense motion estimation out of local minima. Our framework uses a robust direct data term (AD-Census). It is implemented with a powerful second order Total Generalized Variation regularization [4] with external and self-occlusion reasoning [5]. Our framework achieves state of the art performance in several cases (standard optical flow benchmarks, wide-baseline stereo and non-rigid surface registration). Our framework has a modular design that customizes to specific application needs.

CONCEPT AND RESULTS
The framework introduced in this article allows us to greatly extend the scope of applicability of variational optical flow techniques. We combined a modern powerful discontinuity preserving regularizer with robust direct data term and features integration. We generalized and extended the model of [1] to support any point features and introduced the novel concept of a priori matches to enable the use of weakly localized features such as segments [2]. Self-occlusion detection [5], rarely accounted for in optical flow estimation, further increases the robustness of our approach. This allowed us to showcase state of the art results on standard narrow-baseline optical flow and wide-baseline stereo. Preliminary results on non-rigid surface detection (Table 1 and Table 2) compare favorably with other methods and suggest promising use cases. Future work involves the improvement of each building block: higher-order regularization, richer direct data term and new features such as contours.

Table 1. Results on the KITTI benchmark. Our method ranked 5 is the top true 2D optical flow method. Methods 1–4 are actually not true 2D optical flow methods as they use a priori information about the scenes.

<table>
<thead>
<tr>
<th>R</th>
<th>Method</th>
<th>Average Err</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PR-SF+E</td>
<td>1.7</td>
</tr>
<tr>
<td>2</td>
<td>PCBP-Flow</td>
<td>2.2</td>
</tr>
<tr>
<td>3</td>
<td>MotionSLIC</td>
<td>2.7</td>
</tr>
<tr>
<td>4</td>
<td>PR-Sceneflow</td>
<td>3.3</td>
</tr>
<tr>
<td>5</td>
<td>Ours with TGV2</td>
<td>4.5</td>
</tr>
<tr>
<td>6</td>
<td>fSGM</td>
<td>5.7</td>
</tr>
<tr>
<td>7</td>
<td>Ours with Census</td>
<td>12.2</td>
</tr>
</tbody>
</table>

Table 2. Benchmark on the training sequence of the Middlebury dataset with average end-point errors (px): it shows that adding sparse matches does not affect overall performance for small baseline based data.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Average Err (px)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDOF [1]</td>
<td>0.36</td>
</tr>
<tr>
<td>Ours</td>
<td>0.33</td>
</tr>
<tr>
<td>Ours with matches [6]</td>
<td>0.33</td>
</tr>
</tbody>
</table>

References:
HUMAN ACTION SEGMENTATION USING DEEPLY OPTIMIZED HOUGH TRANSFORM

RESEARCH TOPICS: COMPUTER VISION, ACTION AND ACTIVITY RECOGNITION, ACTION SEGMENTATION
ADRIEN CHAN-HON-TONG, LAURENT LUCAT, PATRICK SAYD, CATHERINE ACHARD
PARTNERSHIP: INSTITUT DES SYSTEMES INTELLIGENTS ET DE ROBOTIQUE

Hough-like methods (Implicit Shape Model, Hough forest,...) have been successfully applied in multiple computer vision fields like object detection, tracking, skeleton extraction or human action detection. However, these methods are known to generate false positives. To handle this issue, several works like Max-Margin Hough Transform (MMHT) or Implicit Shape Kernel (ISK) have reported significant performance improvements by adding discriminative parameters to the generative ones introduced by the Implicit Shape Model (ISM). We propose to use only discriminative parameters that are globally optimized according to all the variables of the Hough transform. To this end, we abstract the common vote process of all Hough methods into linear equations, leading to a training formulation that can be solved using linear programming solvers. Our new Hough Transform significantly outperforms the previous ones on HoneyBee and TUM datasets, two public databases of action and behaviour segmentation.

DEEPLY OPTIMIZED HOUGH TRANSFORM (DOHT)

In the context of temporal signals segmentation and recognition, the Hough Transform is composed of three steps:
1. Feature extraction and quantization to form codewords
2. Each codeword votes for each time (in a large neighborhood) and each label according to a specific learned weight
3. All the votes are agglomerated to form the Hough score from which segmentation decisions are taken.

In our DOHT approach, we focus on the optimization of the weights used during the vote process and so to the link between codewords and votes which is generic whatever the features and codewords used. Thus, the proposed method can be employed in the Hough forest context by substituting the weights estimated by ISM by the weights optimized by our proposed method.

We propose to use Hough transform to segment and recognize temporal series. In a non parametric context, the training of Hough transform consists to properly select the weights used in the voting process. The simple way (Implicit Shape Model) consists in computing some probabilities on the training database, leading to a generative model. Some methods (Max-Margin Hough Transform, Implicit Shape Kernel) propose to add some parameters optimized on a training database in a discriminative way. We propose to skip the first step based on a generative model and to globally learn all the parameters of the Hough transform on the training database, resulting a deeply discriminative model. This required to reformulate the voting process to express it in a linear form in order to use linear programming solvers.

EXPERIMENTAL RESULTS ON HUMAN ACTION SEGMENTATION

TUM is a multi-sensor dataset and in particular it contains skeleton streams. It is composed of 19 sequences around 2 minutes each containing 9 kinds of actions (each action is a label) like Lowering an object, Opening a drawer, etc. performed by 5 people. We decide to consider only skeleton based features (8 articulations). Hence, the input signal of our algorithm is the 3D positions of each articulation at each time. In our experiments, DOHT significantly outperforms ISM, MMHT and SISM and achieves equivalent performance than a SVM based on the same features and codeword applied on the optimal segmentation (obtained from the ground truth). Hence, for this dataset, we achieve equivalent performance than the best published (82.6% against 84.3%) without using the optimal segmentation.

Experiments have also been conducted on the Honeybee dataset. The Honeybee dataset provides tracking output of honey bees having 3 kinds of behaviours correlated with their trajectories. In this experiment, DOHT achieves equivalent performances than the best published results.

This work opens new perspectives for other contexts such as object segmentation in image, video spatio-temporal segmentation, automatic speech segmentation, sign language segmentation.

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy on TUM</th>
<th>Accuracy mean on Honeybee</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISM</td>
<td>58.4</td>
<td>71.9</td>
</tr>
<tr>
<td>MMHT</td>
<td>69.6</td>
<td>78.8</td>
</tr>
<tr>
<td>SISM</td>
<td>68.5</td>
<td>77.5</td>
</tr>
<tr>
<td>DOHT (ours)</td>
<td>82.6</td>
<td>86.5</td>
</tr>
</tbody>
</table>

Table 1: Results on TUM and Honeybee datasets

References:
IMPROVING PERSON DETECTION PERFORMANCE USING CONTEXT

RESEARCH TOPICS: COMPUTER VISION, VIDEO SURVEILLANCE, PERSON DETECTION, CONTEXT, MACHINE LEARNING
THIERRY CHESNAIS, YOANN DHOME, NICOLAS ALLEZARD, THIERRY CHATEAU
PARTNERSHIP: INSTITUT PASCAL

With the rise of videosurveillance systems comes a logical need for automatic and real-time processes to analyze the huge amount of generated data. Among these tools, pedestrian detection algorithms are essential, because in videosurveillance locating people is often the first step leading to more complex behavioral analyses. Classical pedestrian detection approaches are based on machine learning and pattern recognition algorithms. Thus they generally underperform when the pedestrians’ appearance observed by a camera tends to differ too much from the one in the generic training dataset (Figure 1). We study the concept of the contextualization of such a detector. This consists in introducing scene information into a generic pedestrian detector. The main objective is to adapt it to the most frequent situations and so to improve its overall performances.

The key hypothesis made here is that the camera is static, which is common in videosurveillance scenarios. The goal is to identify every part of the detector which can benefit from the approach in order to fully contextualize it. To make the contextualization process easier, our method is completely automatic and is based on semi-supervised learning methods. First of all, data coming from the scene are gathered. We propose different oracles (Figure 2) to detect some pedestrians in order to catch their appearance and to form a contextualized training dataset. Then, we analyze the scene geometry, which influences the size and the orientation of the pedestrians and we divide the scene into different regions. In each region, pedestrians as well as background elements share a similar appearance. In the second step, all this information is used to build the final detector which is composed of several classifiers, one by region. Each classifier independently scans its dedicated piece of image. Thus, it is only trained with a region-specific contextualized dataset, containing less appearance variability than a global one. Consequently, the training stage is easier and the overall detection results on the scene are improved.

We compared performances of the contextualized detector with a state of art one. We chose the Dalal and Triggs detector detector available in OpenCV and applied the same evaluation criterion. When the learning dataset has a very different point of view, the Dalal and Triggs detector is not very successful in detecting most of pedestrians. Our method, using an oracle formed by basic classifiers, can achieve significantly better performance because it is contextualized.

Figure 1: A detector trained on a generic dataset (left) reach lower performances than a contextualized classifier (right).

Figure 2: Oracle diagram

Figure 3: Recall-precision curve on PETS2007

References:
PEOPLE REIDENTIFICATION IN NON OVERLAPPING CAMERAS

RESEARCH TOPICS: COMPUTER VISION, VIDEO SURVEILLANCE, PEOPLE REIDENTIFICATION, TRACKING, CAMERA NETWORK
BORIS MEDEN, PATRICK SAYD, FRÉDÉRIC LERASLE
PARTNERSHIP: LAAS

We propose a pedestrian tracking system that uses re-identification to monitor non-overlapping cameras. As tracking, re-identification is an assignment problem, the difficulties being to generate an accurate representation and to prune unlikely pairings.

The assignments are realised in two stages. First, a Markovian multi-target tracking-by-detection framework which includes identification in the search space is run in the cameras. This generates tracks in the cameras and a first assignment between them thanks to the local identification. This solution is then optimized globally by a network supervisor benefiting from coarse topology knowledge over a sliding window with MCMC sampling. The tracking results obtained on a large ground-truthed dataset demonstrate the effectiveness of the approach.

PROPOSED APPROACH

The main steps of our algorithm are:
1. At the camera level, automatic distributed trackers based on HOG detections and track targets for each camera. The appearance model used is composed of horizontal stripes of HSV histograms weighted by their distances to the symmetry axis. The use of topology allows to instantiate new identities from the feeding areas in an identity database, which we compare with to perform re-identification. The mixed-state formalism uses that database and sample in this identity space. That way the tracker produces a tracklet and re-identification probabilities in the database representing the belief of the tracker. The resulting tracklets are sent to the supervisor along with their probabilities of identity, their time of existence and their areas.
2. At the network level, the supervisor resorts to deferred logic to optimize the assignment between the received tracklets using re-identification distributions and network topology information. The combinatorial space is efficiently explored through MCMC sampling. Tracks output by the supervisor are optimized to represent the activity of the same person. The local identification of the distributed mixed-state filters is fused with coarse topology knowledge to yield the likelihood function of the MCMC data association.

EXPERIMENTAL RESULTS

The NOFOV sequence presents a total of 12 pedestrians wandering between 5 cameras.

Figure 2 gives an overview of the network. The dataset is 4000 frame-long. The identity database is built at runtime using models of the trackers starting in Cam0, entry area.

We compare here the method based only on colorimetric information and particle filtering, with the supervised system which optimizes the tracklets with topological constraints. Table 1 presents true re-identification rates of the supervisor applied to real data optimizing the output of our tracking-by-reidentification module. Fusing topological constraints with identities distributions allows our MCMC formulation to increase the assignment quality.

Table 1: Results of reidentification rates on the NOFOV sequence

<table>
<thead>
<tr>
<th>NOFOV Sequence</th>
<th>cam0</th>
<th>cam1</th>
<th>cam2</th>
<th>cam3</th>
<th>cam4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tracking-by-Reidentification</td>
<td>88.1%</td>
<td>65.3%</td>
<td>58.5%</td>
<td>54.6%</td>
<td>54%</td>
</tr>
<tr>
<td>Tracking-by-Reidentification supervised</td>
<td>90.6%</td>
<td>76.2%</td>
<td>68.2%</td>
<td>63.9%</td>
<td>62%</td>
</tr>
</tbody>
</table>

References:
HIERARCHICAL SCENE MODEL FOR SPATIAL-COLOR MIXTURE OF GAUSSIANS

RESEARCH TOPICS: COMPUTER VISION, VIDEO SURVEILLANCE, BACKGROUND SUBTRACTION, VISUAL TRACKING

CHRISTOPHE GABARD, LAURENT LUCAT, CATHERINE ACHARD

PARTNERSHIP: UNIVERSITÉ PIERRE ET MARIE CURIE

The proposed approach uses the principle of Spatial Gaussian Mixture Models (SGMM), also known as SMOG. In this work, a mixture of Gaussian distributions is used, but unlike the well-known Stauffer and Grimson method, each Gaussian Mode represents a group of pixels instead of only a single pixel. Then a single list of modes is created and updated for each frame to represent the whole scene. To introduce spatial and color consistency, each pixel observation is described by both spatial and color coordinates: X = [x, y, R, G, B]. From these observations, the scene is characterized by a set of Gaussian distributions in a 5-dimensional space. A simplified representation of the so-obtained model is shown in Figure 1.

The proposed system aims to reach a moving object detection with a pixel-size precision, and robust to realistic difficult environment conditions like acquisition noise or camera vibration. First, the scene is modeled in a hierarchical way, with a pixel-wise temporal modeling and a global representation encoding the spatial consistency between pixels. This allows, one the hand to overcome some problems encountered with the SMOG approach and secondly, to make detection decisions on consistent sets of pixels, leading to robustness. Another contribution is the control of mode creation and evolution. It avoids color dispersion of modes enforces spatial density and so eliminates important drift with time. Finally, as the scene is modeled as an entity, the Gaussian mixture represents both background and foreground areas and a decision step is implemented and combined with a local shadow detection process to label each mode as background or object.

Local modeling performed independently on each pixel is typically proposed by conventional approaches of motion detection. Using a descriptor defined on blocks of pixels of predetermined shape and size leads to more robustness, but penalises the detection precision. To overcome these limitations, the described approach combines the accuracy of the pixel information with the robustness of a consistent pixels set decision. The complete scene modeling (background + foreground) also allows to manage some ambiguities when moving object has an appearance similar to the one of the background.

Where conventional algorithms have only background model, the proposed method introduces two models (background and object). This allows to transform the detection task into a classification problem with two classes.

The experimental results show that the described approach improves both methods from which it derives thanks to the rigorous control of mode consistency, avoiding model drifts.

References:
LOCALITY-CONSTRAINED AND SPATIALLY REGULARIZED CODING FOR SCENE CATEGORIZATION

RESEARCH TOPICS: COMPUTER VISION, SCENE CATEGORIZATION
AYMEN SHABOU, HERVÉ LE BORGNE
SPONSORSHIP: CVPR 2012

In recent works addressing object recognition and scene classification tasks, the bag-of-words (BoW) is one of the most popular models for feature design. Inspired by the seminal work of [1], different approaches have been proposed to improve both its generative property to describe accurately images and its discriminatory power for classification. Despite remarkable progresses, it remains challenges concerning the extraction of local descriptors, codebook design, local descriptors coding and pooling, including a spatial layout into the final feature, and the final classification.

Given a training dataset, the first step of the BoW method consists in extracting local features, such as SIFT, HOG or SURF [1], from images. Then a codebook, which is a set of visual words, is built to represent them. Initial methods are based on clustering techniques, such as K-means [1]. Despite their efficiency, the obtained codebooks suffer from several drawbacks such as distortion errors and low discriminative ability. Other approaches have rather attempted to improve the discriminative power of the codebook while compacting it relying on a supervised method. However, recent works of [2] show that, for the recognition task, codebook design is less critical than the next stages (coding, pooling and spatial layout).

Coding consists in decomposing local features over a codebook in order to satisfy some desirable properties. Various strategies are proposed in the literature. The earliest one is the hard coding, a voting scheme that is simple yet highly sensitive to reconstruction errors induced by the codebook. A more robust voting approach is the soft coding, which assigns a descriptor to all the visual words according to their distances. Authors of [3] introduced another coding property, called locality that ensures sparsity while remaining efficient.

The next step of BoW design is pooling the obtained codes to obtain a compact signature. Usually, the maxpooling operation is used, leading to signatures that are appropriate to linear classifiers. Finally, the Spatial Pyramid Matching (SPM) step, proposed in [4], is usually exploited to include some spatial layout information to the BoW. Such vectors of fixed size can then feed a machine learning algorithm such as SVM or Boosting.

In the current work, the local feature coding step is investigated. While several techniques have outperformed the classic hard assignment by introducing either the locality or the similarity constraints in the feature space, we propose a new formalism that implicitly preserves these properties while adding the local contextual information from the spatial domain of the image. Figure 1 shows a schematic comparison. The proposed coding approach is divided into two steps. The first step is an optimal basis selection for each local feature, formulated as a labeling problem. For this purpose, we introduce a novel objective function that includes locality and similarity (or coherency) constraints in both the feature space and the spatial domain of the image. Furthermore, we provide an appropriate efficient optimization algorithm, called knn-expansion, which is inspired from the fast optimization tools dedicated to Markov Random Field (MRF) based energy minimization task. The second step consists in assigning responses (or values) to the selected optimal bases.

This new approach enriches the BoW signature leading to more accurate features for classification than the state-of-the-art methods. Furthermore, it is generic and can thus be added to several recent coding strategies. The proposed method improves the performances of several state-of-the-art coding schemes for scene classification on three publicly available benchmarks (UIUC 8-sport, Scene-15 and Caltech-101).

Figure 1: Schematic comparison of basis selection methods to code dense descriptors. The first configuration is the one adopted by some recent coding approaches. The second configuration corresponds to the proposed LCSR method.

References:
IDENTIFYING BAD SEMANTIC NEIGHBORS
FOR IMPROVING DISTRIBUTIONAL THESAURUS

RESEARCH TOPICS: NATURAL LANGUAGE PROCESSING
OLIVIER FERRET
SPONSORSHIP: ACL 2013

Distributional thesauri are now widely used in a large number of Natural Language Processing tasks. The work we present in this article focuses on the automatic building of a thesauri from a corpus. As illustrated by Table 1, such thesaurus gives for each of its entries a list of words, called semantic neighbors that are supposed to be semantically linked to the entry. Generally, each neighbor is associated with a weight that characterizes the strength of its link with the entry and all the neighbors of an entry are sorted according to the decreasing order of their weight.

Following work such as [1], a widespread way to build a thesaurus from a corpus is to use a semantic similarity measure for extracting the semantic neighbors of the entries of the thesaurus. Three main ways of implementing such measures can be distinguished. The first one relies on handcrafted resources in which semantic relations are clearly identified. Work based on WordNet-like lexical networks for building semantic similarity measures. These measures typically exploit the hierarchical structure of these networks, based on hypernymy relations. The second approach makes use of a less structured source of knowledge about words such as the definitions of classical dictionaries or the glosses of WordNet. WordNet’s glosses were used to support Lesklike measures and more recently, measures were also defined from Wikipedia or Wiktionaries [2]. The last option is the corpusbased approach, based on the distributional hypothesis: each word is characterized by the set of contexts from a corpus in which it appears and the semantic similarity of two words is computed from the contexts they share. This perspective was first adopted by [1].

The work we present in this article takes place in the framework defined by [1] for implementing the distributional approach but proposes a new method for improving a thesaurus built in this context based on the identification of its bad semantic neighbors, that is to say the neighbors of the entry that are actually not semantically similar to the entry, rather than on the adaptation of the weight of their features. By discarding these bad neighbors or at least by downgrading them, the rank of true semantic neighbors is expected to be lower. This makes the thesaurus more interesting to use since the quality of such thesaurus strongly decreases as the rank of the neighbors of its entries increases, which means in practice that only the first neighbors of an entry can be generally exploited.

The approach we propose for identifying the bad semantic neighbors of a thesaurus entry relies on the distributional hypothesis, as the method for the initial building of the thesaurus, but implements it in a different way. This hypothesis roughly specifies that from a semantic viewpoint a word in context, i.e. in a sentence, from all other words and more particularly, from those of its neighbors in a distributional thesaurus that are likely to be actually not semantically similar to it. The underlying hypothesis follows the distributional principles: a word and a synonym should appear in the same contexts, which means that they are characterized by the same features. As a consequence, a model based on these features that can identify a word in a sentence is likely to identify also a synonym of this word in a sentence, and by extension, to identify a word that is paradigmatically linked to it. More precisely, we found that such model is specifically effective for discarding the bad neighbors of the entries of a distributional thesaurus.

The proposed method was tested on a large thesaurus of nouns for English and led to a significant improvement of this thesaurus, especially for middle and low frequency entries and for semantic relatedness.

| abnormality | defect [0.30], disorder [0.23], deformity [0.22], mutation [0.21], prolapse [0.21], anomaly [0.21] . . . |
| agreement   | accord [0.44], deal [0.41], pact [0.38], treaty [0.36], negotiation [0.35], proposal [0.32], arrangement [0.30] . . . |
| cabdriver   | waterworks [0.23], toolmaker [0.22], weaponer [0.17], valkyry [0.17], wang [0.17], amusement-park [0.17] . . . |
| machination | hollowness [0.15], share-price [0.12], clockmaker [0.12], huguenot [0.12], wrangling [0.12], alternation [0.12] . . . |

Table 1: First neighbors of some entries of the distributional thesaurus

References:
SPACE-TIME ROBUST VIDEO REPRESENTATION FOR ACTION RECOGNITION

With the constant expansion of visual online collections (such as Youtube) and video protection, action recognition has become an important problem in computer vision. It is a difficult task since online videos are subject to large visual diversity. Robust to such variability, Bag-of-Features (BoF) [1] has been adopted as the main paradigm for representing a video. A BoF is computed in 3 steps: (1) local feature extraction, (2) local feature coding and (3) local feature pooling.

This paper focuses on the third step that aims at summarizing the feature code distribution in a fixed length vector. We propose a novel content driven pooling that leverages space-time context while being robust toward global space-time transformations. Traditional pooling considers each local feature independently. Such an algorithm discards the local feature position information in the video space-volume. However, this space-time context has been proven useful for classification [2]. Indeed, discriminative information is not equally distributed in the video space-time domain as shown by Figure 1. To benefit from this context, spatial pooling [2] divides a video using fixed segmentation grids and pools the features locally in each grid cell. Despite the performance improvement, spatial pooling loses the BoF space-time invariance. Different action instances with various localizations in the space-time volume can result in divergent representations. This problem is severe for the actions which have dramatic space-time variance as illustrated in Figure 2. In this case, spatial pooling divides one action across different grid cells which may lead to a significant performance drop. A BoF representation robust to space-time variance is therefore critical for action recognition.

In this work, we propose to take advantage of the space-time discriminative context with an emphasis on retaining the space-time robustness. Beyond standard spatial pooling which uses fixed segmentation grids, we segment a video according to its content through saliency maps. Our algorithm relies on the idea that the discriminative information has a non-uniform distribution in saliency spaces. For example, “Running” is more likely to be distinguished from “Walking” by regions subject to high motion. In addition, different saliencies can highlight different regions in the video space-time volumes. They may capture complementary information which can be appropriately fused. Based on those observations, we propose two main contributions.

We introduce a novel space-time invariant pooling which leverages the space-time context. We first extract video structural cues using various saliency measures. We then aggregate the local feature statistics over fixed saliency subregions, each sub-region defining a structural primitive. Focusing on different structural aspects, cornerness, light and motion saliences are investigated. Cornerness highlights regions repeatable under geometric transformations, motion identifies regions with strong dynamics and light provides coarse object segmentation. To automatically determine the optimal structural primitives combination associated to a specific action, we introduce a sparse feature weighting regularizer, which is able to assign optimal weights to different feature groups.

We evaluate our approach on standard action datasets (KTH, UCF50 and HMDB). We improve state-of-the-art performances on every dataset. Most noticeably, the accuracy of our algorithm reaches 51.8% on the challenging HMDB movie dataset which outperforms the state-of-the-art of 7.3% relatively.

References:
BUILDING SPECIALIZED BILINGUAL LEXICONS USING LARGE-SCALE BACKGROUND KNOWLEDGE

RESEARCH TOPICS: MACHINE TRANSLATION AND CROSS-LINGUAL INFORMATION RETRIEVAL

DHOUHA BOUAMOR, ADRIAN POPESCU, NASREDINE SEMMAR, PIERRE ZWEIGENBAUM

SPONSORSHIP: EMNLP 2013

PARTNERSHIP: CNRS LIMSI

The plethora of textual information shared on the Web is strongly multilingual and users’ information needs often go well beyond their knowledge of foreign languages. In such cases, efficient machine translation and cross-lingual information retrieval systems are needed. Machine translation already has a decade long history and an array of commercial systems was already deployed, including Google Translate and Systran. However, due to the intrinsic difficulty of the task, a number of related problems remain open, including: the gap between text semantics and statistically derived translations, the scarcity of resources in a large majority of languages and the quality of automatically obtained resources and translations. While the first challenge is general and inherent to any automatic approach, the second and the third can be at least partially addressed by an appropriate exploitation of multilingual resources that are increasingly available on the Web.

In this paper we focus on the automatic creation of domain-specific bilingual lexicons. Such resources play a vital role in Natural Language Processing (NLP) applications that involve different languages. At first, research on lexical extraction has relied on the use of parallel corpora [1]. The scarcity of such corpora, in particular for specialized domains and for language pairs not involving English, pushed researchers to investigate the use of comparable corpora [2]. These corpora include texts which are not exact translation of each other but share common features such as domain, genre, sampling period, etc. The basic intuition that underlies bilingual lexicon creation is the distributional hypothesis which puts that words with similar meanings occur in similar contexts. In a multilingual formulation, this hypothesis states that the translations of a word are likely to appear in similar lexical environments across languages. The standard approach to bilingual lexicon extraction builds on the distributional hypothesis and compares context vectors for each word of the source and target languages. In this approach, the comparison of context vectors is conditioned by the existence of a seed bilingual dictionary. A weakness of the method is that poor results are obtained for language pairs that are not closely related. Another important problem occurs whenever the size of the seed dictionary is small due to ignoring many concept words. Conversely, when dictionaries are detailed, ambiguity becomes an important drawback.

We introduce a bilingual lexicon extraction approach that exploits Wikipedia in an innovative manner in order to tackle some of the problems mentioned above. Important advantages of using Wikipedia are:

– The resource is available in hundreds of languages and it is structured as unambiguous concepts (i.e. articles).
– The languages are explicitly linked through concept translations proposed by Wikipedia contributors.
– It covers a large number of domains and is thus potentially useful in order to mine a wide array of specialized lexicons. 

Mirroring the advantages, there are a number of challenges associated with the use of Wikipedia:

– The comparability of concept descriptions in different languages is highly variable.
– The translation graph is partial since, when considering any language pair, only a part of the concepts is available in both languages and explicitly connected.
– Domains are unequally covered in Wikipedia and efficient domain targeting is needed.

The approach introduced in this paper aims to draw on Wikipedia’s advantages while appropriately addressing associated challenges. Among the techniques devised to mine Wikipedia content, we hypothesize that an adequate adaptation of Explicit Semantic Analysis (ESA) [3] is fitted to our application context. ESA was already successfully tested in different NLP tasks, such as word relatedness estimation or text classification, and we modify it to mine specialized domains, to characterize these domains and to link them across languages.

The evaluation of the newly introduced approach is realized on four diversified specialized domains (Breast Cancer, Corporate Finance, Wind Energy and Mobile Technology) and for two pairs of languages: French-English and Romanian-English. This choice allows us to study the behavior of different approaches for a pair of languages that are richly represented and for a pair that includes Romanian, a language that has fewer associated resources than French and English. Experimental results show that the newly introduced approach outperforms the three state of the art methods that were implemented for comparison.

References:
CEA LIST’S PARTICIPATION AT MEDIAEVAL 2013 PLACING TASK

RESEARCH TOPICS: MULTIMEDIA CONTENT GEOTAGGING
ADRIAN POPESCU, NICOLAS BALLAS
SPONSORSHIP: MEDIAEVAL PLACING TASK 2013

Multimedia content geotagging is potentially useful in a wide variety of applications. Although an increasing number of existing devices include geotagging options, a wide majority of online content is still not geotagged and methods for efficient automatic geotagging are needed. In this article we describe CEA LIST’s participation to the MediaEval 2013 Placing Task [1]. In 2012 [2] we submitted runs that exploit either textual or visual information to place videos on the map. The main innovation with respect to state of the art methods was to create user geotagging models and to rely on them if useful textual annotations are missing. In 2013 we focused on improving our 2012 participation in four directions: (1) exploit a larger geotagged dataset in order to improve the quality of a standard geolocation language model, (2) model machine tags, (3) estimate the geographicity of tags associated to geolocated photos and (4) exploit user cues in order complement language models whenever these last are likely to fail.

Language models were successfully introduced as an alternative to gazetteer based geolocation and refined progressively in different editions of MediaEval Placing Task. The best performing state of the art systems combine language models and user modeling [2]. The search space in Placing Task is very wide (the physical word) and it is only partially covered by the training data provided by the organizers. We complemented this training set with external Flickr data, from which we removed all test set items, in order to study the effect of dataset size. If properly modeled, machine tags give very precise information about a photo’s location and here we propose a method to exploit them in priority. Geographicity (i.e. the geographic intent) of textual annotations was poorly studied and we modeled the geographicity of individual tags using spatial statistical technique. Finally, user modeling introduces a supplementary constraint since we need to have user data available but this condition is fulfilled for most social networks.

We have submitted the following runs, using a cascade of techniques in the order presented below: RUN1 – machine tag detection and location models based exclusively on internal training data; RUN2 – machine tag detection and location models based all training data; RUN3 – machine tag detection, location models based on all training data, geographicity and user modeling; RUN4 is RUN3 and the use of temporal cues. We present the performances of the different runs in Table 1.

The comparison of RUN1 with the other runs indicates that adding external data to language models has a positive effect on performances. In particular, RUN2 is similar to RUN1 but it exploits a much larger training set. The use of supplementary data results in more robust language models and we hypothesize that adding even more supplementary training data would further improve results. The superior performances of RUN3 indicate that adding user modeling is beneficial since precision is improved at all scales. RUN3 and RUN4 have nearly equal performances up to 10 km precision and the introduction of temporal cues is only useful at larger scales. This result is probably explained by the fact that it is usually improbable for users to move in regions of size greater than tens of kilometers.

In conclusion, in this article we approached multimedia content geotagging problem. We proposed an approach that exploits either textual or visual information to place videos on the map. The main innovation with respect to state of the art methods was to create user geotagging models (including machine tags, geographicity and user cues) to rely on them if useful textual annotations are missing. We also exploit a larger geotagged dataset in order to improve the quality of the language model. All modifications proposed this year have a positive effect. A ”standard” based only on the training data has the poorest performance, with P@1km=0.268, while P@1km=0.434 when all cues are used. This means that almost 43% of the images from flickr may be localized within one kilometer.

<table>
<thead>
<tr>
<th>Run</th>
<th>P@0.1</th>
<th>P@1</th>
<th>P@10</th>
<th>P@100</th>
<th>P@1000</th>
<th>err@1</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>0.074</td>
<td>0.26</td>
<td>0.43</td>
<td>0.5</td>
<td>0.63</td>
<td>98.8</td>
</tr>
<tr>
<td>#2</td>
<td>0.067</td>
<td>0.38</td>
<td>0.58</td>
<td>0.67</td>
<td>0.79</td>
<td>3.45</td>
</tr>
<tr>
<td>#3</td>
<td>0.133</td>
<td>0.43</td>
<td>0.62</td>
<td>0.71</td>
<td>0.81</td>
<td>2.07</td>
</tr>
<tr>
<td>#4</td>
<td>0.132</td>
<td>0.43</td>
<td>0.63</td>
<td>0.72</td>
<td>0.83</td>
<td>2.08</td>
</tr>
</tbody>
</table>

Figure 1: Multimedia content geotagging.

Table 1: P@X precision at X km. err@1 – median error at 1 km.

References:
EMBEDDED SYSTEMS
ARCHITECTURES & EMBEDDED SOFTWARE
AUTOMATIC DEPLOYMENT ON EMBEDDED PARALLEL SYSTEMS

RESEARCH TOPICS: PARALLELISM, AUTOMATION, ADEQUACY, EXPLORATION, BENCHMARKING
M. KRUMPE GOLDSZTEJN, Y. LHUILLIER, J. FALCOU (LRI), L. LACASSAGNE (LRI)

In order to be able to let developers write efficient and portable programs, we propose a programming approach based on generative programming and algorithmic skeletons.

In our approach, generative programming (or metaprogramming) is used to separate architectural specific constructs from algorithmic description.

A single algorithm description allows our tool to generate the deployment for multiple architectures, and even multiple topologies for a same target.

Although Parallel architectures have being studied for several decades the trend is still the multiplication and diversification of these architectures which also applies to their programming models. This diversity is particularly remarkable in embedded systems that are designed for specific applications. Such architectures have various power and size constraints that make parallel embedded architectures highly specialized, heterogeneous and require specific programming. This not only means complexity to program but also complicates comparing of these systems. Moreover the programmers need to learn a new programming model for every architecture.

In order to address both the issues of portability and benchmarking we have designed a tool implementing an automatic methodology for the deployment of code [1]. We propose a programming approach based on generative programming and algorithmic skeletons. The structure of our solution is presented in Fig.1 and consists of 3 levels.

The front-end is based on an algorithmic skeletons approach: the user writes a skeleton description of the application which is translated into an internal representation. At the middle-end all possible semantics is extracted using architecture agnostic transformations of the internal representation. Finally the back-end produces target architecture specific files required for compilation.

The CEDAR Architecture (Fig. 2) is a Configurable Embedded Distributed Architecture with an adaptive routing strategy based on ACO (Ant Colony Optimization).

It offers a high degree of flexibility and can handle any interconnection topology (Fig. 3). Routing paths for remote data transfers are defined at runtime and allow a homogeneous distribution of traffic, avoiding deadlocks and contentions.

The first experiments of automatic deployment with a CEDAR back-end [3] have proven that most specific codes can be generated using our approach. It is also possible to generate multiple topologies without re-writing any code.

References:
Related publications
SNET, A FLEXIBLE, SCALABLE NETWORK PARADIGM FOR MANYCORE ARCHITECTURES

RESEARCH TOPICS: MANYCORE ARCHITECTURES, SCALABLE NETWORK, ANT COLONY OPTIMISATION.
C. AZAR, S. CHEVOBBE, Y. LHUILLIER, J.-P DIGUET

ABSTRACT: A scalable communication paradigm for manycore architectures, called SNet (Scalable NETwork), is presented. It offers a wide range of flexibility by exploring the routing paths in a dynamic way, taking into consideration the network load. It is then followed by the data transmission phase through the chosen path.

One among the limits of the gigascale SoCs (System-On-Chip) will be the ability to efficiently interconnect pre-designed functional blocks and to accommodate their communication requirements in a scalable manner [2]. We propose to dedicate a set of Processing Elements (PE) to achieve routing tasks and a new PE co-processor, called DMC (Direct Management of Communications), to handle data transfers. With simple software libraries uploaded in PE memories, the user defines a specific topology, which is based on existing communication links between PEs. Then the routing is dynamic and implemented in a distributed manner in order to create paths from producer to consumer according to data dependencies. The whole concept is called Snet.

The routing strategy in Snet, implemented in the software library, is based on a distributed ACO (Ant Colony Optimization) algorithm in which dynamic paths exploration is achieved in parallel for all communicating tasks. More details on the platform and the ACO algorithm are found in [3].

We follow hereby an evaluation methodology developed in [4] to extract the characteristics of different routing topologies applied to SNet, by setting fixed simulation parameters to all tested topologies and modeling traffic with Poisson injection distribution. 16 computational nodes are mapped to the platform, as shown in Figure 1, with a varying number of routing nodes.

Each node of the platform injects 640 messages of n flits in the network, the flit width being 32 bits. In order to test the limits of the network saturation, n varies from 10 to 1000 flits, which is equivalent to 25 KB and 2.44 MB respectively. The simulations are conducted on our ISS simulator, which is based on an ISS array of RISC processors and model one instruction per cycle.

Measured latencies for the studied routing topologies under spatially uniform traffic distribution are plotted in Figure 2 for best mapping in which each two communicating tasks are placed as close as possible. In the actual implementation, the network presents a bandwidth of 9.6 Gbps at 300 MHz, considering that one transmission can be achieved simultaneously by the DMC module.

SNet efficiently implement a novel routing paradigm designed for scalability, flexibility, and ease of programmability.

It performs path exploration before data transmissions using the ACO distributed algorithm, without interrupting the application running. Results show that, though offering high flexibility and scalability, SNet manages to score high injection rate values reaching 0.9, and low message latencies. In the next project steps, time reduction of the ACO paths creation stage will be feasible by introducing specific hardware modules for ant agents propagation. Multitasking execution will increase computing efficiency and pave the way of dynamic migration.

Figure 1: Routing networks tested on the manycore platform. Dark gray refer to routing nodes, light gray to computing nodes, and white to unassigned cores.

Figure 2: Latency variation following the Poisson injection distribution.

References:
ENHANCING CACHE COHERENT ARCHITECTURES WITH ACCESS PATTERNS FOR EMBEDDED MANYCORE SYSTEMS

RESEARCH TOPICS: SHARED MEMORY, COHERENCE PROTOCOLS, MANY CORES, MEMORY ACCESS PATTERNS

J. MARANDOLA (USP), S. LOUISE, L. CUDENNEC, J-T ACQUAVIVA, D.A. BADER (GATECH)

ABSTRACT: One of the key challenges in advanced micro-architecture is to provide high performance hardware-components that work as application accelerators. In this paper [1], we present a Cache Coherent Architecture that optimizes memory accesses to patterns using both a hardware component and specialized instructions. The high performance hardware-component in our context is aimed at CMP (Chip Multi-Processing) and MPSoC (Multiprocessor System-on-Chip). We also provide a first evaluation of the proposal on a representative embedded benchmark program, which shows that we can achieve over 50% computing speedup and reduce memory throughput by nearly 40%.

Shared memory paradigms are gaining interest to program multicores systems: the main C compilers already embed support for OpenMP. Indeed, such programming concepts allow improving on legacy code to obtain a reasonable and efficient multicore support. But the age of simple multicores is reaching an end: as the number of cores grows, single buses are replaced by Networks on Chip (NoCs), distributed memory, and distributed data-paths: bus spying techniques used to ensure cache coherence are no more applicable.

With distributed caches and NoCs, the usual MESI (Modified, Exclusive, Shared, Invalid) protocol for cache coherence must be modified to refer to a given (reference) core called Home Node (HN) which tracks the MESI state of a given cache line for the whole chip. But this technique does not scale well, and is not adapted to embedded devices and applications. First, it can be very talkative as seen in Figure 1, and, second, it does not take advantage of regular memory accesses.

Such regular accesses can be represented as memory access patterns and a research effort was engaged which led to a patent deposit [2]. Improving on the baseline protocol which is the state of the art of shared memory mechanisms for manycore systems, a hardware structure and a specific protocol was designed specifically to handle the pattern based access. An example comparing the same series of data accesses for both protocols can be seen in Figure 2.

Even for such a simple pattern with only 3 elements (the difference grows linearly with the size of the pattern), the number of messages is reduced, and a speculative prefetch is done: once the first element of the pattern is detected, the remaining parts of the pattern is fetched and updated without waiting for any other memory access. Hence, future accesses are automatically prefetched and ready for use, reducing both throughput and memory latency.

A first real-size evaluation of the supposed advantage was done on a simple simulation instrumented with a in-house modified version of the pintrace Pintool memory analyzer, from Intel’s Pin framework. We showed on that on a two-pass image filter that was chosen for it stresses memory accesses, we obtained a reduction of 37% of message throughput and an acceleration of the application by more than 50% with regards to the baseline protocol alone.

Hence, this protocol, taking advantage of regular memory accesses (patterns), was validated on a program representative of embedded applications. The results show that such an apparatus significantly reduce message and memory throughput and accelerate applications. Such breakthrough can be vital for the future of manycore systems, their programmability and their performance.

Figure 1: A write message transaction with the baseline protocol.

Figure 2: Comparison between baseline protocol and the pattern approach (speculative-hybrid protocol).

References:
AN OPTIMIZATION APPROACH FOR THE SYNTHESIS OF AUTOSAR ARCHITECTURES

RESEARCH TOPICS: DESIGN, AUTOMOTIVE SYSTEMS, SCHEDULING, DEPLOYMENT, OPTIMIZATION

C. MRAIDHA, S. TUCCI PIERRIOVANNI

The synthesis of automotive architectures is a complex problem that needs an automated support. AUTOSAR, standard for the specification of automotive architectures, defines a synthesis process of software components and their connections in a set of fixed-priority OS tasks distributed over a network of ECUs.

The synthesis process, according to AUTOSAR, can be divided into four main activities. The first one is called allocation as it consists in placing atomic software components on the ECU-s and exchanged signals on the buses. Each software component encapsulates the implementation of a specific functionality which is exposed to the outside world by means of ports. Internal behaviour of each atomic software component is represented by a graph of runnable entities, which in turn represent schedulable units of computation. Second, third and fourth activity are partitioning, scheduling and ordering. Runnables and signals are being partitioned in OS tasks and messages which are then scheduled by the assignment of static priorities. Moreover the order of runnables inside a task is defined with respect to the functional constraints.

It is common and recommended practice in this demarche, to specify end-to-end timing constraints at the highest level, between input and output ports of the highest level component of the architecture, usually representing the system under study (see Figure 1). For each external stimulus, consumed by an input port, the constraint specifies a deadline for the response to be produced on the output port. Each end-to-end constraint is progressively refined by specifying the end-to-end chain of runnable entities (traversing one or more atomic components) that are activated to produce a system response triggered by the given stimulus.

By knowing runnables allocation their partitioning and order within the tasks as well as priorities of the tasks, the designer can now compute the response times of runnables end-to-end chains to see if end-to-end timing constraints are met. If some deadlines are violated, the designer has to find another configuration. This is a cumbersome process as the synthesis is an NP-hard problem. Hence appropriate support is required. In the current state of practice, only partial solutions exist as none of them handle all four dimensions (allocation, partitioning, scheduling and ordering) at once. Most common solutions take as input a task model which means that the partitioning is already known or is done manually based on an engineer’s expertise. This severely minimizes the design space to explore and therefore might exclude feasible solutions.

CEA LIST has proposed two techniques for the synthesis of AUTOSAR architectures in its entire form. The first technique is based on mixed integer linear programming (MILP). It returns the optimal solution but is limited to small size systems. In order to improve scalability – to address industrial size systems - a second technique, based on genetic algorithms (GA) is proposed. For both techniques optimization criteria relate to end-to-end responses and consumed memory.

Techniques proposed by CEA LIST allowed to obtain better results than those obtained by current partial solutions Figure 2 shows that our GA solution is able to find architectures with better response time (fitness value) than those obtained with current approaches, improving response time, respectively, by 34.87% and by 16.48% with respect to approaches excluding partitioning and considering only partial partitioning.

References:
We generating code from application models, based on executable components, specifying their structure, behavior and interactions. The objective of the approach is to support a set of concepts that enable application components that are independent of the underlying OS and middleware via an extensible tool chain.

We start with a description of the system model, since its properties have an impact on code generation. This model consists of three parts: a model describing the software components, a model describing the hardware platform and a model that describes component instances, including their configuration and allocation to hardware components.

The software structure is described by a de-composition of the system into components. These may interact with each other via connectors linking interaction points that are owned by the components. Well-formed models must respect the compatibility between ports. The component behavior has two different aspects: The algorithmic aspect describing the services provided by a component and the control aspect describing the orchestration of calls as well as different modes of a component. The control aspect is typically modeled by a state machine, the algorithmic aspect via action languages.

If the hardware structure is modeled in a rather similar way as the software structure: By means of hierarchical components. In this case, connectors represent physical connections and ports physical interaction points such as I/O ports.

The task of the deployment part is twofold: (1) it allows to identify to which hardware element a software component is allocated and (2) it allows to configure instances of software components. For each hardware component, a subset of the original model, source code and finally the binary compiled from this code is created. The source code generation has to support the different ways in which behavior is described, i.e. action languages and state-machines (while preserving their execution semantics).

The complexity of the code generation is reduced by separating it into different tasks: the support of state-machines, component concepts (connectors and ports) and the code-generation from the “OO-subset” of UML. For the latter, existing standard code generators can be used. The OO subset is created from the application model by a sequence of automated model transformations that are executed immediately before code generation.

The figure below shows the different steps for the generation of a binary starting with a software component model, a platform model and a deployment description. It is supported by an extension of the open source UML modeler Papyrus called Compass Designer.

Separating the steps towards code generation has an additional advantage: it enhances the flexibility as different parts of the chain might be exchanged. In the presented approach, the flexibility is primarily achieved by feeding model libraries as input in the container and connector transformations, not by changing the transformation chain itself. For instance, in case of the container transformation, an application component applies a container rule. This rule, typically defined in a model library controls which elements are added to the container of the application component.

In case of the OO transformation, the default transformation creates an operation for each consuming (“in”) flow port. Yet it might be useful to change this mapping for data that is consumed by a flow port: the robotic middleware Orocos uses a mapping in which all consuming ports are mapped to a single operation (which has a separate in-parameter for each flow port). While keeping a generic component to OO transformer, the transformation takes the underlying pattern as input.

References:
In a software development context of real-time embedded systems, the designer makes various architectural choices, at the design level, to describe the realization of their software part. Then, a verification of timing properties is performed to assess these choices. This verification step requires an abstraction of some information related to the underlying Real-Time Operating System (RTOS) such as scheduling policy, communication mechanisms, etc. In fact, the design model is a Platform-Independent Model (PIM), thus most of the verification tools used to validate this model make assumptions about the target RTOS and consider a software platform offering unlimited resources. In that case, the refinement of the design model to an RTOS-specific model, which corresponds to a deployment phase, is a non-trivial transformation because the assumptions made may not be verified for the targeted RTOS.

CEA LIST has proposed a model-driven approach [1] to guide the designer by proposing solutions whenever the refinement of the original design model to an RTOS-specific model is not feasible. Each proposed solution is based on a software deployment pattern. A pattern is defined as a model transformation aiming to solve a particular deployment problem under some assumptions considered in the original design model.

CEA LIST in collaboration with the Université de Bretagne Occidentale and the Ecole Nationale d’Ingénieurs de Sfax have initiated the population of a software pattern base that allows to deal with the most common problems encountered in the deployment of real-time design architectures on different real-time operating systems families.

Moreover, this methodology integrates a refactoring phase [2] whose objective is to guide the designer by proposing solutions after the pattern application. This approach illustrated on Figure 1 called Design Refinement toward Implementation Methodology (DRIM), integrates two steps: (1) a deployment feasibility tests step and (2) a mapping step.

The choice of the pattern to apply for a given deployment problem can have an impact on the real-time system performance. For this reason the refactoring integrates a performance evaluation phase that consists in estimating the impact on the performance after the pattern application.

This work has contributed to the automation of real-time design models refinement and opens good perspectives for the RTOS dependent optimizations of real-time software architectures.

Figure 1: DRIM process overview.

Figure 2: Refactoring phase.

References:
A STEP-FORWARD IN FORMALIZING UML: PRECISE SEMANTICS OF UML COMPOSITE STRUCTURES (PSCS)

RESEARCH TOPICS: EXECUTION SEMANTICS, MODEL-DRIVEN ENGINEERING
A. CUCCURU
PARTNERSHIP: THALES, AIRBUS, INCOSE, NASA JPL

Model Driven Engineering helps in managing design complexity of modern real-time systems through models specified at appropriate abstraction levels. Being able to execute such models naturally contributes to management of complexity. By observing an execution, engineers can more easily figure out the constraints that models place on the runtime structure and behavior of the final system. Observations can be used to make relevant decision choices, from the very early steps of the system design flow.

Developing tools for execution of models requires semantics of the underlying formalisms to be unambiguously specified. In a joint effort with Thales, Airbus, INCOSE or NASA JPL (among other participants), the LISE took the initiative to formalize the need for a new standard specification that would address this subject [1], for one of the most widely used standard modeling language: the Unified Modeling Language (UML). After issuance of this RFP, the LISE took the lead of a task force (essentially composed by the actors that where involved in the preparation of the RFP), which has been working two years on the elaboration of a joint response to this RFP.

The resulting specification [2] finally got adopted by the OMG in March 2014. This specification deals with the execution semantics of Composite Structures, a key UML concept to deal with component-oriented modeling. In order to keep homogeneity and ensure complementarity with existing standards, this specification extends fUML (foundational UML), which defines execution semantics of a UML subset with a formal execution model. The extension strategy of fUML, depicted in Figure 1, consists in extending the UML subset in order to include UML Composite Structures and then extend the execution model accordingly.

This specification not only specifies precise execution semantics for this important subset of UML. It is also the first OMG specification to include a normative suite of test cases. This test suite consists in a set of executable models (defined within the subset of UML identified by this specification) whose execution by conforming tools shall satisfy properties which are explicitly part of those models. Tool vendors can then use these executable models to demonstrate conformance to this specification, by successfully executing all the tests with their own implementation.

The LISE is now chairing the finalization task force of this specification. The availability of a test suite will be a major advantage for future evolutions of the specification (bug corrections, improvements), since it will enable a test-driven methodology for the evolution of the specification itself. It will typically avoid regression issues, which are usually faced in evolutions of documents as complex as OMG specifications.

While PSCS is definitely an OMG specification on its own, it shall really be seen as a second volume in a series of specification aiming at formalizing all the executable parts of UML. Currently, the LISE is thereby strongly involved in the preparation of a roadmap on these subjects. The next major step in the area will be the formalization of UML state machines, the most suited UML concept for the description of state-based behaviors.

In support of these standardization efforts, the LISE develops Moka, a model execution environment integrated in the UML/SysML Papyrus modeler (also developed by the Lab). Moka includes an implementation of the fUML execution model, as well as its extensions for composite structures, and already enables simulation of systems modeled within this subset of UML.

References:
**MARTE, A STANDARD MODELING LANGUAGE DEDICATED TO CPS**

RESEARCH TOPICS: MODEL-DRIVEN ENGINEERING, CPS, COMPLEXITY

SÉBASTIEN GÉRARD

PARTNERSHIP: BRAN SELIC (MALINA SOFTWARE CORP.)

Interest in model-based methods for designing software has grown significantly since the introduction of the Unified Modeling Language (UML) and its adoption as an industry standard in the late 1990’s. Quite naturally, this interest has extended to the real-time domain, resulting in the adoption of several real-time oriented modeling language standards based on UML. The first of these, the UML Profile for Schedulability, Performance, and Time [1], was supplanted by the more advanced MARTE standard [2, 3], which was aligned with the most recent major revision of UML, UML 2 [4].

MARTE is an abbreviation for Modeling and Analysis of Real-Time and Embedded systems, and is a profile of UML. It is a domain-specific modeling language intended for model-based design and analysis of real-time and embedded software portions of cyber-physical systems. It is designed as a supplement to UML, providing capabilities that are either inadequate or missing in UML.

A UML profile is a domain-specific interpretation of the general UML language that specializes some of its concepts to reflect domain phenomena and concerns. Because a properly defined profile is conformant to the rules and concepts of standard UML, it has the potential to reuse some of the existing UML tools, methods, and expertise. Since MARTE is defined as a profile of UML, the integration of the two languages is seamless.

MARTE provides broad coverage of the real-time domain and is the result of the collective work of numerous domain experts.

MARTE extends UML then with the ability to specify precisely various kinds of physical values, such as time-related information, resource limitations and capacities, etc. Needless to say, these play a crucial role in the design of any real-time system. Finally, concluding Part II is a description of two topics that are the very core of MARTE and which provide the essential bridge between the logical world of software and the physical world in which it must operate: the MARTE representations of time and resources.

MARTE can be used to accurately model those parts of cyber-physical systems that involve software. This covers not only modeling of the application, but also modeling of the underlying platform, as well as specifying the relationships that exist between the two (i.e., deployment).

MARTE can be exploited for analyzing designs in order to predict or validate their run-time characteristics, such as their performance and timing properties as shown in the following figure. MARTE provides a generic framework that supports a broad category of possible types of analyses that conform to a common pattern. Two of the most common of these, schedulability and performance, are directly supported by the standard.

Finally, users of MARTE working within a particular application domain may find that MARTE is not always precise or complete enough to meet their needs for expressing the concepts in their particular project or application domain. For example, while MARTE provides facilities for modeling operating system threads, these are relatively general, covering the common features of threads across a broad range of different operating systems. However, when it comes to implementation-level detail, most operating system implementations of threads differ from each other. If there is a need for models to include platform-specific details (e.g., for automated code generation), then a general MARTE concept will have to be extended to incorporate the missing information.

---

**Figure 1: A generic schema for model-based analysis**

---

**References:**


MODEL-DRIVEN SAFETY ASSESSMENT OF ROBOTIC SYSTEMS

RESEARCH TOPICS: SAFETY ASSESSMENT OF ROBOTIC SYSTEMS

NATALIYA YAKYMETS, AGNES LANUSSE, SAADIA DHOUIB

PARTNERSHIP: ALDEBARAN ROBOTICS

Modern robotic systems (RSs) are complex and capable to perform sophisticated tasks in different domains. In order to cope with system complexity, engineers consider new approaches to system development based on formalized modeling. Model driven engineering (MDE) is expected to significantly simplify the support of system requirements, design, analysis, verification and validation through a system life-cycle. The use of system level models enables simulation of the overall performance and behavior of complex RSs. In addition, MDE gave birth to domain specific languages like Robotic Modeling Language (RobotML) [1] to target system development in different domains. RobotML is built as a UML profile to design, simulate and deploy robotic applications. With RobotML, RSs can be defined using appropriate notations, abstractions and facilities to automatically generate executable code. RobotML could be a solution for robotics experts to deal with variability problems and to hide the lower level programming details.

Our research has been inspired by RSs for safety-critical applications. Such RSs are expected to satisfy a high level of safety. Standards concerned with the development of safety-critical systems require an application of specific design flows where system engineering is conducted in parallel with safety assessment (SA). This allows the concept, design and implementation of safety-critical RS to be developed with respect to the safety aspects. Each phase of SA flow implies application of a set of specific methods and activities. Although these well-established methods provide an efficient support for safety engineers, they could greatly benefit from a tighter coupling with system modeling environments.

MDE offers facilities to annotate models with information needed for SA, to perform validation according to dedicated rules, to write transformation rules towards formal languages to permit their analysis by formal tools. It becomes thus possible to perform model driven SA by incorporating existing SA methods and tools into a uniform MDE environment.

We propose a methodology [2] for model driven analysis of RSs in the preliminary SA phase (Figure 1). The safety standards for robotics are not mature enough and some of them have been still under development. Therefore the methodology follows the IEC61508, a generic standard on functional safety design, and describes SA using FTA approach. It leverages features of RobotML (i) to capture information required for formal analysis (ii) to propagate SA results back into the MDE environment. The use of the proposed methodology allows the safety engineer to start SA from the early phases of RS development which can significantly reduce time and cost constraints.

First, a RS is designed with the Papyrus platform using RobotML language. Second, we define the sufficient finest level of RS architecture where SA will be conducted and then annotate a RobotML model with the possible failure behavior at this level. While defining failure modes of the components, information on the possible hazards derived from the hazard analysis is taken into account. RS dysfunctional behavior is annotated using analytical expressions. Once the annotation has been done, the failure states and events related to the component failure modes are automatically extracted and the RobotML model is converted into the AltaRica language. The checking of the AltaRica model is performed by the ARC tool. This tool also computes minimal cut sets for the considered model. Based on this information we automatically generate fault trees (FTs) and represent them in the Open-PSA format.

Finally, we perform FT qualitative and quantitative analysis according to IEC61508 and compute a set of factors (like probability of the top FT event, contribution of minimal cut sets, etc.) to evaluate system safety. In order to make SA results more representative, we display FTs in RobotML modeling environment using dedicated FT profile.

References:
MOKA: AN OPEN AND FLEXIBLE FRAMEWORK FOR MODEL EXECUTION IN PAPYRUS

Since many years, the LISE has been leading the development of the reference open-source UML/SysML modeler Papyrus, in the context of the Eclipse community. The Papyrus offer has been recently enriched by a new module for model execution: Moka. This module includes aspects related to control of model executions (e.g., suspending/resuming executions after breakpoints have been encountered) as well as to observing states of executed models at runtime (e.g., emphasizing graphical views of model elements on which execution has suspended, retrieving and displaying any state information about the runtime manifestation of these model elements). The ability to execute models and observe their execution can help engineers to make relevant modeling choices (and thereby improve the quality of their models), from the very early steps of the system design flow.

Moka comprises two main parts. The first part concerns a generic API and framework contributing to and extending the Eclipse debug API, with the idea that elements to be executed, controlled, and observed are EMF models and model elements instead of programs and program statements. The second part uses these API and framework to provide basic execution and debugging facilities for foundational UML (fUML [1]) and its extensions for composite structures (PSCS [2]), an executable subset of UML with precise and standard semantics. A continuous activity underlying the development of this second part consists in improving these basic execution capabilities with appropriate model libraries, where new control and observation functionalities are brought by fUML compliant models rather than tool-level extensions of the execution engine.

One of the motivating ideas behind this approach is to limit as much as possible modifications to the core execution engine, in order to preserve by construction compliance with standard fUML semantics. An example has been published in [3]. The purpose of this article is to sketch the principles for designing control entities using fUML and apply a simple design pattern to extract control from system models so that it can be delegated to these control entities. These basic principles are the key to any simulation environment.

By using these principles, we have designed a fUML model library that can be used to constrain execution orders inside a fUML model according to the semantics of the Discrete-Event (DE) MoC. In addition to introduce additional control mechanisms in the execution model, this article has also demonstrated that it was feasible, using this approach, to introduce a model of time in the execution without any modification to the implementation of the tool, even if fUML is time-agnostic.

A similar work has been done to prototype an execution environment for the ROOM profile (which is extensively used by our partner Ericsson) or some subset of standard profiles such as SysML and MARTE. The approach has also been applied to prototype co-simulation environments with Ptolemy and Simulink. Indeed, the execution model of fUML has been extended through this approach to solve typical co-simulation problems such as synchronization and communication.

The first official release of Moka is scheduled for June 2014, in conjunction with the launch of Papyrus 1.0. Based on more recent research work [4], a second version is already under development. This version will integrate facilities and automated tooling to enable a parameterization of the execution engine with fUML models capturing the semantics of UML profiles.

References:
A NEW APPROACH FOR REAL-TIME TASKS SYNTHESIS AND THEIR PLACEMENT ON DISTRIBUTED ARCHITECTURES

RESEARCH TOPICS: OPTIMIZATION, PLACEMENT, PARTITIONING, SCHEDULING

S. TUCCI PIERGIOVANNI, C. MRAIDHA

PARTNERSHIP: SCUOLA SUPERIORE SANT’ANNA (ITALY), MCGILL UNIVERSITY (CANADA)

The emergence of distributed embedded architectures is a phenomenon observed in very competitive domains like automotive, infotainment systems or telecommunication. This emergence strengthens the problem of resources optimization while keeping or even increasing the quality of service requirements of these systems.

In practice, as illustrated on Figure 1 the design of these architectures relies on three main levels:

– The application functions level;
– The (hardware and software) platform resource level (computing and communication resources needed to realize the functions);
– The intermediate tasks level (software resource used to virtually parallelize treatments on the same resource).

The design of the final system call for the realization of two allocation levels and a validation:

– The placement consists in the allocation of functions to platform resources;
– The partitioning consists in the allocation of functions to tasks (to decide on the appropriate level of parallelization)
– The scheduling that configures tasks priorities to guarantee timing requirements of functions with available platform resources.

As we can see in Figure 2, to face the combinatorial order of this multi-level optimization problem (which is dependent of the number of functions, signals, processors and buses), existing approaches used to treat only one or two parts of this problem, but not the problem in its integrity. This partial resolution leaves the designer handle the integral optimization problem with a trail-error approach that might be rather expensive in terms of development time without any guarantee on the quality of the final result.

The approach [1] developed by CEA LIST in collaboration with two international experts from Scuola Superiore Sant’Anna and McGill University, is a multi-staged optimization approach using at each stage either mathematics formulations in Mixed Integer Linear Programming (MILP) or meta-heuristics implemented using Genetic Algorithms. This approach allows the global treatment of this optimization problem (placement, partitioning and scheduling). Its scalability was demonstrated on systems with hundreds of functions deployed on tens of processors, while non-multi-staged approaches based only on exact mathematical formulations are able to treat only cases with two or three processors.

References:
SOPHIA FRAMEWORK FOR MODEL-BASED SAFETY ANALYSIS

RESEARCH TOPICS: SAFETY ANALYSIS
NATALIYA YAKYMETS, AGNES LANUSSE, YUPANQUI MUNOZ JULHO

Standards concerned with the development of safety-critical systems define specific design and safety assessment flows. According to safety standards, system engineering is conducted in parallel with various safety analysis (SA) activities. This allows the concept, design, and implementation of safety-critical systems to follow and respect the safety aspects. Each phase of SA flow implies application of specific methods and activities including hazard analysis, fault tree (FT) generation and analysis (FTA), failure mode and effects analysis (FMEA). Generally, safety assessment activities are conducted within environments decoupled from the design. As a result, information on system model is captured manually from textual documentation. These well-established methods have been used for many years by safety engineers; however, they are error-prone and drastically time consuming in the case of complex systems. To overcome these drawbacks, a first idea is to capture information automatically from design models so that safety analysis can be performed on an image of the design model, avoiding useless double capture of information and improving the quality of the produced safety models. Such an approach is defined as Model-Based (MB).

Our research focuses on integration of SA techniques into a MB system engineering (or MBSE) environment, as well as methodological approaches to conduct model-based SA in this context. We propose a methodology and associated framework, called Sophia, for MB SA of complex systems [1] [2] [3]. Sophia leverages features of SysML modeling language and includes facilities to perform FTA and FMEA in the uniform MBSE environment supported by the Papyrus editing tool. Figure 1 gives an overview of the proposed methodology. The methodology includes several steps.

Step 1. A system is analyzed by the safety engineer who defines the SA objectives, methods to be applied, safety evaluation metrics, as well as the finest architectural level sufficient to perform SA in terms of time, cost, safety requirements, etc.

Step 2. The system architecture is expanded till the selected SA level and components are annotated with the possible failure modes. While defining failure modes of the system components, information on the possible hazards derived from the hazard analysis is taken into account. Failure modes are introduced by either analytical expressions or state machines.

Step 3. Once the annotation has been done, the failure states and events related to the failure modes of the system components are automatically extracted, and the model is converted into a formal language like AltaRica, SMV, etc. This model can be further analyzed using various SA methods such as FTA, FMEA, etc. recommended in IEC61508.

Step 4. SA starts from the granularity of the top-level representation of a system. To assess system safety, we define a set of metrics, as well as their boundary values. These values are used as stop criteria during SA: if values of safety metrics estimated for a given component satisfy the stop criteria then SA is not conducted for this component; otherwise, a component is expanded for further SA. Such an iterative process continues until the granularity of the selected SA depth is reached for all safety critical components. These components might be further analyzed until the finest level of system hierarchy is reached. In this case, the SA depth is updated and new information on component failure modes should be added into the model if needed.

Step 5. To describe system failures and then to derive the obtained SA results back into the modeling environment, we exploit the UML extension mechanism based on profiles. The use of UML profiles allows separation of different SA aspects related to various SA methods (FTA, FMEA, etc.) and types (qualitative and quantitative analysis).

References:
Most approaches to model based testing (MBT) assume that a single tester interacts with the system under test (SUT) and this tester observes all inputs and outputs. However, many systems such as web services and wireless sensor networks interact with the environment at multiple physically distributed interfaces. This has led to interest in distributed testing, where there is a separate local tester at each interface, a tester only observes events at its interface, and there is no global clock.

This approach to distributed testing was formalized by ISO as the distributed test architecture. It is known that the use of the distributed test architecture affects software testing. However, only recently has the effect been formalized as implementation relations for FSMs [HU08] and IOTS [HMN12]. Previous works showed that the distributed test architecture causes additional controllability and observability problems. A controllability problem is a situation where a local tester does not know when to apply an input. Observability problems refer to situations where one cannot distinguish between the global trace produced by the SUT and that expected despite these being different.

In [GHL13], conjointly with Rob Hierons of the Brunel University and Pascale Le Gall from ECP, we have explored how to extend the previously cited approaches to the case of timed distributed systems. Distributed system specifications are given as tuples of cooperating timed input output transition systems (TIOTS): IOTS extended with time.

We assume that testers have local clocks that are not synchronized but clocks progress at the same rate; it should be straightforward to adapt this to the case where the clocks can drift.

We characterize observations (as well as behaviors of system specifications) as tuple of traces, called multitraces, where each element of the tuple denotes an observation of a tester at a given interface. Instead of defining the implementation relation with some projection mechanisms, we directly use the notion of multitraces to define a new implementation relation dtioco for distributed testing.

Reasoning with timed multitraces rather than simple untimed traces with projection techniques (as it was done in [HU08] and [HMN12]) has the advantage of solving some observability and controllability problems, which permits to use already existing centralized testing techniques ([BEGL12]) at each interface of a distributed system in order to perform distributed testing. Indeed, a compositionality result says that dtioco holds if and only if all multitraces of the SUT are such that exchanged messages respect some communication rules and also that the local executions of the SUT conform to their associated local components of the specification. Thus, solving the oracle problem mainly becomes a multitrace analysis problem.

References:
GUARANTEED INTEGRATION OF ORDINARY DIFFERENTIAL EQUATIONS USING AFFINE ARITHMETICS

RESEARCH TOPICS: STATIC ANALYSIS OF CYBER-PHYSICAL SYSTEMS
O. BOUISSOU, S. MIMRAM, A. CHAPOUTOT (ENSTA PARISTECH)
PARTNERSHIP: ENSTA PARISTECH

Cyber-physical systems are characterized by a strong interaction between a discrete system (the software) and a continuous plant (the system to be controlled). For the analysis of such systems, it is necessary to be able to safely predict the trajectories of the continuous plant. Most generally, the trajectories of the plant are implicitly defined by a set of ordinary differential equations (ODEs). So, to safely predict the trajectory of the plant, one must compute guaranteed bounds on the solution of ODEs, which is the goal of the techniques of guaranteed integration of ODEs.

There are mainly two kinds of such techniques. First, for linear ODEs, one can explicitly solve the ODE and use precise and efficient representation of sets to compute over-approximations. This is done in the SpaceEx tool for example. The second technique, applied for non-linear equations, consists in using interval arithmetics and the Taylor expansion of the ODE to compute safe bounds on its solution. For example, the VNODE tool uses this technique.

These two methods present drawbacks that limit their application for the analysis of real-world control-command systems: SpaceEx only considers linear ODEs while VNODE only computes boxes that contain the solution of the ODE and does not compute relations between the variables of the plant.

In [1], we propose a new method to perform guaranteed integration of non-linear, non-polynomial ODEs that uses affine arithmetics to introduce relations between the variables of the plant. Affine arithmetics is an extension of interval arithmetics, first introduced by Stolfi in 1994 that allows for implicit relations between variables by using common symbolic variables (called noise symbols) that range in the interval [-1,1]. For example, if we have three variables x, y and z, and two noise symbols e1 and e2, then the expressions

\[
\begin{align*}
    x &= 3 - 2e_1 + e_2 \\
    y &= -1 + e_1 - 0.5e_2 \\
    z &= -14 + 10e_1 - 5e_2
\end{align*}
\]

represent the set of all tuples (x,y,z) such that

\[
\begin{align*}
    x &\in [0,6], y \in [-2.5,0.5], z \in [-29,1], \\
    2y + x &= 1, 4x - 2y + z = 0.
\end{align*}
\]

The method proposed in [1], which was then implemented in the HySon tool [2], uses affine arithmetics and variations of the well-known Runge-Kutta numerical methods to obtain guaranteed enclosures of the solution of any ODE.

These two innovations (using numerical methods and affine arithmetics) make our method perform very well on stiff equations and complex dynamics. For example, we could solve the non-linear ODE

\[
\begin{align*}
    y' &= z \\
    z' &= (z \cdot z) - (3 / (0.001 + (y \cdot y)))
\end{align*}
\]

up to t=50s, while VNODE cannot go beyond t=20s.

Moreover, as we use affine forms, we introduce more and more relations between the variables of the system allowing deriving complex invariants on the trajectories. For example, for the ODE

\[
\begin{align*}
    u' &= v \\
    v' &= (u^2)
\end{align*}
\]

starting from \(u\in[0.95,1.05], v\in[-1.05,-0.95]\), the trajectories become ellipsoidal, as shown in the picture below.

In [1], we develop a prototype tool that applies these ideas on a text-based description of differential equations. We showed that the use of numerical methods outperforms the state-of-the-art tool for guaranteed integration of ODEs.

In the HySon tool presented in [2], the ODEs can be represented as Simulink models and the guaranteed integration is linked with a simulation algorithm for hybrid systems that allow computing precise bounds of the trajectories of a complex, cyber-physical system.

An extension of these methods to Differential Algebraic Equations (DAEs) is currently studied in collaboration with ENSTA ParisTech.

References:
A HOMOTOPICAL COMPLETION PROCEDURE WITH APPLICATIONS TO COHERENCE OF MONOIDS

RESEARCH TOPICS: REWRITING THEORY, ALGEBRA
Y. GUIRAUD (UNIV. PARIS 7), P. MALBOS (UNIV. LYON 1), S. MIMRAM
PARTNERSHIP: UNIV. PARIS 7 ET LYON 1

In mathematics, one often works with monoids, which are an algebraic structure consisting of a set equipped with a way of multiplying its elements. Monoids are generally infinite, and in order to perform computations on them (to compute some of their properties, such as their homology for instance) a finite description of them is usually provided by a presentation, which consists of a set of generators and relations generating the monoid. However, these are not very easy to work with directly since they require working with equivalence classes of words: one would rather like to have canonical representatives of those equivalence classes. When the relations of the presentation can be oriented in order to obtain a rewriting system which is convergent (i.e. both terminating and confluent), the normal forms of the rewriting system (words that cannot be reduced by the rules) provide canonical representatives of the words modulo the relations. Algorithms computing convergent rewriting systems are thus fundamental in order to study monoids and the present work improves existing ones by extending them in three directions: we consider changing the generators of the presentation, we compute a homotopy basis, and we can remove superfluous generators as explained below.

One of the most used algorithm in rewriting theory is the Knuth-Bendix completion procedure which starts from a terminating rewriting system and iteratively adds rules to it, trying to produce an equivalent convergent rewriting system. Here, we are interested in extending this procedure in order to retrieve information about the low-dimensional homotopy properties of a monoid. We therefore consider the notion of coherent presentation, which is a generalization of rewriting systems that keeps track of the two-dimensional cells generated by confluence diagrams. We extend the Knuth-Bendix completion procedure to this setting, resulting in a homotopical completion procedure. This algorithm is based on a generalization of Tietze transformations, which are a complete set of operations that can be iteratively applied to relate any two presentations of the same monoid. These transformations consist in adding (or removing) both superfluous generators and relations. We first noticed that only the second kind of operations is used in the Knuth-Bendix completion procedure. However, we show that allowing a change of generators can enable one to make a presentation convergent: as witnessed by the completion of A2 below, there are some monoids which admit no finite convergent rewriting system on a given set of generators, but do admit one if we do add generators. Moreover, Tietze transformations can be used to remove useless generators, rules, or confluence diagrams in a coherent presentation, thus leading to a homotopical reduction procedure which makes a presentation smaller and thus computationally more tractable. Finally, we applied these techniques to the study of some examples coming from representation theory, to compute minimal coherent presentations for them: braid, plactic and Chinese monoids.

This work gave rise to an implementation, which has since then been used by various other researchers in order to experiment with conjectures about some properties of given monoids. Below, an example of the use of the software is given on a simple example, computing a completion of the monoid A2, by first adding a generator, and computing a homotopy basis.

Generate the Artin monoid of type A2, also known as the braid monoid.
# artin a2
a,b | A : bab -> aba
We have 1 generator and 1 rule.
Add a new generator.
# rgen ba
a,b,ba | A : bab -> aba, B : ba -> (ba)
We have 3 generators and 2 rules.
Perform a Knuth-Bendix completion.
# kbc
a,b,ba | B : ba -> (ba), D : (ba)b -> a(ba), E : a(ba)a -> (ba)(ba), G : (ba)(ba)a -> b(ba)(ba)
We have 3 generators and 4 rules.
Compute homotopy generators.
# hbasis
ba(ba)a --B(ba)a--> (ba)(ba)a --G--> b(ba)(ba)
ba(ba)a --bE--> b(ba)(ba)
(ba)ba --Da--> a(ba)a --E--> (ba)(ba)
(ba)ba --(ba)B--> (ba)(ba)
[...]
We have 4 homotopy generators.

The homotopical reduction procedure further reduces the number of homotopy generators to two instead of four.
TIME REVERSAL REFLECTOMETRY FOR CABLE AGING CHARACTERIZATION

RESEARCH TOPICS: REFLECTOMETRY, TIME REVERSAL, CABLE AGING
L. EL-SAHMARANY, F. AUZANNEAU AND P. BONNET (UNIVERSITE BLAISE PASCAL, CLERMONT-FERRAND)

ABSTRACT: We investigate the effects of aging (i.e. slow homogeneous degradation) on electrical cable characteristics by the use of a new method based on time reversal. In case of a global cable aging, the commonly used methods such as reflectometry provide non-relevant or inaccurate information. Through theoretical study and numerical simulations, the benefits of this new method called Time Reversal Reflectometry (TRR) are presented. TRR is experimentally shown to be successful for the detection and quantification of cable aging.

Aging is described as a slow structural modification which gradually decreases the efficiency of an object, information or organism to provide its functions. Therefore, this paper overcomes reflectometry's limitations by proposing a new approach based on time reversal applied to reflectometry’s fundamental principle. It focuses on the detection and estimation of electrical cable aging.

This new method is based on the principles of time reversal and standard reflectometry method [1]. Instead of using a predefined signal (Gaussian pulse) like standard reflectometry method, it uses an adapted signal that allows characterizing more precisely the cable’s electrical parameters (RLCG) modifications due to aging. The adapted signal will be insensitive to dispersion which distorts signals and decreases the ability of cable aging detection and estimation. The detection of cable aging using time reversal is summarized by the following process:
1) Inject a (symmetrical) pulse signal into a healthy cable.
2) If needed, truncate and shift in time, then normalize the reflected signal.
3) Apply time reversal, and then save as adapted signal.
4) Inject the adapted signal into the aged cable.
5) Process the reflected signal, calculate the Skewness Coefficient noted SC, and estimate the cable aging. SC is calculated by quantifying the signal’s distortion on the left “a” or the right side “b” of its maximum as presented on Fig.1. Then, SC = b/a. A value of SC close to 1 means the cable under test is healthy. Otherwise, if the reflected signal is asymmetrical or SC value is far from 1, this means the cable is aged and the value of SC enables to quantify the aging.
6) Loop steps 4 and 5 when needed.

In order to investigate this method a comparison by changing the per unit length capacitance value was performed, the simulation is done by using a RLCG frequency model of a cable using MATLAB®. The reflected signal from step 5 for the healthy cable (Capacitance is C0) and for three simulated aged cables (Capacitance values of aged cables are 0.2, 0.6 and 1.2 times C0) were calculated. Table I presents the values of SC and shows the effect of aging. When the cable is healthy, SC is equal to 1 and when it is aged (0.2 * C0) SC is down to 0.5127.

Thermal aging experiment was performed on a 100 m long coaxial cable subject to thermal aging [2]. Table 2 shows the effect of aging via the variation of the skewness coefficient. It was noted that the increase of SC with time led the reflected signal to lose its symmetry (as illustrated in Fig.2).

The proposed method presents a simple and more accurate technique to estimate cable aging. It can help monitor the health of the cables and the safety of an entire electrical system.

References:

Table 1: Values of skewness coefficient SC for different simulated aged cables

<table>
<thead>
<tr>
<th>Simulations</th>
<th>1.2*C0</th>
<th>C0</th>
<th>0.6*C0</th>
<th>0.2*C0</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC</td>
<td>1.09</td>
<td>1</td>
<td>0.7943</td>
<td>0.5127</td>
</tr>
</tbody>
</table>

Table 2: Values of skewness coefficient during aging

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC</td>
<td>1</td>
<td>1.28</td>
<td>1.29</td>
<td>1.31</td>
</tr>
</tbody>
</table>
ON A USEFUL TOOL TO LOCALIZE JACKS IN WIRING NETWORK

RESEARCH TOPICS: WIRED NETWORK DIAGNOSIS, SOFT FAULT DETECTION
M. FRANCHET, N. RAVOT, O. PICON (UNIVERSITÉ PARIS-EST, ESYCOM)

ABSTRACT: To efficiently monitor and maintain wired networks, their topology has to be known. Most of the time a wiring network is made of several cables linked to each others with connectors. So knowing where they are is valuable information for going back to the topology. Then the damaged portions of the network can be localized relatively to the jacks, which will facilitate and accelerate maintenance. New data processing techniques based on a time frequency transform are shown to improve the detection capacity of both soft defects and connectors in wires.

Electrical cables are everywhere in many fields where the transfer of energy and information is necessary to guarantee the performance of a system. One day or another, a cable network will show signs of weakness or aging involving the appearance of defects. These anomalies can be at the origin of dysfunctions and imply serious consequences for the system or the environment. This is why diagnosis methods for wired networks have been thoroughly studied in the past few years.

Reflectometry based methods have proven to be the best suited, as they provide detection and localization information, while requiring only one connection to the network. But, Time Domain Reflectometry (TDR) or Frequency Domain Reflectometry (FDR) methods are well suited for hard defects (i.e. defect that prevent any signal from going further away). But soft defects, such as localized damage to the insulation or shielding of a wire, are much more difficult to diagnose. This kind of defect account for 30% to 50% of all detected wiring faults, and is the premises of future hard defects.

A new method, called JTFDR (Joint Time Frequency Domain Reflectometry) takes benefit of the advantages of both TDR and FDR while avoiding their limitations by the use of innovative signal processing. It is based on the use of the Wigner Ville transform (WVT) coupled to a normalized Time Frequency Cross-correlation function (TFC) applied to TDR measurements, which greatly enhances the connector’s and soft defects signatures.

Fig. 1 shows that standard TDR measurement cannot efficiently detect and locate a connector or a soft defect in a line, as their peaks have very weak amplitudes [1].

The WVT has previously shown great ability for time–frequency localization of chirp-like signals. For this reason, the Pseudo WVT is combined with a normalized Time Frequency Cross-correlation function (TFC), defined below:

\[
C_{TFC}(t) = \frac{2\pi}{E_s(t) E_r} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} PW_s(t', \omega) PW_r(t' - t, \omega) d\omega dt'
\]

In this formula, Es(t) and Er are normalization factors. The first normalization term provides local amplification of the weak signals isolated by the WVT. As a result, the weak peaks of the connector and the soft defect are enhanced to a level similar to the end of line reflection (an open circuit), making them much easier to detect and localize [2] – figure 2.

So knowing where these jacks are is valuable information for going back to the topology. Once this information is known, the damaged portions of the network can be localized relatively to the connectors; this greatly facilitates and accelerates maintenance.

Besides, in a wiring network, the jacks themselves can be sources of damage. So it is important to be able to monitor their condition. Knowing where they are is the first step to study their health and anticipate their ageing.

References:
SPIKING NEURAL NETWORK FOR EMBEDDED IMAGE SENSORS

RESEARCH TOPICS: EMBEDDED SPIKING NEURAL NETWORK, NANO-DEVICES, MEMRISTORS, CBRAM

D.Roclin, O.Bichler, C.Gamrat, J-O.Klein (IEF PARIS-XI)

ABSTRACT: Our work focuses on the implementation of a Spiking Neural Network (SNN) for an embedded image sensor. To make it embeddable, there are 2 major constraints: power and area. We first analyzed each of the three parts of a SNN, the synapse, the neuron and the learning method, in order to optimize its implementation in a digital architecture and to gain in silicon area and power consumption. To further improve this gain, we chose to implement the synapses of the SNN with emergent nano-devices. Usually realized with flash memory, we wanted to study the gain by designing the first neuromorphic chip with CBRAM devices as synapses.

In a previous work [3], we have shown that Spiking Neural Network (SNN) is an effective mean for processing large natural data sets. The aim of this work is to investigate possibilities for embedding such a network in an image sensor in order to retrieve high level information at the sensor output. We analyzed the main parts of the SNN architecture: synaptic array, neuron and learning mechanism, with respect to silicon area and power consumption. We used an event-based simulator for the task and we started from a previously established simulation, which emulates an analog spiking neural network that can extract complex and overlapping, temporally correlated features in a vision application.

Starting at the synapse level, our first objective was to quantize the lowest bit resolution that maintains the learning ability to and the detection of visual features. We first demonstrated that the SNN was still functional with only 2-bit synaptic weights. Such a bit resolution saves area and energy in a digital architecture, or reduces the need for multi-level storage in a memristor based analog architecture.

Neurons in the SNN circuits are generally of the Leaky Integrate and Fire (LIF) type; hence implementing the leak is a major component of the neuron behavior. Indeed an exponential leak on a digital system is space consuming and adds latency due to the calculation complexity. Therefore, we proposed to implement a linear leak to eliminate these costs. This change eliminated the exponential calculation and replaced it with a multiplicative calculation. This lowered the area and latency of the Arithmetic Logic Unit of the system. We proved that a linear leak did not affect the learning capability [1].

Spike-Timing-Dependent-Plasticity (STDP) or a derivative is generally used as the learning paradigm in SNN. The STDP mechanism requires the implementation of a timestamp for each input neuron that stores the time of the last spike. When an output neuron fires, it compares each synaptic activation timestamp to the LTP window time and decides if the synapse will undergo a Long-Term-Potentiation (LTP) or a Long-term-Depression (LTD). This implementation is space consuming due to the large number of required comparators and counters. Our hypothesis is that the order of arrival of the presynaptic spikes is as important as their precise arrival time. We proposed to replace the time-based LTP window with a First In First Out (FIFO) buffer. When an output neuron fires, it performs a LTP on the synapses present in the FIFO and a LTD on all the others.

This work showed that a FIFO of 200 events is already efficient. This implies that implementing a FIFO memory shared by all output neurons is a valid solution [1].

To further improve the area and power consumption of the SNN we chose to replace the common used Flash memory that implements the synapses with emergent Conductive-Bridge RAM (CBRAM) memristors. This technology has many benefits as for example its non-volatility (no need to refresh the stored data and it is kept even if the power supply is turned off), its low programming power consumption, its scalability and its compact structure when arrange in matrix (1T-1R) or crossbar (1R).

We designed a test chip that includes a 4*4 CBRAM matrix and a 4*4 CBRAM crossbar as well as single CBRAM devices (Fig.2). The chip also includes CMOS circuits that drive the different structures making the chip controllable with digital inputs.

Our neuromorphic hybrid CMOS/CBRAM test chip will further be used for various experimental characterizations on SNN circuits:
- validate the use of memristive devices such as CBRAM for implementing synapses
- retrieve power consumption data
- test various learning schemes for SNN including stochastic techniques
- evaluate the difference in area and power consumption compared to synapses implemented with standard CMOS.

References:
GESTURE RECOGNITION ON SMART CAMERAS

RESEARCH TOPICS: GESTURE RECOGNITION ALGORITHMS, EMBEDDED SYSTEMS, SMART CAMERA

A. DZIRI, S. CHEVOBBE, M. DAROUICH

ABSTRACT: Gesture recognition allows a natural interaction without using complex devices. However, most real-time methods are designed to work on computers with high computing resources and memory. This work aims to analyze the relevant methods and investigates the ability of hand gesture recognition on smart camera. Indeed, two gesture pipelines are designed and implemented on embedded processors. The results obtained show that gesture recognition with an acceptable recognition rate (70-80%) can be executed in real-time (4-30 fps) and uses about 200 kB of memory on embedded processors that can be used in smart cameras.

Several gesture recognition methods are developed to allow a natural human-machine interaction. However, these methods are designed without considering the computing resources and the amount of memory available in the embedded systems used in smart cameras. Furthermore, recognition systems use sophisticated sensors like Kinect, Time-of-Flight camera or stereo cameras. These sensors are more expensive and less compact than a smart camera. Smart camera is a compact vision system integrating an image sensor, low power embedded processors and low capacity memory on the same chip. This work consists in investigating the gesture recognition for ultra-integrated smart cameras. Gesture recognition on smart camera is a challenging problem because of the low computing resources available: low clock frequency for the processor, no floating point unit (FPU) and a small amount of memory (<1 MB).

To meet the challenge, we proposed two gesture recognition pipelines. Each pipeline is composed of two main blocks as shown in the Fig.1. In the first pipeline, the gesture recognition method is based on the use of invariant moments. The second one is based on finger tips detection. Both pipelines are implemented to comply with the constraints of embedded processors that we find in smart cameras. Indeed, floating-point to fixed-point conversion was performed and memory usage reduced. After, we analyzed for each pipeline its performances (recognition rate, frame rate and memory) on embedded processors.

In this study we targeted two embedded processors Antx [2] and ARM Cortex-A9. Antx is a simple low footprint processor that integrates neither hardware multiplier nor FPU. This processor is close to the embedded processors that we find in the ultra-integrated smart cameras. In the other hand, Cortex-A9 is powerful embedded processor that can be used for high performance smart cameras.

The results of performances analysis of both pipelines are shown in the Fig.2. The recognition rate is given by the Fig.2.a. and the execution time on the embedded processors is given by the frame rate in the Fig.2.b. The memory analysis from our implementation shows that less than 200 kB is required for each pipeline.

The results obtained show that the gesture recognition can be executed on embedded processors for few gestures in real-time with an acceptable recognition rate. Future works will consist in implementing the pipeline of finger tips detection on the real smart camera to perform human-machine interaction. Another possible study would be to modify the pipeline to improve the recognition rate.

Figure 2: Performances of finger tips and invariant moment pipelines: IM: invariant moments, FT: finger tips.

References:
A NEW APPROACH OF SMART VISION SENSORS

RESEARCH TOPICS: SMART IMAGERS, ADAPTIVE PROCESSING, FEEDBACK
J. BEZINE, M. THÉVENIN, R. SCHMIT, M. DURANTON, M. PAINDAVOINE (LEAD)

Today’s digital image sensors are used as passive photon integrators and image processing is essentially performed by digital processors separated from the image sensing parts. This approach imposes to the processing part to deal with definitive pictures with possibly unadjusted capture parameters. This work presents a self-adaptable preprocessing architecture concept with fast feedback controls on the sensing level. These feedbacks are controlled by digital processing in order to adapt the exposure and processing parameters to the captured scene parameters. This innovative way of designing smart vision sensors, integrating fast feedback control enables new approaches for machine vision architectures and their applications.

Nowadays, in most image processing systems, the sensor is separated from the image processing parts, pixel values being sent serially. First, photons are integrated for a predefined exposition time; next, a control circuit reads and converts the pixel values from analog to digital sequentially. Finally, pixel values are sent to an image processor for image enhancement or computer vision applications. Thus, image processing systems consider pixel values after the end of full exposure. In that way, corrections such as dynamic range enhancement or image stabilization need to be added in order to suppress the effects of unadjusted image capture parameters. This is particularly true in vision applications such as obstacle detection, or target tracking, the image sensor being used on moving vehicles, suffering from their vibrations and often analyzing difficult scenes (highly contrasted, or suffering of bad weather conditions).

During the last decade, image processing systems tend to link sensing parts with the processing units. Near-pixel processing were introduced in smart sensor, at analog or digital level, in order to refine or adapt captured images before final processing, thus optimizing it. To further improve silicon and energy efficiency, this work proposes to associate even more closely image capture and image processing by adding fast and local feedback controls in the usual image capture process. This adaptation of the usual image capture process is presented in Fig. 1. It is firstly based on the close control of the image capture parameters (exposure time, conversion gain and pixel reset), during photo integration time. This introduces the use of frame sub-exposures to construct a full frame. These sub-exposures may be considered as sampled continuous readout. To deal with the control needed for our approach, a hardware architecture adaptation relying on 3D stacking technologies to process pixel quickly enough to enable capture control – by feedback – during the image construction is proposed. It associates a 2D preprocessing elements matrix to the photo-sensitive layer, separated in pixel blocks. These preprocessing elements are designed to do generic vision pre-computing in order to provide a preprocessed image, or specific image features to the high level processing unit associated. However, the innovative purpose of this layer is to control locally the photo-sensitive layer by processing incoming pixel values on the fly, and sending back adapted capture parameters.

This work was presented in [1]. The first results of such an application on proposed feedback controlled design were obtained. Fig. 2 shows an application of our approach for motion detection in a highly contrasted environment. As image processing algorithms are designed for traditional architecture that processes images after their acquisition, new algorithms must be designed in order to benefit from this smart sensor architecture. Further work will investigate such designs, and enhance our smart sensor adaptation capabilities and flexibility.

References:
Digital technologies are extensively used in the design of modern instrumentation and control (I&C) architectures. The safety assessment of I&C systems relying on digital equipment includes usual measures applied throughout the equipment lifecycle, e.g. quality and configuration management, testing... For systems important to safety, complementary assessments can be required as Independent Validation & Verification measures. For software dependability assessment, such measures include the formal demonstration of classes of properties. In recent works [1,2], we have focused on so-called intrinsic properties, i.e. properties that can be verified without knowledge about the functional purpose of the component under analysis. Interesting properties are typically the freedom from intrinsic faults; e.g. deadlocks, memory leaks,... This led us to present a taxonomy of intrinsic faults that might be postulated in safety or safety related I&C software [1] and some methodologies and case studies regarding specific classes of intrinsic properties in this field [2].

Frama-C [3] is a modular static analysis platform for programs written in the C programming language. Frama-C allows several static analyzers to build upon the results computed by each other, generally enhancing the accuracy of analyses. Recent work [4] was performed as part of an ongoing research project between CEA, AREVA and EDF. This project aims at evaluating the use of the Frama-C platform to address the formal demonstration of three classes of properties in the software of real-life systems important to safety. The property types include the following.

1. Intrinsic properties, as explained above. This direction is intended to investigate whether the state of the art in formal verification of the freedom from intrinsic software faults can be improved: can the process be further automated, the coverage be extended?

2. Structural properties: the challenge here is to identify and formalize the properties that can contribute to a demonstration of independence or non-interference — typically: independence of the system software with respect to its environment in normal operation, non-interference of selftests and clock interruptions on the system software.

3. Functional properties: as opposed to intrinsic properties, this class of properties is directly derived from the system functional specification. The objective is to express functional requirements at the software level under a mathematical form, and then formally prove the functional correctness of the software under analysis. In a first stage, this direction is to be addressed as a feasibility study.

The practical work reported in [4] is performed on a case study representative of real-life software important to safety in nuclear power plants (mainly controller system software, but also application software). The software under verification contains approximately 50k lines of C code, excluding comments. This work is mainly concerned with the first class of properties, namely freedom from intrinsic faults. It discusses the coverage offered by Frama-C with respect to a reference typology of intrinsic faults, presents the results of the case study, and outlines some issues related to the memory model used for the analysis. Partial results regarding the number of potential faults found are presented in Figure 1.

References:
E-ACSL: AN EXECUTABLE SPECIFICATION LANGUAGE

RESEARCH TOPICS: SOFTWARE SPECIFICATION AND VERIFICATION
M.DELAHAYE, N.KOSMATOV, G.PETIOT, J.SIGNOLES
SPONSORSHIP: FUI9

A usual input of a software verification tool includes a program with its (partial) specification. Testing tools need at least a precondition (or test context) specifying admissible input data on which the program should be tested, and usually require an oracle, deciding if the results of the execution on a given test are correct. Detecting potential runtime errors by abstract interpretation also needs a precondition to improve its precision. Tools for deductive verification require a formal specification (function contracts) with pre/postconditions, loop invariants, etc. Runtime assertion checking assumes that program properties to be checked are expressed in the form of assertions (or, more generally, annotations). Although the specification is extremely important for the verification process, its format varies from one tool to another, especially between static and dynamic analysis tools. That makes it difficult to combine them in a completely automatic way.

Recent research [1,2,3] showed that combinations of static and dynamic analyses can be beneficial for software verification. This work addresses C program specification in Frama-C [3], a platform dedicated to analysis of C programs, that comes with ACSL, a formal behavioral specification language shared by various Frama-C analyzers.

The primary objective of this work is to specify E-ACSL [4,5], an expressive sub-language of ACSL that can be translated into C, compiled and used as executable specification. Our second goal is to develop its automatic translator into C. This approach brings several benefits.

E-ACSL is the first formal behavioral specification language for C that builds a bridge between static and dynamic analysis tools and avoids manual rewriting of a formal program specification for testing. Second, choosing a sub-language of ACSL has the advantage of being supported by existing Frama-C analyzers. Third, translating into C rather than into a specific format of a particular tool allows the usage by other analysis tools. Fourth, the possibility to observe the status of an annotation during a concrete execution may be very helpful while writing a correct specification of a given program, e.g. for later program proving. Finally, an executable specification makes it possible to check runtime assertions that cannot be verified statically and to establish a link between monitoring tools and static analysis tools.

In order to support memory-related annotations for pointers and memory locations in E-ACSL (such as being valid, initialized, with a particular offset, etc.), we have developed a monitoring library for recording and retrieving validity and initialization information for the program’s memory locations [6]. The underlying store of memory block metadata relies on a compact prefix tree, also called a Patricia trie. A simple example of a Patricia trie storing four 8-bit base addresses of allocated memory blocks in its leaves is illustrated by Figure 1. We describe the implementation of the store, efficient updates and lookups in it, a static analysis based optimization of monitoring, and our initial experimental results in [6].

The E-ACSL specification language and related Frama-C plugins have been presented in tutorials on runtime assertion checking and its combinations with other analyzers at several international events, such as RV 2013 [5] and TAP 2014.

---

**References:**


Technological systems are complex systems built with many components interacting with each other and combining multiple physical phenomena: thermodynamic, hydraulic, electric, etc. More specifically we deal with systems which have to provide as an output an instruction given as an input. For example a fuel cell of an electric car is a system of that kind: given a power instruction it has to provide this power as an output (Fig1). Faults, which are un-observable damages affecting components of a system, can occur due to many causes: wear, dirtying, breakage, etc. Some are serious and must require to stop the system, or to put it in a safety mode; while others have limited impact and should only be reported for an off-board repair. Thus, it is necessary to achieve on-board the detection of these faults and to discriminate them with the best precision in order to take the appropriate decision. An embedded diagnosis system, complementing the controller, is a suitable solution to do this. However, the problem is then to ensure that this diagnosis system will always be able not only to detect any fault when it occurs (does the fault induce an observable behavior distinct from the normality?), but also to assign a unique listed fault to a divergent observable behavior (do some faults induce the same observable behavior?). This problem is known as diagnosability. In many works [2,3] diagnosability is studied just as if the diagnostic system was a demure having at any time the capacity of knowing the current state of the system, computing the reference state and all faulty states and comparing them altogether. But this is of course not possible in practice because such a diagnostic system would have to run many reference models or embed a huge database. The purpose of such studies is only to assert that if a fault is not diagnosable for such an ideal diagnostic system, it is not for any real diagnostic system accessing less precise information.

In our work [1] we propose a tractable model-based approach of diagnosability which also extends the diagnosability concept: diagnosability is not only the ability of detecting and recognizing faults with no ambiguity, it is also the ability of achieving this identification within a delay compatible with the protection of the system. The main ideas of our approach are the followings. We elaborate models of faults (wear, dirtying, breakage…) and immerse them into the elaborated models of faults (wear, dirtying, breakage…) and immerse them into the model of the system, obtaining that way a set of faulty models, and in such a way that a fault can be activated at will. Then, thanks to a learning process based on simulation, we assign a signature to each fault, which is a set of temporal expressions satisfied by the system behavior when the fault occurs and whose satisfaction can be checked very fast. For example, with respect to Fig1, the fault f1 could be characterized by the following property: after the instruction change (Top ins) then in the time interval [a,b] the output remains steady (x = V[1]x assuming that V[1]x denotes the value of x one time unit after) and meanwhile (on interval [0,a]) there exists at some time a high slope of the output (|V[1]x – x| > S). Gathering all those elements, the signature of this fault is expressed in ARTiMon language by the expression:

$$E_{[0,a]} (V[1]x - x) > S \cup_{[a,b]} (x = V[1]x)$$

The diagnosability of a set of faults can then be expressed in terms of signatures satisfaction: if a fault occurs its signature must be satisfied and not all the others (including the normal-signature) and the normal-signature must be satisfied by any normal behavior. If those conditions are fulfilled the diagnostic can operate like this: it checks all signatures at runtime. If the normal-signature is unsatisfied then the fault that occurred is the one whose signature is true. In our approach the technology used to study diagnosability is identical to the one used to check signatures online, namely this is the ARTiMon technology of the CEA LIST. This ensures the whole consistency of the approach, and the real time performance required by our definition of the diagnosability.

Our work for diagnosability study has been successfully applied on a fuel cell system developed by Sherpa Engineering. Real-Time performances of ARTiMon for embedded diagnostic have validated been on a dSPACE target.

**References:**


TOWARDS VERIFIED CLOUD COMPUTING ENVIRONMENTS

RESEARCH TOPICS: SOFTWARE VERIFICATION
F.Loulergue (LIFO), F.Gava (LACL), N.Kosmatov, M.Lemerre (CEA LIST)
SPONSORSHIP: CARNOT
PARTNERSHIP: LIFO, LACL

With the development of mobile and internet applications, cloud computing becomes pervasive in our lives. More and more of our data are stored in the cloud. It is therefore necessary to have reliable, safe and secure cloud environments.

While the certification of programs in critical systems is an old concern, a recent trend in this area is to formally verify both the programs and the tools used to produce them [1] (and even the tools used to analyze them) using automated and interactive provers, i.e. to have a program and a machine-checked proof that this program meets its specifications. Verified programs could take several forms.

One could use a methodology based on the specification of software and refinement down to an implementation, such as the B method [2]. One could also use methods based on Hoare logic and associated tools.

Another possibility is to take advantage of the fact that the logic of some interactive provers (such as Coq [3]) contains a functional programming language. One writes the program to verify as a functional program, and proves that the function meets its specification (usually written in the logic of the proof assistant that corresponds to usual mathematical logic).

A program runs in an environment: when a program is proved correct in its source form, the proof is done with respect to the semantics of the source language. If the compiler is incorrect, even a proved program could go wrong. A verified compiler itself makes some assumptions about the operating system, and so on. The Trusted Computing Base, or TCB, is the software (and hardware) that is assumed to be correct, without having to prove its correctness, in a computing environment. The trust of an environment increases as the size of the TCB decreases. In the area of distributed systems, the focus is often on security rather than functional correctness. Security is of course important, but security properties can rely on some other properties at different levels of the environment.

From this point of view verifying software with respect to functional specifications is a sub-domain of verifying security properties. But verifying the correctness of a program with respect to a functional specification is interesting by itself and is not necessarily related to security matters. Thus we mostly focus in this work on formal verification of software with respect to functional specifications.

A usual software stack in cloud computing environments is depicted in Figure 1. In [4], we discuss the verification of different layers of the stack. We argue that verification at all these levels is possible and existing work makes us think it is manageable, even if it is not at all in the reach of a single team effort.

In particular, we address the problem of verification of Cloud hypervisors [4,5] and illustrate it on Anaxagoros, a microkernel and Cloud hypervisor developed at CEA LIST, partially verified using Frama-C [6], a platform dedicated to analysis of C programs.

References:
FRAMA-C FROM A SOFTWARE ANALYSIS PERSPECTIVE

RESEARCH TOPICS: SOFTWARE VERIFICATION

P. CUOQ, F. KIRCHNER, N. KOSMATOV, V. PREVOSTO, J. SIGNOLES, B. YAKOBOWSKI

SPONSORSHIP: ANR, FUI, FP7

The past forty years have seen much of the groundwork of formal software analysis being laid. Several angles and theoretical avenues have been explored, from deductive reasoning to abstract interpretation to program transformation to concolic testing. While much remains to be done from an academic standpoint, some of the major advances in these fields are already being successfully implemented [1, 2, 3, 4] – and met with growing industrial interest.

The ensuing push for mainstream dissemination of software analysis techniques has raised several challenges.

Scaling is predictably important from the point of view of adoptability. Handling large problems is a prerequisite for the industrial diffusion of software analysis and verification techniques. Scaling also requires better understanding of how language idioms (e.g. pointers, unions, or dynamic memory allocation) influence the underlying architecture of large software developments. Overall, achieving scalability in the design of software analyzers for a wide range of software patterns remains a difficult question.

Interoperability enables the design of elaborate program analyses. Recent work on the interplay between program analyses and transformations [5], the complementarity of forward and backward analyses [6], or the precision gain afforded when combining static and dynamic approaches [7] has demonstrated the value of interconnected approaches. Yet running multiple source code analyses and synthesizing their results in a coherent fashion requires carefully thought-out mechanisms.

Soundness is a strong differentiator for formal approaches. By using tools that over-approximate all program behaviors, users are assured that none of the errors they are looking for remain undetected. This guarantee stands in stark contrast with the bug-finding capabilities of heuristic analyzers, and is paramount in the evaluation of critical software. But the design and implementation costs of such high-integrity solutions are hard to expend.

The Frama-C software analysis platform provides a collection of scalable, interoperable, and sound software analyzers for the industrial analysis of ISO C99 source code. The platform is based on a common kernel, which hosts analyzers as collaborating plug-ins and uses the ACSL formal specification language as a lingua franca. Its architecture is illustrated by Figure 1.

Frama-C includes plug-ins based on abstract interpretation, deductive verification, and dynamic analysis; and a series of derived plug-ins which build elaborate analyses upon the former. In addition, the extensibility of the overall platform, and its open-source licensing, have fostered the development of an ecosystem of independent third-party plug-ins. The article [8] is intended as a foundational reference to Frama-C’s kernel and main analyses.

References:
BINARY-LEVEL TESTING OF EMBEDDED PROGRAMS

RESEARCH TOPICS: BINARY-LEVEL SOFTWARE VERIFICATION
S. BARDIN, P. HERRMANN (P. BAUFFRETON, N. CORNUET, S. LABBÉ)
SPONSORSHIP: ANR, EDF
PARTNERSHIP: EDF, SAGEM

Formal methods are increasingly being recognized as a powerful approach for building very high-confidence software systems. Existing methods work either on models of the program to be analyzed or on its source code - typically C or Java. More recently, a few teams begin to design analysis tools working on binary code (executable files) [1,5].

Companies working in safety-critical domains could benefit in several ways from automated binary-level testing, for example if (part of) the source code is not available (COTS, legacy code), or if the certification process does require binary-level analysis (DO-178B when optimizing compilers are used). Yet, binary-level analysis is much more challenging than source code analysis.

In the past few years, we developed our own binary-level testing tool OSMOSE [3] geared toward safety-critical programs. Its main output is a set of test data with their guaranteed trace executions. The underlying technology relies on dynamic symbolic execution [3], bitvector constraint solving [2] and formal semantics of machine code [4].

We recently add original features [6] in order to meet certain specific concerns of the safety-critical industry, such as the ability to achieve very high coverage (possibly with a little guidance from user) and to provide elements of trust in the tool’s results. These new functionalities include a generic search API, search directives, test suite replay & completion and modular reasoning.

Thanks to the new features, OSMOSE has been successfully used in the following industrial contexts from aeronautics (SAGEM) and energy (EDF) [6]:
– system-level testing of a small but complex program (Sagem): OSMOSE achieves full coverage, while random testing stayed below 50% coverage;
– understanding and testing a third-party software (EDF, handcrafted example from vendor): OSMOSE allowed to pinpoint problems in documentation (ack. by vendor).

Moreover, experimental results carried out at EDF R&D demonstrate that binary-level coverage stresses the program under test in a more thorough way than the most advanced source-based coverage criteria.

These results are part of a larger effort on binary-level program analysis. Future work comprises wider experiments as well as exploring applications of binary-level analysis to security-critical programs.

Figure 1: binary-level program verification
Figure 2: OSMOSE

References:
GATEL: A V&V PLATFORM FOR SCADE MODELS

RESEARCH TOPICS: TESTING AND VERIFICATION OF SCADE MODELS
B. BLANC, B. MARRE, P. MOUY
SPONSORSHIP: IRSN
PARTNERSHIP: ESTEREL TECHNOLOGIES

Scade is a specification and programming language for the description of synchronous data-flow computations. It is used for reactive control/command systems, mainly for aeronautics control and electrical power production applications. This language allows mixed graphical and textual descriptions into the corresponding Scade SUITE [1].

GATEL performs testing activities from Scade models and verification property on Scade models. It interprets the language constructs as boolean, numerical and temporal constraints using constraint logic programming [2]. An implementation of a k-induction procedure is used for verification activities.

Test sequence generation and proof of properties are directed by specific objectives which are tied to safety properties or to declarative characterizations of some interesting states of the models expressed with a Scade boolean property (including or not a specific observation of the past). These objectives define the behaviour to cover or the negation of the property to prove. Then, the model itself and the objective are translated into a constraint system to use constraint logic programming techniques. The use of a specialized constraint solver allows for a precise fit to user needs. If a sequence is generated, it represents a test sequence that covers the behaviour to observe, or a counterexample of the property. The objective can also be proved unreachable statically or by using a k-induction procedure. If no solution exists due to GATEL parameters (number of cycles, bounds of domains…), the tool indicates it to user.

One main characteristics of GATEL is to allow the user to define his own selection strategy. There are two different ways to describe objectives in GATEL and consequently, to guide or help test or proof strategies. The first one relies on an interactive unfolding of Scade operators of the current constraint system (test cases are defined by a structural decomposition of the initial constraint system). For example, by unfolding the If operator in a constraint S= If Cond then ExpThen else ExpElse, two sub-domains are automatically derived – one for each branch. The second way to describe objectives is the use of predefined functional scenarios attached to variables. A scenario can be seen as a high level splitting method, since the user identifies which parts of the domain to explore, only exhibiting particular instances of behaviour among all possible ones. Each scenario represents one expected case of the decomposition, and is defined as any boolean Scade observation of the past: (*I split Var with [case1,…,casen] *).

These mechanisms are used during verification by case-splitting the objective into easier to prove sub-goals.

The GATEL resolution proceeds as a standard DPLL algorithm but some dedicated treatments improve performance and distinguish GATEL from other similar tools. One characteristic of GATEL is the dedicated treatments done for temporal operators, automata and associated clocks [3]. Clock declarations define a hierarchy and for each clock, the highest number of cycle and the associated status is memorised. The use of properties (order, initial or not, complementarity…) on this clock hierarchy improves performance on equivalent models (automata version vs. boolean version). GATEL also uses affine abstraction for counters. If a counter is identified in the model, it is abstracted with affine intervals to define lower and upper bounds of variation for this counter. At the same time, a reduced interval is computed. This abstraction makes it possible to compute an over-approximation of the necessary number of cycles needed to meet the objective. Another particularity of GATEL is the large choice of arithmetic models. Three models are available for the real type: simple floating point, double floating point or real arithmetic. More details concerning floating point arithmetic are given in [4]. These various arithmetic models can be used to identify particular defaults of the Scade models or for example, to identify discrepancies between ideal and concrete semantics in a given computation.

References:
[1] www.esterel-technologies.com
STATIC ANALYSIS OF NUMERICAL PROGRAMS AND SYSTEMS

RESEARCH TOPICS: PROGRAM VERIFICATION, HDR DEFENSE
S. PUTOT

Among program verification techniques, static analysis consists in establishing properties of programs that hold for all possible executions of the program, given sets of inputs and parameters. For instance, one property of interest is that a program cannot produce run-time errors. This is only a first step towards correction of numerical systems: we also want the program to compute what is expected! Correctness implies verifying functional properties of the algorithm. But it also supposes that the use of finite precision does not impact too much the behaviour of the program. Many other sources of uncertainties exist, whose effect on the system need to be taken into account.

The approximate knowledge of the environment of the program, usually modelled using differential equations, is for example one of these sources of uncertainty.

My Habilitation Thesis (HDR) [1] is a survey on the work conducted by the Fluctuat team on these subjects during the last 12 years, in the LSL then LMEASI laboratories, and their implementation in the Fluctuat static analyzer.

For safety-critical programs, having a proof of good behaviour is crucial. And automating this proof is important in order to analyze large programs. The approach chosen here, abstract interpretation, consists in abstracting the properties of interest in an ordered structure, in which operations and fixpoint computations can be transferred. A key step is to define a tight and sound over-approximation of the property, that can be computed efficiently. For numerical properties of programs, most classical abstract domains rely on sub-polyhedra.

We defined a family of zonotopic abstract domains, in which arithmetic operations rely on affine arithmetic. This domain is of low complexity, and well suited to the analysis of numerical expressions, involving non-linear operations. The explicit parametrisation by affine forms shows a strong analogy with the Taylor methods used in guaranteed numerical computations or hybrid systems analysis. A challenge was the definition of order-theoretic operations, union and intersection, needed whenever the execution of the program can take different paths. Zonotopes are expressive enough to synthesise invariants for many stable recurrent sequences such as linear recursive filters. Fluctuat is able to determine zonotopic invariant sets of the iterates, as shown for an order 2 filter in Figure 1.

We also developed some variations of these domains. For instance, an inner approximation, to compute sets of values of the outputs, that are sure to be reached for some inputs in the specified ranges.

Or probabilistic affine forms, which deliver guaranteed sets of distribution of values of program variables. For instance, on the filter example, we can prove that reachable values for the output are in [-1.09,2.76], but outside [-0.25,1.75] only with very low probability.

We extended this abstract domain to study the propagation of rounding errors and uncertainties in programs. One difficulty is the possibility for a test to be unstable: when, for a given input, the finite precision control flow can differ from the control flow that would be taken by the same execution in real numbers. Not taking this possibility into account may be unsound if the difference of paths leads to a discontinuity in the computation. For instance, in the scheme analyzed in Figure 2, Fluctuat indicates a discontinuity error (in purple) on the stopping criterion of the loop, signalling that it may not converge in the same number of iterations in finite precision and in real numbers.

Figure 1. Invariant set for linear filter iterates
Figure 2. Discontinuity error at stopping criterion

References:
Information security is the preservation of confidentiality, integrity and availability of information. Over the last decade, it has become a critical field of concern for individuals, companies, and states. For individuals, information security has a significant effect on privacy. For companies, should confidential information about a business’ customers or finances or new product line fall into the hands of a competitor or a black hat hacker, a business and its customers could suffer widespread, irreparable financial loss, not to mention damage to the company’s reputation. For states, information security is a key element of national security as shown by the recent announcements about the growth of the ANSSI (Agence Nationale pour la Sécurité des Systèmes d’Information).

Non-interference [1] provides confidentiality and integrity of information by ensuring that an attacker cannot observe (for confidentiality) or modify (for integrity) sensitive data. For instance, a credit card number must not be revealed to anyone knowing the name of the card’s owner. Figure 1 illustrates one variant of non-interference for confidentiality: if attackers control the public insensitive inputs of a program and may observe its public insensitive outputs, they must not be able to deduce any information about its private sensitive inputs. To guarantee this property, no flow of information must go from private inputs to public outputs.

There are nowadays several approaches which aim to guarantee non-interference properties of computer programs [2, 3]. But none is best and each has its own advantages and drawbacks depending on target programs and goals. Also there is still no prominent tool to ensure non-interference of large-scale applications.

Based on this observation, we developed the following ideas. First, since no technique is best, we would like to let the end-users choose not only the best technique, but the best combination of techniques. Next, since getting a new tool able to handle industrial applications requires years of costly development, we would like to rely on the existing Frama-C framework developed in our lab and which is already able to verify safety properties of large-scale critical applications by combining several analysis [4].

We then proposed an innovative approach to tackle this problem. It is shown in Figure 2 and consists in a program transformation which converts the non-interference property on an input program Pin to a set of safety properties on another program Pout. This way, the end-user can choose a combination of existing analysis embedded in Frama-C (e.g. Value, WP, or E-ACSL) to establish the (un)safety of Pout. He can then conclude whether the non-interference property of Pin is valid.

Our work is both theoretical and practical: we proved that our program transformation is the first sound information flow monitor for programs handling pointers and aliasing, while we developed a new Frama-C analysis implementing this transformation. Experiments are ongoing.

This work received two best student paper awards [5, 6] in national and international conferences.

References:
RESULTS FOR COMPOSITIONAL TIMED TESTING

RESEARCH TOPICS: COMPOSITIONAL TESTING, TIMING CONSTRAINTS, SYMBOLIC EXECUTION, MODEL-BASED TESTING, COMPONENT-BASED SYSTEMS, REAL-TIME SYSTEMS

BOUSTHEINA BANNOUR, CHRISTOPHE GASTON.

SPONSORSHIP: ARTEMIS PROJECT MBAT: HTTPS://WWW.MBAT-ARTEMIS.EU/HOME/

Building test harness for systems composed of a large number of interconnected units can be a difficult task. This may be due to the fact that monitoring exchanges between subsystems requires too much technological efforts. This situation can typically occur when systems are distributed. The lack of internal instrumentation can prevent some faults from being discovered. In order to have a partial observation of such internal behaviours, one may try to define test cases whose execution results, even though observed only at the system interface level, are informative about those internal behaviours. However, this requires the system to be sufficiently controllable, which is often not the case, especially of distributed or asynchronous systems. Moreover, once a fault is observed at the testing interface level, further observability capabilities are often needed to identify which subsystem(s) caused the fault. Another technical limitation is related to side-effecting instrumentation which can be complex enough to modify the system runtime performance. Such influence questions the quality of the observation and hence the accuracy of the testing process.

In order to tackle those issues, compositionality results may be useful. In a compositional framework, a system is seen as an assembly of interacting units whose correctness implies the correctness of the whole system by construction. Thanks to such a result, testing the whole system comes mainly to testing its units. At the system level, one only has to check that the system is assembled in a way that reflects the structuring operators considered in the compositionality result. When testing is used for that purpose, this is the simplest form of what is called integration testing. In this work, we study such a compositionality result in a model-based testing framework dedicated to timed systems [Schmaltz2008]. Models are particular automata, named Timed Input Output Labeled Transition systems (TIOLTS) whose transitions are labeled by inputs, outputs or durations. In this framework, the correctness of a system is modelled by a so-called conformance relation named tioco. TIOLTS can be composed to constitute models of systems. For such system models, we show how to extract the set of possible executions of its units in the context of the system. For each unit, it is characterized, thanks to projection mechanisms, as a TIOLTS whose set of traces is a subset of the one characterized by the stand-alone TIOLTS that models the unit. Our projection techniques are based on those that we introduced in [Faivre07] and latter adapted to deal with time [Gaston2012]. We show that under a certain hypothesis, if each unit conforms to its TIOLTS obtained by projection then the composition of unit models conforms to the system one. As usual in testing, such a result is more exploitable in its converse expression: if a composition of units does not conform to a system model, then there is at least one unit that does not conform to its TIOLTS obtained by projection. Therefore assuming that the units are correctly composed, any potential fault of the system can be identified by testing its units at the unitary level. The process is illustrated in the following figure.

The TIOLTS framework is adequate for theoretical reasoning, however as soon as one wants to really model systems, it is not expressive enough. Indeed, it is classical to reason about durations in terms of first-order constraints, which is impossible with TIOLTS. In previous papers [Gaston2012], we presented a symbolic framework called Timed Input Output Symbolic Transition Systems (TIOSTS) to deal with such constraints, and more generally with data in a symbolic manner. We defined testing algorithms based on TIOSTS. Our compositionality result can be used with those algorithms to test systems. This requires to transpose the notion of projection to the symbolic context and to define an algorithm for assessing the satisfaction of the compositionality result hypothesis.

References:

ABSTRACT: Providing security in the cloud has been approached in different angles and with different upside and downsides in the trade-off between security and efficiency. One of the most serious candidates is Homomorphic Encryption, as it allows to make some computations on the encrypted data, while its security has been proven to be very strong.

Our work consists in identifying the computational hot spots resulting of the use of Homomorphic Encryption and find ways to mitigate them. Eventually we demonstrated that Homomorphic Encryption was already a realistic approach (and probably the best) for some algorithms.

Homomorphic encryption has been introduced in 2009 and it has been one of the most promising cryptographic research line since. Indeed, specific computations on the ciphertexts can be passed on to the underlying plaintexts thanks to the special design of a homomorphic encryption scheme.

For example, we can see in Fig.1 the main advantage of homomorphic encryption compared to classical encryption. When the ciphertexts cannot be used for anything else than decryption in classical encryption, they can be added or multiplied when using homomorphic encryption, and the additions or multiplications impact on the plaintexts.

Plaintexts: \( p_1, p_2, p_3 \)
Ciphertexts: \( c_1, c_2, c_3 \)

This property is particularly interesting in the cloud computing paradigm. In the client/server model, a client encrypting its data using homomorphic encryption allows to perform computations on said data, without gaining any knowledge on it. Homomorphic encryption thus appears to be the perfect solution to guarantee outsourced data security, while keeping the possibility to process it.

Of course, using homomorphic encryption efficiently in the cloud computing models requires a lot of work on both the cryptographic primitives themselves and on the process of homomorphic-encrypted data. Our research focuses on the latest issue, and has been published in two articles in 2013 [1,2].

Our first concern was to gauge the possible computations that could be done on homomorphic-encrypted data. Starting with a homomorphic encryption scheme that allowed XOR and AND operations on encrypted bits, we wrote a small set of simple algorithms, using only XOR and AND gates, that can now be run in the encrypted domain, such as a bubble sort, a threshold, a FFT, etc. [1].

We built our platform in order to be relatively independent of the homomorphic encryption scheme used, since the state-of-the-art is not stable yet. Indeed, we can use a new external implementation of a homomorphic encryption scheme without changing the non-cryptographic part of our platform.

Once we had seen the extent of the expressivity we could achieve using homomorphic encryption, we focused on reducing the overhead of homomorphic computation on the non-cryptographic end.

All the computationally realistic homomorphic encryption schemes are based on lattice-based cryptography, and have the same computational hot spots.

When computing on homomorphic-encrypted data, the XOR gate is nearly free in comparison to the AND gate, which has a big overhead. Another limiting factor is the multiplicative depth of the algorithm we want to run in the encrypted domain (by multiplicative depth we mean, seeing an algorithm as a boolean circuit, the maximum number of AND gates in a path, for all the paths).

We showed in our most recent article [2] how the number of AND gates and the multiplicative depth are crucial to the efficiency of computation on homomorphic-encrypted data. Minimizing theses values are therefore a very big perspective for our future research.

We already started to rebuild Boolean circuits in order to have less AND gates, and smaller multiplicative depth on some test algorithms. In this manner, we have been able to demonstrate a (fictitious) medical diagnosis algorithm that compares (encrypted) medical test results from a patient (blood pressure, cholesterol, etc.) to standard values and give a risk/no-risk answer. Using homomorphic encryption, such a test was completed in less than 2 seconds on an 8-cores processor with a security equivalent to 128-bits.

We have strong hope to be able to perform in the close future more sophisticated algorithms in less time, while keeping the security to the same level.

References:
TIME- AND ANGLE-TRIGGERED REAL-TIME KERNEL AND ITS USE FOR POWERTRAIN APPLICATIONS

RESEARCH TOPICS: TIME-TRIGGERED PARADIGM, REAL-TIME OPERATING SYSTEM
D. CHABROL (KRONO-SAFE), D. ROUX (KRONO-SAFE), V. DAVID, M. JAN, M. HMID, P. OUDIN (DELPHI) ET AL.

ABSTRACT: The time-triggered (TT) approach provides a predictable and reproducible execution of real-time systems but cannot cope with tight temporal constraints (around 100µsec) and does not allow to directly specifying the temporal behavior of the system based on angles. We have extended the PharOS technology to combine several time domains (time and angle triggered) allows designing and executing powertrain controllers in a deterministic way on multi-core architectures. We have developed a prototype of a subset of a powertrain controller from Delphi based on PharOS.

Embedded safety-critical real-time systems are often implemented as hard real-time periodic tasks within a Time-Triggered (TT) real-time operating system (RTOS). The schedulability analysis of such systems must be performed assuming the worst-case demand. However, instrumentation and control multitasking systems can have tight temporal requirements. Besides, the physical law of the system to control and command may dynamically change within a specified range. The use of the TT paradigm to design such systems dramatically oversizes the required processing capability, if not making systems unschedulable.

In the automotive field, the management of an engine is ensured by a powertrain controller. It is an electronic device that assists the control of vehicle’s engine by measuring multiple events and enabling real time adjustments of fuel, air, spark and gear shift to help the powertrain system to efficiently, reliably and economically operate. Its main function is to compute when the injection sequences of fuel in the motor should occur (as shown by Figure 1). The frequency of injection depends on the speed of the engine. In order to reduce of clatter of the motor as well the gas emissions and the fuel consumption, the injection sequence should occur at precise instants. However, the execution time on an injection sequence is independent of the engine speed. As the speed of engine may vary, the starting point of an injection sequence should then be dynamically adjusted in order to be the most accurate. This can lead to temporal of a few hundred microseconds.

We have proposed a new paradigm, called eXternal-Triggered (xT), which generalizes the TT approach to external events and therefore unifies this model to the Event-Triggered paradigm. The triggering of tasks can be linked to the occurrence of events from the controlled system, in addition to the classical physical time scale of the TT paradigm. Input events, such as temporally mastered interrupts, are used by the xT paradigm to produces ticks available for specifying the triggering of tasks. However, each event may not be linked to a tick; several events may be required to produce a tick. This feature is implemented by defining a filtering function for each source of input events. It must be provided by the designer of a domain as it knows the meaning of events from the controlled system. It can also be used to freeze the current time of a domain by filtering any event that will occur. This can be used to implement a synchronization strategy between the value of the current time of a domain and the physical process being control and command or simply in case of hardware failure of the sensors that generate events.

We have extended our PharOS Real-Time Operating System (RTOS) technology to include the support of the xT paradigm and we applied it to develop a subset of a powertrain controller from Delphi over a PowerPC MPC551x evaluation board. Figure 2 shows the temporal behavior of the designed application.

References:
A HEURISTIC ALGORITHM FOR STOCHASTIC PARTITIONING OF PROCESS NETWORKS

RESEARCH TOPICS: GRAPH PARTITIONING, CHANCE-CONSTRAINED OPTIMIZATION, COMPIlATION

O. STAN, R. SIRDEY, J. CARLIER, D. NACE (UTC)

ABSTRACT: In this work, we study the problem of partitioning networks of processes under chance constraints. This problem arises in the field of compilation for multi-core processors. The theoretical equivalent for the case we consider is the Node Capacitated Graph Partitioning with uncertainty affecting the weights of the vertices.

For solving this problem we propose an approximate algorithm which takes benefit from the available experimental data through a sample-based approach combined with a randomized greedy heuristic, originally developed for the deterministic version. Our experimental results illustrate the algorithm ability to efficiently obtain solutions of good quality within an acceptable execution time which are also robust to data variations.

The development of 100+ cores microprocessor architectures has triggered a renewed interest for the so-called dataflow programming models in which one expresses computation-intensive applications as networks of concurrent tasks interacting through (and only through) unidirectional FIFO channels. In [1], we present a heuristic algorithm dedicated to the resource-constrained graph partitioning problem which crops up when mapping networks of dataflow processes on a parallel architecture assuming the resource consumptions of the processes are uncertain.

Known, in the deterministic case of single-dimensional weights, as the NP-hard problem of Node Capacitated Graph Partitioning, the assignment of the weighted vertices of a dataflow graph to a fixed set of partitions, has, to the best of our knowledge, received little attention from the stochastic programming community.

In order to respect as close as possible the real context of our application, a qualitative analysis of the sources of uncertainty, mainly the execution times, was performed. This preliminary analysis showed the inherent difficulty of obtaining an analytical description of the distributions of the execution times.

Even if it is reasonable to assume that the probability distributions of execution times have a bounded support (no infinite loops), we have to cope with the fact that these distributions are intrinsically multimodal (due to the presence of data dependent control). Also, in the case of process networks, we cannot overlook the problem of dependencies between these random variables.

The approach we propose is justified by the theory of statistical hypothesis testing and takes into account the important role of experimental data. Additionally, for solving a chance constrained problem, no assumptions are being made about the joint distribution of the random variables, in particular with respect to the independence of these variables.

Our algorithm design methodology consists in leveraging an existing heuristic for the deterministic case without significant restructuring (i.e. at small cost in terms of software engineering) and with acceptable performance hit. Furthermore, this non-parametric method we introduce for solving our chance-constrained partitioning of process networks is applicable, in combination with other approximation algorithms (e.g. meta-heuristics), to other optimization problems.

Figure 1 shows the statistically significant approximation model to an initial chance-constrained program, obtained with our robust binomial approach. $X_i$ are binaries variables and their sum follows a Binomial distribution, $k$ is a parameter determined in function of $NS$, $\varepsilon$ and $\alpha$, and $L$ is a constant of large size, depending on the problem structure but generally easy to find.

For our problem, the objective $f$ is to minimize the communication inter partitions and the probabilistic constraints are on the capacities of the clusters. The heuristic we adapted for treating the stochastic graph partitioning is an already available greedy randomized affinity-based heuristic, easy to modify and quite efficient for the placement of the processes in the deterministic case.

The task weights are random variables (memory footprint or computing core occupancy) for which we dispose of a relevant sample of $NS$ independent and identically distributed realizations.

By using the statistical hypothesis testing within a heuristic approach, we overcome the computational effort of taking into account the uncertainties of the weights of the vertices. Concerning the complexity, we remark a linear increase with a factor of $NS$ in comparison to the deterministic version.

This approach can solve, with an acceptable solution quality, confidence level and computation time, problems representative in size of our application context. The overall solutions have a quality comparable to those of the heuristic for the deterministic case and moreover they are statistically guaranteed at a confidence level $1-\alpha$.

References:

ADAPTING JUST-IN-TIME COMPILATION TO EMBEDDED SYSTEMS

RESEARCH TOPICS: JIT COMPILATION, EMBEDDED SYSTEMS
A. CARBON, Y. LHUILLIER, H-P. CHARLES

Just-In-Time (JIT) compilation is today widely employed in many application domains and massively transferred to embedded systems. However, JIT compilation complexity lead to important performance loss for embedded processors due to their lack of mechanisms to manage JIT compilation algorithm irregularities in terms of control and data. Managing these irregularities, associative arrays and dynamic memory allocation still represent 25% of the LLVM bytecode compiler execution time, despite many existing software optimizations. To reduce their impact on execution time, our ongoing work consists in the proposition of hardware dedicated resources to accelerate them, based on standard libraries and replacing these software optimizations.

Just-In-Time (JIT) compilation has become a major topic for academic and industrial researchers in the last 15 years. JIT compilation technologies consist on executing all, or parts of, compilation stages dynamically during the application execution.

The main reasons of this growing interest are the following:
- Increasing dynamism of applications and their workloads
- Increasing interactions between applications
- Increasing portability and security requirements
- Increasing performance requirements

Based on a state-of-the-art analysis, we identify four main technologies using JIT compilation:
- Virtual machines (eg. Java Virtual Machines)
- Dynamic binary translation (eg. Apple Rosetta)
- Multistage dynamic compilation, consisting in deporting compilation phases to runtime (eg. LLVM framework)
- JIT compilation for dynamically-type languages (eg. JavaScript, Python)

In all these technologies, the efficiency of Just-In-Time compilation depends on the ability to compensate its overhead with execution speedups obtained on the generated code. Compilation algorithms are complex and particularly difficult to handle for embedded processors, with important performance loss introduced, as presented in Figure 1.

We profiled different JIT compilation algorithms [1], extracted from the highlighted technologies, run them on a x86 processor, and compared their behavior to regular algorithms extracted from miBench benchmarks. Figure 2 shows a common control irregularity compared to regular algorithms with a slight increase of misprediction rates for indirect branches and instruction cache. This highlights a common complexity on control and a high instruction/data ratio contrary to regular algorithms, in which the amount of managed data is far bigger than the amount of instructions.

Table 1 shows a common data irregularity with a slight increase of instruction depths for JIT compilation algorithms (LLVM in our case), highlighting the fact that JIT compilation algorithms are significantly point-intensive.

These irregularities, already visible on x86, are important issues for embedded processors due to a lack of mechanisms to handle them (especially concerning predictions).

To highlight the parts of the code responsible for these irregularities, we profile the LLVM bytecode compiler. Results obtained show that associative array and dynamic memory allocation represent on average 25% of its execution time [2], despite many existing software optimizations: LLVM developers provide more than 8 specific data-types re-implementing standard data types of the C++ STL library. Our ongoing work deals with the development of hardware dedicated resources, based on standard libraries. We are looking for accelerating associative array management and dynamic memory allocation, replacing existing software optimizations to reduce their impact on execution time.

Table 1: Algorithm indirection depths (x86 instrumented simulator).

<table>
<thead>
<tr>
<th>Standard</th>
<th>Compression jpeg</th>
<th>Djukstra</th>
<th>LLVM compiler</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;2</td>
<td>192</td>
<td>75541</td>
<td>7733</td>
</tr>
</tbody>
</table>

Figure 1: Execution time slowdown comparison between conventional and JIT compilation algorithms.

Figure 2: Misprediction rates for indirect branches and instruction cache, relative to the total number of instructions (x86).

References:
DYNAMIC CODE GENERATION:
LARGE SPECTRUM, MANY APPLICATIONS

RESEARCH TOPICS: DYNAMIC CODE GENERATION, COMPILATION, OPTIMIZATION
H.P. CHARLES, D. COURROUSSE, Y. LHUILLIER, V. LOMULLER, A. CARBON

To generate an executable program, programmers generally use a compiler which transforms source code to binary code (compile time), and during the execution (run-time) the binary program does not change. This is the classical way to produce and execute binary programs, called static compilation scheme.

There is an increasing number of situations where this classical scheme insufficient, and sometimes counterproductive.

Due to the increasing complexity of memory hierarchy and parallelism levels, program performance is increasingly tied to the characteristics of the data sets. Data size, data alignments and data values have a deep impact on program behavior. Iterative compilation is a technology allowing a binary code to be adapted to a running data set, but the adaptation is only valid for one data set, not the unknown that a user can provide. The adaptation of a running code to a given data set parameter is one of our motivations.

Portability issues on embedded systems, such as cell phones and set-top boxes, is another motivation. Many compilation infrastructures are used such as JIT compilation for Android Java based applications, for javascript in browsers for smart phones and set-top boxes, for graphic rendering on Android GPUs, etc. In these situations, binary generation has to be fast, occupy a small memory footprint, generate efficient code, but use power sparingly.

We have developed a tool and experiments that try to tackle these problems. Usage examples are listed in the following items:

ISA dynamic adaptation: we have developed a small code generator embedded in the Scilab mathematical solver that is able to determine at run-time if a code should run either on a CPU or on a GPU. On the GPU side, our code generator is able to dynamically adapt the code according to the matrix size and based on initial benchmarking.

The results of the experimentation are shown in the figures 1 & 2. Figure 1 shows the performance of MAGMA, the reference mathematical library for matrix multiplication on GPUs. For clarity we have only plotted the results superior to 145 GFlops for matrix sizes between 64x64 and 2000x2000. In figure 2 we have plotted the results from the same experiment using our library dynamic adaptor. [1]

JIT hardware acceleration: [3] JIT compilation requires a lot of memory access, which uses hash tables and tree balancing. This impacts the performance because code generation is done at run-time. We have shown in our article that we could mix data and programs in order to accelerate searches in these trees.

VLIW dynamic bundling: [2] Many VLIW processors use bundled instructions (grouped processor instructions) to implement instruction-level parallelization. The compilation process tries to maximize bundle usage. We have shown that we could use dynamic "bundesization" by using running parameters and improving code performances.

Tool for dynamic code generation: [4] dynamic code generation is a difficult task owing to the lack of general tools. We have developed an infrastructure which helps to build dynamic code generators.

References:

Figure 1: Matrix product GFlops (> to 145) obtained on NVIDIA GPUs by the MAGMA BLAS library.

Figure 2: Matrix product GFlops (> to 145) obtained on NVIDIA GPUs by our dynamic library.
FAST AND AUTOMATIC CITY-SCALE ENVIRONMENT MODELING FOR AN ACCURATE 6DOF VEHICLE LOCALIZATION

RESEARCH TOPICS: COMPUTER VISION, AUGMENTED REALITY
VINCENT GAY-BELLILE, DORRA LARNAOUT, STEVE BOURGEIOS, BENJAMIN LABBE, MICHEL DHOME
PARTNERSHIP: INSTITUT PASCAL

To reduce the cognitive effort of the drivers, navigation systems are moving towards Augmented Reality applications. Most of existing solutions use the geolocation from the GPS. Nevertheless, GPS provides in-plane data (2D locations and therefore the yaw angle) whose accuracy is insufficient to a high quality service in urban context. To overcome GPS uncertainties, 6DoF vision-based localization solutions exploiting geo-referenced 3D landmarks database can be employed. To ensure an accurate localization at city scale, large scale databases have to be created while maintaining a high accuracy of their landmarks position. Moreover, to provide a good service quality, database update should be frequent. Ideally, it must rely on a collaborative approach where the end-user can extend the existing database. Therefore, solutions using RTK GPS, catadioptric cameras or LiDAR are inadequate since these sensors are unavailable on consumer cars.

We propose a complete framework that automatically builds an accurate city scale database of landmarks using only a standard camera, a GPS and Geographic Information System (GIS). We exploit two components of the GIS: the Digital Elevation Model (DEM) and the 3D buildings model to create an accurate geo-referenced database. We introduce a coarse-to-fine approach to fuse all these information which works in three steps:

1. An initial geo-referenced reconstruction is obtained by fusing visual SLAM with GPS data and Digital Elevation Model (DEM) [1].
2. After this initialization step, the most important errors reside in the in-plane camera parameters (position and yaw angle) due to the GPS inaccuracy. These errors are corrected in priority by using the constraints provided by the buildings models. They are used to refine the in-plane DoF of the camera poses and therefore improve the 3D point cloud [2].
3. A last 6DoF refinement is performed on the in-plane and out-plane parameter (altitude, pitch and roll angle) of the camera poses using constraints provided by both the buildings model and the DEM [2].

Results. To evaluate the modeling process, two large scale databases are built from two different sequences recorded in the districts of Versailles in France (the first sequence: about 2400 meters, the second sequence: about 1800). These sequences are recorded in normal driving conditions (50 km/h). For that, the vehicle has been equipped with both a standard GPS (1Hz) and an RGB camera providing 30 fps with a 90° field of view. GIS models with an uncertainty of 2 m, are provided by the French National Geographic Institute. The resulting database are illustrated in Figure 1, only few minutes are required to obtain accurate database using a non-optimized code executing on an Intel(R) Xeon(R) CPU quad cores 2.4 GHz.

Once the databases are created, they can be used to ensure an online and accurate localization and therefore an Augmented Reality application such as Navigation Aided through Augmented Reality.

Figure 1: The databases built with the proposed solutions. Note that, even if they do not cover the whole city of Versailles, it is possible to merge several databases to create a city scale database.

References:
AN INTERACTIVE AUGMENTED REALITY SYSTEM: A PROTOTYPE FOR INDUSTRIAL MAINTENANCE TRAINING APPLICATIONS

RESEARCH TOPICS: COMPUTER VISION, AUGMENTED REALITY
BASSEM BESBES, SYLVIE NAUDET-COLLETTE, MOHAMED TAMAAZOUSTI, STEVE BOURGEOIS, VINCENT GAY-BELLILE, MICHEL DHOME
PARTNERSHIP: INSTITUT PASCAL

We created a prototype to investigate the application of AR to industrial education and training. It features a tracked monocular Optical See-Through Head Mounted Display (OST HMD) HMD that solves the 3D registration issues and allows to accurately superimpose virtual maintenance procedures on an industrial object. The training system guides a learner step by step with virtual sequences of an assembly/disassembly procedure for a specific object of interest. The user can interact with the system by simply pointing to select a specific object component with an ordinary laser pointer or his finger.

PROTOTYPE DESCRIPTION

Our prototype involves two main vision-based modules: camera localization and user-interaction handling. The first module includes markerless trackers for camera localization, which can deal with partial occlusions and specular reflections on metallic object surfaces. The tracking solution used is based on our previous work [1] that estimates the camera trajectory by combining SLAM and model-based localization techniques in a single framework. It results in an accurate and robust localization which is crucial in our context to deal with the fast displacement of the head. The second module [2] is a fast image processing method for red laser dot or finger tracking. By combining these processing elements, the proposed system is able to deduce the element pointed to by a ray tracing of the workpiece.

EXPERIMENTATION

The prototype incorporates a monocular OST HMD and a DELL PC (Intel quad-core Xeon X5482 3.2GHz, 3.25 GB RAM). The OST HMD can easily display additional information such as 3D CAD objects or simple virtual animated sequences on the natural field of vision of the user. This system allows to augment the users vision field with virtual information accurately co-registered with the real world. To keep greatly benefit of the potential of this hand free device, the system combines the tracker module [1] with a simple user-interaction vision-based module [2] to provide overlaid information in response to user requests. As depicted in Figure 2, the system is capable of interactively augmenting the selected component with an assembly/disassembly animated virtual sequence.

Figure 1. Overview of the proposed system for Industrial Maintenance training through Augmented Reality

Figure 2. Interaction illustration: 3D augmentation concerning an assembling procedure is activated in response to the user finger designation.

References:
Nowadays, many robotic applications require an accurate model to perform tasks where dynamics is significant. The friction model discussed in this paper aims at improving the existing rigid robot model. The losses in joint transmission originate in friction between moving parts in contact or between moving parts and the ambient fluid. Commonly, robotic identification models represent joint transmission friction force as a viscous friction force, depending on the velocity, added to a constant dry friction force. However, the tribology science field teaches that friction in general depends on load (reaction force normal to the contact surface). It is important to consider this dependence when variable loads are applied on the joint transmission (external payloads, inertial load and gravity forces). Since these mechanisms are lubricated, it is appropriate to refer to the Stribeck curve (rather than Coulomb). This curve describes the friction coefficient as a function of both inertial and external forces. Since these mechanisms are lubricated, it is appropriate to refer to the Stribeck curve (rather than Coulomb). This curve describes the friction coefficient as a function of both inertial and external forces.

In this paper, we use the Force Transfer Diagram (fig.1), a representation that provides an implicit representation of friction in a mechanism and its dependence with load and velocity variation, the thresholds, the dissipative quadrant as well as the power asymmetry inherently present in some mechanisms (incline family) [1].

An efficient method for the dynamics model estimation is based on writing the inverse dynamic model (IDM) as linear with respect to the dynamics parameters and using properly designed experiments. Least squares algorithm (LS) is used to obtain the estimates of these dynamic parameters. This method has been applied and improved on a large number of robots and prototypes to identify inertial, gravity, and friction parameters. The usual friction model used in the IDM distinguishes the dry friction from the viscous friction. The viscous friction is directly proportional to the velocity (Fr) and dry friction is modeled as a constant (Fc sign()). This dry friction constant is identified without external load or with a given payloads [2]. However, the Stribeck curve plots the variation of the friction coefficient according to the Hershey parameter combining the effects of load, velocity and viscosity.

This paper proposes a new expression of the load-velocity friction model, in order to identify a serial n degrees of freedom (DOF) robot. The friction force of this new inverse dynamic identification model is a linear function of both inertial and external forces. An experimental validation on an industrial manipulator used as a force feedback telerobot in nuclear plant is exposed in this paper.

Table 1: Comparison relative norm of error (validation set)

<table>
<thead>
<tr>
<th></th>
<th>usual</th>
<th>new</th>
<th>% improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>all</td>
<td>0.049306</td>
<td>0.036022</td>
<td>27%</td>
</tr>
<tr>
<td>axe</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>q&lt;10%</td>
<td>0.0810</td>
<td>0.1112</td>
<td>23%</td>
</tr>
<tr>
<td>q&lt;50%</td>
<td>0.0620</td>
<td>0.0606</td>
<td>22%</td>
</tr>
<tr>
<td>q&gt;50%</td>
<td>0.0575</td>
<td>0.0590</td>
<td>27%</td>
</tr>
<tr>
<td>q&gt;50%</td>
<td>0.0520</td>
<td>0.0570</td>
<td>2.7%</td>
</tr>
</tbody>
</table>

References:
ASSISTANCE TOOLS FOR GENERIC DEFINITION OF ITER MAINTENANCE TASKS AND SCENARIOS IN ADVANCED SUPERVISORY CONTROL SYSTEMS

RESEARCH TOPICS: ROBOTIC SUPERVISORY CONTROL
S. ZIEBA, F.-X. RUSSOTTO, M. DA SILVA SIMOES, Y. MEASSON
SPONSORSHIP: F4E. PARTNERSHIP: ITER, EFDA GOT-RH

ITER maintenance requires a combination of scheduled and unscheduled replacement tasks on the machine during periodic shutdown times, and to carry out component maintenance in the hot cell both during machine shutdown and operational periods [1]. Due to restricted visibility conditions, such tasks require the use of advanced Supervisory Control Systems able to provide operators with efficient virtual representation of the environment and context-based assistance functionalities, optimizing operations effectiveness, reliability and safety. Complexity and diversity of maintenance tasks require the use of a large set of remote handling (RH) equipment gathered in a control room and organized in work cells.

This paper aims at providing a unified formalism and a generic set of tools to define the ITER maintenance tasks and scenarios regarding in-vessel and hot cell operations. A three-layered approach is proposed to model these tasks and scenarios. A set of relevant physical actions is defined for every elements of the scene. From these physical actions, behaviors are defined to represent relevant high-level functionalities. Finally, functioning modes are defined to represent human-machine interactions.

This model aims at describing all the relevant actions required for the maintenance scenarios. Firstly, available operations are defined via physical interfaces attached to the scene elements. Such interfaces define the nature of the action that can be performed, the location and extent of the interaction relatively to the object and, optionally, an assistance function for the action achievement. Description of the scene with these interfaces provides a function-centered model of the environment. Actions available on an object are thus identified according the provided interfaces. Context-based assistances can then be dynamically proposed to the human operator to highlight these opportunities of action. Interfaces shall provide physical data to enable actions between elements of the scene. Such data allow defining precisely the location and the modalities of every relevant physical interaction with objects of the scene. Such a unified representation for the elements of the scene eases the planning activity for the human operator by allowing simple modifications of the original action plan and offers learning perspectives for the remote handling system, the operator being able to teach the automated system some new tasks.

Context-based assistances are useful for situations where cooperation is required to perform a global task. Cooperation can be initiated between operators in a same work cell or between different work cells. HMI must gather information about the progress of each work cell in the procedure and indicate the points where a transition will take place. From the point of view of the supervisor, HMI of the workstation must clearly identify the tasks of each work cell and the planned interactions in the procedure. Such assistances concern more especially the supervisor workstation with indications of the type of cooperative activities.

Blanket remote handling system (Figure 2) can be considered as a relevant illustrative example regarding the needs for cooperation between work cells. For this maintenance task, a rail, composed of two semi-circular sections, is deployed around the equatorial region. On this rail, a vehicle manipulator travels so that a blanket handling gripper can access the in-vessel areas.

Perspectives concerning advanced assistances shall be related to the development of an intention recognition module. The objective of intention recognition is to relieve the operator mental workload by computing sequence of actions to be performed based on the procedure followed by the operator, the movement of the operator, the current tool, the current grasped object… According to the level of automation [2], the intention recognition module may suggest a list more or less restrained of alternatives or initiate the most probable intention without waiting for the operator’s consent.

References:
OPTIMAL DESIGN OF COMPLIANT JOINT, GRIPPERS AND ACTUATOR FOR MINIATURE AND PORTABLE ROBOTIC DEVICES

RESEARCH TOPICS: MINIATURE AND MICRO-ROBOTICS

C. ROTINAT-LIBERSA, B. SOLANO, A. GRIRA, B. LEGRAND\(^1\), L. BUCHAILLOT\(^1\), R. EL KHOURY MOUSSA, M. GROSSARD, M. BOUKALLEL, A. HUBERT\(^2\), N. CHAILLET\(^2\), C. VIDAL\(^3\)

SPONSORSHIP: ANR TECSAN
PARTNERSHIP: \(^1\)IEMN, \(^2\)FEMTO ST, \(^3\)ENDOCONTROL

Miniature scale robots development is needed to address force feedback manipulation applications, such as grippers for miniature objects manipulation or characterization, minimal invasive surgical (MIS) instruments or hand prosthesis. Indeed, all these devices have to be small, powerful and lightweight, even portable, and so require new structural, actuation (and sensing) technologies with a high level of energy density and potential integration.

Thus, concerning the development of grippers for mini- or micromanipulation, the CEA LIST Flexin software [1] for the optimal design of compliant structures has evolved. On one hand, it allows the synthesis of a monolithic PZT gripper integrating structural, actuation and force sensing functions [2]. On the other hand, the method has been adapted for the optimal synthesis of MEMS (Micro Electro Mechanical Systems) [3]. It has been applied to design a silicon-based micro-gripper integrating inter-digital electrostatic actuators and sensors, with outstanding performances regarding the state-of-the-art: two time larger stroke (5\(\mu\)m) and a theoretical force resolution three order of magnitude better (55pN). In addition, Flexin software has been used to design a monolithic compliant joint structures have been developed [4].

Furthermore, the development of miniature compliant articulations with locally embedded actuation articulations with locally embedded actuation is a scientific and technological important issue which should allow a better integration, size reduction and portability, without decreasing the dexterity. Firstly, new compliant bending joint and Universal joint structures have been developed [4].

Secondly, locally embedded actuation principle has to provide large forces within small volume. That's the reason why new fluidic miniature actuators with high mechanical power density have been developed, based on the McKibben muscle principle [5, 6]. One allows producing a traction force up to 10 N and a 20% contraction rate, with a 1.5 mm cross section diameter actuator. The mechanical behavior of this actuator has been analytically modeled [5] and has the potential to be used soon to control the actuator without using any exteroceptive sensors, thus making the miniature robotic devices integration easier. In addition, thanks to its natural compliance, this locally embeddable actuator remains a good candidate for actuation of robotic fingered hand or other miniature devices to be used in close interaction with human.

The new compliant joints designed and the McKibben type miniature muscle developed have been integrated into a portable dexterous MIS instrument [6] for complex suturing tasks. In this device, the fluidic muscle has been locally embedded in order to actuate the intracorporeal compliant needle holder, avoiding parasitic mechanical couplings with the other motorized degrees of freedom of the device.

References:

A FRAMEWORK FOR THE CLASSIFICATION OF DEXTEROUS HAPTIC INTERFACES BASED ON THE IDENTIFICATION OF THE MOST FREQUENTLY USED CONTACT AREAS

RESEARCH TOPICS: DEXTEROUS HAPTIC INTERFACES
F. GONZALEZ, F. GOSSELIN, W. BACHTA
SPONSORSHIP: IEEE WHC 2013. PARTNERSHIP: CEA LIST, ISIR

Haptic interfaces are purpose for the simulation of physical interactions with a virtual environment in a way as realistic as possible. Ideally, the user should be able to grasp and move the interface in the same way as real objects and get the same feedback as for real interactions. However in their design, a trade-off is usually made between interaction capabilities and technological complexity. In particular, dexterous haptic interfaces, which are adapted to the manipulation of virtual objects with the fingers, are usually designed with a focus on the task to simulate.

This paper is aimed at providing a tool to help in the choice of this design, through the observation of the hand contact areas to focus on considering an identified set of tasks. Hand interactions are usually divided into manipulation, in which the hand is used to involve changes on the environment e.g. moving an object, and exploration, characterizing movements performed in order to gain some knowledge on the environment e.g. rub the surface of an object to sense its roughness. The manipulation taxonomies have been investigated along with the exploration ones to identify 39 elementary interaction patterns [1, 2]. They are analyzed in order to define 25 associated hand-environment interaction areas, which are merged together and then broken in elementary contact areas, from the sole pinky fingertip to the whole inner hand surface.

The elementary interaction areas and frequencies of use are merged in order to draw some interaction maps as in Fig. 1. The percentage of use of a contact area over others is given both in figures and grayscale. At first glance, this map fits the intuition: the fingertips are the most utilized parts, with a prevalence of thumb and index. The rest of the fingers is next. Finally, the palm is the least used.

The interaction map gives an insight on how long and which areas are utilized by a particular worker. Although it gives some cues on which elementary areas should be included in the design of a dexterous interface, this information is not sufficient since it does not allow sorting the capabilities of an interface as a function of its complexity. To make this classification possible, a tree is built, which sorts the interaction areas and some of their combinations in function of their constitutive elementary contact areas and the percentage of time they allow to interact naturally with the environment.

Fig. 2 gives an application example of this tree, of which an excerpt is presented at the right, to the sorting of the state-of-the-art dexterous haptic interfaces. The percentage of achievable tasks is considered as a maximal efficiency rate, since it is attained only if the corresponding haptic interface is able to render every other contact modality, like temperature or roughness. This maximal efficiency, combined with other criteria like maximum achievable effort or workspace volume, can be used to compute an image of the actual efficiency of an interface to simulate a general panel of activities. In this context, Fig. 2 shows here that although the maximum achievable efficiency with three fingers (30.2%) is smaller than with five fingers (45.7%), the h8510 haptic interface can be considered as more efficient than the Cybergrasp since its workspace and maximum force rendering are higher.

Finally, from the performed analysis, some guidelines for the design of dexterous haptic interfaces are given. Force feedback should be made available in every direction on fingertips. The robot kinematics should be designed to enhance the interface’s workspace while preserving high force capabilities. Other aspects of the haptic rendering e.g. temperature and texture should also be taken into account to effectively reach the maximum efficiency indexes.

References:
DOF-DECOUpled ACTIVE FORCE SENSING (D-DAFS): A HUMAN-INSPIRED APPROACH TO TOUCH-BASED LOCALISATION TASKS

RESEARCH TOPICS: OBJECT LOCALISATION VIA TOUCH
NICCOLO TOSI, OLIVIER DAVID, HERMAN BRUYNINCKX
PARTNERSHIP: KULEUVEN (BELGIUM)

In the context of touch-based object localisation in structured environments, solving the full 6D problem is computationally expensive as complexity scales exponentially with the number of degrees of freedom. In particular, literature examples present solutions to tasks whose initial uncertainty is limited, thus keeping the problem complexity tractable for online implementation with finite computational resources.

In order to better understand the decision-making mechanism behind touch-based localisation, a test was carried out to observe human beings while facing a blind localisation task [1]. The objective was to identify possible common patterns to be adopted for improving robot performance. A localisation with a wrist force-torque sensor was simulated. Specifically, 30 subjects were asked to localise a known solid object on top of a table using a 20-cm stick. Other forms of sensing were shut down by wearing:

- thick gloves to prevent finger tactile sensing
- an eye-cover to prevent vision
- a headset to prevent hearing.

Two tests were carried out with different initial uncertainty. In both cases, all the subjects followed the same sequence of actions. In particular, some skipped a few steps but never inverting their order. In addition, the difference on initial uncertainty resulted in a change of selected actions, so that it was possible to infer a correlation between actions and a subset of degrees of freedom representing the pose of the object.

Further to the observation of the experiment, a new planning approach, named DOF-Decoupled Active Force Sensing (D-DAFS), was presented for such robotics applications. Specifically, the whole task is divided into a sequence of subtasks, each of them focused on reducing a piece of uncertainty, so that the resolution of the inference model can be increased accordingly. This allows a better allocation of resources, focusing on adopting the most accurate, thus expensive, scheme only when the uncertainty is sufficiently low for on-line calculations.

In order to validate the presented decoupling methodology, a 3-DOF localisation task was performed with a Staubli RX90 robot equipped with a spherical end-effector and a force-torque sensor. This set up allowed the robot to estimate the contact position and normal vector of the touched face. A solid rectangle was localised on top of a table, dividing the whole task into a first 2D exploration subtask followed by a finer 3D localisation subtask after the first contact. Intuitively, this decoupling scheme allowed us to increase the estimation resolution only when the uncertainty was significantly reduced.

Moreover, a series of simulations have been run to prove the effectiveness of DOF Decoupling in coping with high initial uncertainty avoiding computational overheads.

References:
HIGH LEVEL FUNCTIONS FOR THE INTUITIVE USE OF AN ASSISTIVE ROBOT

RESEARCH TOPICS: ASSISTIVE ROBOTICS

PARTNERSHIP: LIMSI, INSTITUT PASCAL, ROBOSOFT

The growth of ageing population is an incentive for policymakers to promote the development of new services, such as robotics. Such services would include improving quality of life and safety of elderly and disabled people. It also aims at compensating for caretaker’s limited availability. ARMEN (Assistive Robotics to Maintain Elderly People in a Natural environment) project aims at the development of a user friendly robot with advanced functions for assistance to elderly or disabled persons at home. SAM (Smart Autonomus Majordomo). Unlike most of assistive robotics systems able to manipulate objects, SAM has been clinically evaluated.

In this paper, focus is given to the robot SAM (Smart Autonomus Majordomo) and its new features of navigation, manipulation, object recognition, and knowledge representation developed for the intuitive supervision of the robot. Results of technical evaluation are also discussed and the paper also documents details of the clinical evaluations.

The robot SAM is made of two parts: a non holonomic mobile platform, RobuLAB10 from Robotsoft and a six degrees of freedom JACO assistive robotics arm manufactured by Kinova. A stereo rig fixed to the gripper enabled SAM to perform object recognition and to control the arm motion automatically with visual servoing.

Software architecture of SAM relies on client-servers. Services such as object recognition, manipulation or knowledge representation are implemented on different servers. The HMI is a client connected to different services through web services (see figure 2). Combination of services enables SAM to offer extended services. The HMI was developed with help of occupational therapists to be as intuitive as possible.

A virtual conversational agent represented by an avatar allows the user to have an empathetic interaction with SAM is also implemented into the HMI (see figure 3).

Evaluation of SAM required two different campaigns: a technical one and a clinical one. Both of these campaigns took place in rehabilitation centers. During these evaluations SAM had to achieve three scenarios which consist in retrieving an object at different location of an apartment upon request of the user.

The technical campaign allowed us to estimate capabilities of different services SAM offers. During this campaign, each service was individually evaluated. Success of this campaign was necessary to begin clinical evaluations.

Clinical evaluation lasted five months. It was carried on with frail and disabled persons. The main objective was to study the usability of an assistive robotic prototype by severely disabled people. Secondary objective was to conduct a study on the merging needs and expectations expressed by users and their caregivers.

ARMEN project was motivated by the results of the ANSO project. Most of the improvements of the system were made in knowledge representation, object recognition and HMI. Results of clinical evaluations were analyzed by biostatisticians and will enable further improvements to our system.

References:
ROBOT ASSISTANCE SELECTION FOR LARGE OBJECT MANIPULATION WITH A HUMAN

RESEARCH TOPICS: COMANIPULATION
JULIE DUMORA, FRANCK GEFFARD, CATHERINE BIDARD, ASPRAGATHOS N. A., FRAISSE P.
PARTNERSHIP: LIRMM (FRANCE)

This work focuses on human robot haptic joint collaboration for achieving large object manipulation tasks. Each partner has complementary capacities: cognitive for the human vs physical for the robot. Thus, a huge challenge is the integration of the complementary capabilities of both partners to create a system that outperforms the capacities of each of them. The robot should assist the human partner who is able to decide on what to do and to instantaneously plan complex scenarios. To that end, the robot has to detect which motion assistance is needed by the human operator at each instant.

Unlike proactive robots, passive assistant robots benefit from the independence of control parameters from the task: no knowledge about the task is needed to perform a motion with an assistance. They also take advantage of an easy predictability from the human partner. Moreover, these robots verify the passivity constraints, ensuring safe interaction with the environment and the human partner. However, the non-holonomic methods, related to existing passive assistant robots, compel the operator to combine a series of motions in order to perform one movement of the object along a prohibited degree of freedom.

In our previous work [2], we thus proposed a control scheme that allows a human to perform complex manipulation tasks jointly with an active robotic partner that fulfill the passivity requirements but offering a quasi-holonomic behavior to the operator. For that purpose, we assume that the robot has no prior information about the task and the environment. We endow the robot with a library of assistances for performing standard collaborative motions. According to the haptic cues naturally transmitted by the human partner, the robot selects on-line the suitable assistance for the current intended collaborative motion. Assuming that many tasks can be decomposed into a sequence of collaborative motions, this control method allows to perform a wide range of tasks with assistance while the robot has no prior knowledge about the task. A user study highlights the advantages of this method compared with a follower robot. We implemented a basic algorithm of intention detection to switch from the current assistance to the next suitable one. This algorithm is based on the relationships between haptic measures and human intention. These relationships were apprehended in a human-robot haptic interaction model proposed and experimentally validated in a previous work [2].

In [2], thresholds were empirically tuned to determine significant haptic measures, useful for the intention detection. This paper [1] aims at automatically tuning the intention detection parameters based on the naive bayes algorithm and the Matthew Correlation Coefficient. Using this method, the user can choose the confidence level of the decision making that the robot should get to select a new assistance.

The method for tuning the intention detection parameters can be applied without demonstrating the switch between motions. As no signal length is suitably defined by the demonstrations, the proposed method selects the optimal number of demonstration data points that will be used to estimate the intention detection parameters in order to get the best detection performance.

References:
FLEXIBLE ROBOTICS - APPLICATIONS TO MULTISCALE MANIPULATIONS

RESEARCH TOPICS: OVERVIEW IN THE FIELD OF FLEXIBLE ROBOTS FOR MANIPULATION TASKS
M. GROSSARD, S. REGNIER, N. CHAILLET
SPONSORSHIP: JOHN WILEY & SONS, INC.
PARTNERSHIP: CEA LIST, CEA LETI, IRCCYN, FEMTO-ST, EPFL, ICUBE, CSEM, SUPÉLEC, ARAID-EU, LIAS, ISIR

Whatever dimensional scale, robotic manipulation tasks that are characterized by a high degree of precision or high dynamics impose constraints on the design or choice of the robotic device which is dedicated to the task. In the vast majority of cases, these involve, whether by deliberate choice on the part of the designer or not, mechanisms that are characterized by mechanical flexibility phenomena.

The book is structured as follows (Figure 1):
– Chapter 1 deals with the general context of the design of functionally integrated microgripping systems. This approach has resulted in a complex and multidisciplinary approach to the problem that benefits from the structures’ flexibility to respond to the challenges of robotic manipulation on a microworld scale.
– Chapter 2 focuses on the dual notations of modal commandability and observability, which play a significant role in the control authority of vibratory modes that are significant for control issues. Several notable properties related to the reduction of models and actuator/sensor co-localization are introduced for the optimal design of flexible manipulators.
– Chapter 3 presents different modeling tools that allow the simultaneous use of energy and system structuring notations. Modeling using port-Hamiltonian systems is specifically examined as it is currently one of the most advanced tools for structured energy modeling.
– Chapter 4 discusses two sensorless methods that could be used for manipulation in confined or congested environments. Open-loop control strategies for flexible microactuators are examined initially because these approaches are particularly interesting when integrating high performance sensors is difficult. The second part of the chapter discusses how the use of bistable mechanical structures enables the creation of the micromanipulation function.
– Chapter 5 analyzes several appropriate approaches for responding to the specific needs required by versatile prehension tasks and dexterous manipulation. Whether they are mechanical transmissions, actuators, kinematic structures or functional surfaces, mechanical flexibility phenomena inevitably appear during the design of versatile grippers or dexterous manipulators.
– After a classification of compliant tactile sensors focusing on dexterous manipulation, Chapter 6 discusses the development of a triaxial force sensor based on piezoresistive technology.
– Chapter 7 deals with the constraints imposed by submicrometric precision in robotic manipulation. A kinematic analysis method for flexible articulations, key elements in robots and high precision mechanisms is also introduced, highlighting the degree of freedom and behavior of simple guidings.
– Chapter 8 presents the essential steps of modeling, identification and analysis of control laws in the context of serial manipulator robots with flexible joints.
– Chapter 9 provides an overview of models for deformable body manipulators. The approach is based on the generalization of the Newton–Euler formalism in the case of deformable manipulators based on the “floating point” approach.
– The last chapter, Chapter 10, presents a set of contributions that have been made with regard to the development of methodologies for identification and control of flexible manipulators based on experimental data.

References:
MODELING AND MOTION CONTROL OF SERIAL ROBOTS WITH FLEXIBLE JOINTS IN FLEXIBLE ROBOTICS: APPLICATIONS TO MULTISCALE MANIPULATIONS

RESEARCH TOPICS: CONTROL OF LIGHTWEIGHT FLEXIBLE-JOINT ROBOT
M. MAKAROV, M. GROSSARD
PARTNERSHIP: SUPÉLEC (FRANCE)

Throughout the book “Flexible Robotics: Applications to Multiscale Manipulations”, the objective is to provide those interested in the field of flexible robotics an overview of several scientific and technological advances in the practical field of robotic manipulation. In particular, a chapter of that book deals with the modeling and motion control of serial robots with flexible joints.

The perfectly rigid joint assumption, often at the basis of the study of manufacturing robots, can be insufficient in many situations. While the stiffness of such robots is optimized in order to guarantee a good accuracy, flexibilities can nevertheless be emphasized in particular operating conditions, especially while carrying heavy loads. Moreover, we are nowadays witnessing the development of lightweight robots, characterized by their low inertia and more flexible mechanical construction (Figure 1). The arising flexibilities can non-exclusively be concentrated either at the transmissions’ level, in which case we talk about flexible joints (the subject of the present publication), or at the segments’ level, which are then modeled as deformable bodies.

The reduced dynamic model of flexible joint robots is first recalled with its remarkable properties. Several identification approaches of this model are then presented and analyzed in terms of their implementation complexity and the instrumentation required by the experimental protocol. Finally, the main theoretical concepts and their application in several practical control strategies are described.

The flexible joint model aims to represent the elasticities assumed to be concentrated in the mechanical transmission chain between the motors and the actuated segments of the robot considered as rigid (Figure 2). These elasticities mainly result from the transmission elements used, including mechanical elements whose stiffness is not infinite, and can become nonnegligible in several situations.

Beyond a purely static study, an identification of the dynamic behavior of these systems characterizing their transient oscillating behavior is essential in the objective of control design. In that publication, we review several experimental identification methods from literature that have been implemented on real robots with flexible joints. The methodologies are classified by the available instrumentation for experiments. Indeed, conventional industrial robots are often equipped with motor angular position sensors only, which does not allow us to measure all dofs generated by the robot flexibilities. A first family of approaches therefore uses additional sensors in order to measure elastic deformations. Another family of approaches only uses the motor sensors for the identification of models of local or global validity.

The main concepts used for the control of flexible joint robots have also been recalled. Under the hypothesis of high stiffnesses, the singular perturbation approach allows us to make the connection with rigid approaches by separating the control of the robot’s slow dynamics from the robot’s fast dynamics. The feedback linearization transforms the initial nonlinear model into a linear and decoupled system, thus allowing the application of linear control techniques. Given the complexity of the practical feedback linearization implementation in the flexible-joint case, model-based compensations can also be made following an anticipation scheme. Finally, in response to particular problems of vibration damping, robustness to parametric and non-modeled dynamic uncertainties, as well as reduced measurements, several more specific approaches have also been discussed in that publication.

Figure 1: Lightweight ASSIST robot arm of CEA

Figure 2: Flexible transmission elements.

References:
DESIGN OF A NOVEL LONG-RANGE INFLATABLE ROBOTIC ARM: MANUFACTURING, ANALYTICAL AND NUMERICAL EVALUATION

RESEARCH TOPICS: INFLATABLE ROBOT
SEBASTIEN VOISEMBERT, NAZIH MECHBAL1, ALAIN RIWAN, AMEZIANE AOUSSAT1
SPONSORSHIP: J. MECHANISMS ROBOTICS 01/10/2013
PARTNERSHIP: ARTS ET MÉTIERS PARISTECH

The aim of this paper is to present the design of a new long range robotic arm based on an inflatable structure. The CEA LRI has developed long-range robot, such as AIA which is an 8m long multi-link carrier with payload up to 10 kg and a total weight of 150 kg, designed to work inside Tore Supra vessel through a small port of 250mm diameter at ultrahigh vacuum (10-6 Pa) and 150°C.

Inflatable robots are lightweight, they have large payload-to-weight ratios, are shock-proof and soft for the environment, which make them suitable to operate among humans. In addition their low cost can address a wide range of new applications.

Within Euler-Bernoulli beam theory, it can be demonstrated that the thinner the wall of a hollow beam, the smaller the deflection caused by its own weight.

But for very thin membrane, buckling effect dominates as soon as compressive stresses appear. Hopefully, the tube can be pre tensioned by filling it with compressed air so that buckling will only occur if the compressive stresses overcome the pre stress.

Inflated tubes can make light links but joints are needed to get a robot. Classical joint between links would ruin the possibility of making a lightweight robot. As a consequence, the tube itself should be bended to get an articulated mechanism.

The first solution that comes to mind is to squeeze locally the tube to ease the bending (Sanan’s robot arm). But all the mechanical characteristics at this point are dramatically reduced and the robot cannot bear orthogonal load anymore.

Another solution was developed for space-suits that allows a toroidal bending of the tube. Since the volume of the toroid remains constant, in theory no torque is needed to achieve the bending.

Two opposite constant length wires concentrate the tensile stress while the remaining surface is folded to make a bellow.

But this has been tested with an internal pressure lower than 30KPa whereas the minimum pressure needed by an inflated beam with the typical dimensions of a long range robot (10m long and 0.2m of diameter) with a 1kg payload is 130kPa.

Numerical calculation with LS-DYNA, an explicit finite element code.

$$\begin{array}{|c|c|c|}
\hline
\text{Joint type} & \text{Max achievable bending angle (°)} & \text{Transverse deflection (mm)} \\
\hline
a & \text{Restriction} & 47 & 20 \\
\hline
b & \text{CEA proto} & 81 & 11 \\
\hline
\end{array}$$

Compared to the Section restriction concept our prototype gives a higher achievable bending angle and has a lower deflection under orthogonal loads.

MANUFACTURING PROCESS

A lot of work has been done with Warein, a specialized textile company, in order to find the most suitable way of achieving the bellows.

References:

AN ADAPTIVE SMOOTHER FOR COUNTING MEASUREMENT

RESEARCH TOPICS: NUCLEAR MEASUREMENT, SIGNAL PROCESSING
R. COULON, V. KONDRAOSVS, S. NORMAND
SPONSORSHIP: CEA
PARTNERSHIP: IMS INNOVATION & MEASUREMENT SYSTEMS

Counting measurement represents a challenge for nuclear instrumentation due to the stochastic nature of radioactivity. Indeed, event counting has to be processed and filtered in order to display a stable count rate value and to allow variation monitoring of the measured activity (Fig. 1). A filter has been developed by the CEA LIST, improving the response time while maintaining count rate stability [1].

The main filter used in industrial implementations is the Moving Average filter (MA) giving the maximum likelihood estimation of the count rate. This linear low pass filter is suitable for estimating a non-varying count rate, but could become irrelevant when a change in radioactivity occurs. Regarding the experimental conditions and the purpose of the measurement the filter is implemented as a preset time rate meter (to ensure a convenient response time) or preset count ratemeter (to provide a precise measurement). A way of reconciling both requirements may be to introduce a weighted function. Therefore, the Exponential Moving Average filter (EMA) is the most frequent implementation. All these linear filters have shown their limits with regard to the trade-off between smoothing and response time. More recently, edge preserving filters were developed into nonlinear algorithm to deal with the issue.

A nonlinear strategy has been developed by the CEA LIST consisting in adapting the integration time retained for a local Maximum Likelihood estimation in accordance to the detection of any abrupt change in the Poisson statistics. The detection method is based on a hypothesis testing model, and allows for a fast decision to be made. A Centered Skellam Test (CST) has been implemented to address the issue.

A test bench has been conducted in order to evaluate the performance of the CST filter compared to linear filters MA and EMA and a nonlinear filter GLR developed by the NCSR “Democritos”. The counting signal is simulated using a Poisson distribution generated as a function of a count expectation. All filters have been tested with regard to performances (accuracy, precision, response time) defined in the international standard IEC 60325 related to radiation monitoring.

A figure of merit (FOM) for a given filter has been calculated in order to evaluate the general performance of filters in terms of precision versus response time compromise (Fig. 2). A FOM below 1 quantifies the improvement ensured by the filter in comparison to the MA filter as far as the response time versus precision compromise is concerned.

The simplicity of the implementation makes the CST algorithm relevant for a broad range of industrial applications, including process monitoring, area monitoring, health physics monitoring, and embedded detection. The CST filter has notably been implemented into a Geiger-Müller probe and set on a robot designed to address safeguard and security. Innovation & Measurement Systems (IMS) has developed an integrated plug-and-play device based on a Geiger-Müller sensor and implementing the filter developed by the CEA LIST. It provides a gamma dose rate with ensured flexibility and convergence of the system [2].

References:
ZnO NANOWIRES AS AN EFFECTIVE LUMINESCENT SENSING MATERIAL FOR NITROAROMATIC DERIVATIVES

RESEARCH TOPICS: ULTRA-TRACE DETECTION, LUMINESCENT INORGANIC NANOSTRUCTURES


SPONSORSHIP: ANR & CBRN-E R&D PROGRAM

PARTNERSHIP: 1CEA-LIST, 2LNIO (UMR 6279 - CNRS, UTT), 3GEMAC (UMR 8635 - CNRS-UVSQ)

Selective ultra-trace detection of gas analytes is a central challenge in the field of sensors. Current research aims at not only improving the performances of existing devices (reducing detection limit, improving selectivity) but also to propose new technological solutions to meet new needs. With this aim in view, we investigate the potentiality of luminescent inorganic 1D nanostructures to significantly enhance the performances of existing devices [1]. Zinc oxide (ZnO) nanowires were selected as active material according to their optical and morphological properties (UV emission, high refractive index (n = 2.57 at 375 nm) for a good optical confinement and high aspect ratio structures). 2,4-dinitrotoluene (DNT) was chosen as the target molecule.

The remarkably improvement in the ZnO quenching response is attributed to the nanowire geometry (higher surface-to-volume ratio) and to the better crystallographic quality of the nanomaterial compared to the thick layer.

This study is the first example in the literature of sub-ppm TNT detection by luminescent ZnO nanostructures. The results demonstrate the importance of nanostructuration in improving the sensitivity of materials and paves the path towards the realization of efficient low-cost sub-ppm detectors.

Fig. 1 presents a SEM image of the studied ZnO nanowires. A 170 nm ZnO thick film was also considered for comparison.

Fig. 2 shows the evolution over time of the room-temperature photoluminescence (PL) of a ZnO thick film and ZnO nanowires upon exposure to DNT vapors. Compared to the ZnO layer, ZnO nanowires are found to exhibit a strong (95%) and fast (41 s) response to low concentrations of 2,4-DNT (180 ppb) [2].

References:


DATA ACQUISITION FOR UNDERWATER EXPLOSIVE DETECTION

RESEARCH TOPICS: UNDERWATER INSPECTION BASED ON ALPHA PARTICLE TAGGED NEUTRONS

V. KONDRAKOV, K. BOUDERGUI, G. CORRE, G. SANNIÉ, B. PÉROT*, C. CARASCO*, C. ELLEON*

SPONSORSHIP: FP7 UNCOSS PROJECT

PARTNERSHIP: *CEA/DEN/LMN, IRB, ACT, JSI, MPS, PORTS OF DUBROVNIK, VUKOVAR AND BAR

Since World War I, ammunition and chemical weapons have been often disposed of by dumping in the sea. The amount of dumped materials presents a challenging issue regarding the safety and security of coastal and underwater areas, ports and ships. The challenge of the FP7 UNCOSS research project [1] is to provide new specialized products and methods. These advanced tools realize fast and robust non-destructive in-situ inspection of underwater objects in terms of chemical analysis. Besides chemical composition, this analysis can provide information on the size of the object, as well as about the volume of the analyzed substance contained by the object.

A system using a neutron generator setup within a Remotely Operated Vehicle (ROV) for underwater inspection has been developed. The system can inspect objects for the presence of threat materials, such as explosives and chemical agents, by using alpha particle tagged neutrons.

The data acquisition electronics (DAQ) developed by CEA-LIST is based on a modular acquisition system technology. The gamma and the nine alpha signals are processed in ten dedicated electronics modules. They generate arrays of time and energy data for all recorded events. The compact modular DAQ system (80x150 mm) digitalizes signals at 800 MS/s. Fig.1 shows ten acquisition modules and one power supply module. Embedded FPGAs perform real time digital signal processing to provide Time of Flight, gamma and alpha energies. The configurable hardware allows more powerful and flexible data analysis.

These data are stored into the embedded PC in the ROV and transferred to the PC located on the surface boat to be analyzed by dedicated software. The software allows to choose the alpha detector pixels and to select the neutron TOF window associated to the inspected object. The gamma-ray spectrum is unfolded into a linear combination of elementary spectra, which allows determining count fractions related to each element present in the interrogated volume of interest. C, N, and O count fractions are converted into chemical proportions to identify the presence of explosives [2,3].

Final demonstration tests have been performed in Punat, Croatia, on a specific location with a depth of 10 m [4]. The ROV landed above the air plane bomb filled with the Si3C7H3N3O6 TNT surrogate without contact and the hydraulic legs allowed positioning the neutron generator very close to the inspected object, see Fig.2. The UNCOSS project has demonstrated the capability of neutron inspection to check the nature of materials inside containers on the seafloor, as for instance fuel, paint, toxic chemicals, or even chemical warfare. The UNCOSS partners also studied terrestrial applications like the detection of landmines, carbon or pollution in soils, based on the concept of a compact system as the one developed for UNCOSS which can be embedded on a vehicle.

Figure 1. The compact modular DAQ system.

Figure 2. ROV during the air plane inspection at the depth of 10 m.

References:
ACTIVE NEUTRON INTERROGATION BASED ON A LINAC FOR NUCLEAR WASTE PACKAGES CHARACTERIZATION

RESEARCH TOPICS: METHODS FOR NUCLEAR WASTE MANAGEMENT
F. CARREL, A. SARI, F. LAÏNÉ, A. LYOUSSI (CEA DEN/DER)
PARTNERSHIP: CEA DEN/DER, CADARACHE

Nuclear waste management is a major concern from an industrial point of view and characterization of nuclear waste packages is a crucial technical challenge for research institutes. Non-destructive active measurements, based on fission process, are well suited for the detection of actinides (uranium or plutonium isotopes) as soon as non-destructive passive methods encounter severe limitations (strong passive background, attenuation of particles spontaneously emitted by the nuclides of interest). Active neutron interrogation is a reference industrial technique, enabling the quantification of the fissile mass inside nuclear waste packages and mainly based on the use of a d-t neutron generator.

In the frame of a PhD work, CEA LIST studied the possibility of replacing the traditional neutron generator by a linear accelerator (LINAC) [1,2]. Neutrons are emitted by photonuclear reactions in the LINAC conversion target, made of tungsten or tantalum. Several improvements can be expected following such an innovative approach. The first one is related to the intensity of the neutron emission which can be two decades higher than those provided by traditional neutron generators, leading to significant improvement of the detection limit for a given nuclear waste package. The second benefit of this approach is related to the possibility of coupling three different non-destructive techniques, based on a single LINAC: active neutron interrogation, active photon interrogation and high-energy imaging.

The first step of this work concerns the characterization of the photoneutron flux delivered by a LINAC [3]. A great number of activation measurements completed by an extensive simulation work were carried out to determine its main characteristics (neutron flux intensity, energy spectrum). Figure 1 gives an example of photoneutron energy spectra determined using the Monte Carlo code MCNPX.

The second part of this work is focused on an experimental validation of this technique. One of the main challenges concerns the possibility of obtaining pure active neutron interrogation, avoiding parasitic contribution of high-energy photon beam and associated photofission reactions. Several measurement campaigns were carried out, considering prompt and delayed neutron counting as well as delayed gamma-ray spectrometry. Figure 2 gives an example of results obtained during these measurements.

Experimental proof of concept of our approach was obtained during these measurement campaigns. Further step of this work concerns the design of an optimized neutron measurement cell and the validation of this non-destructive approach on mock-up packages.

References:
Localization of radioactive hot spots is nowadays a major issue for radioactive waste management, decommissioning of industrial infrastructures, operational radiation protection or Homeland Security applications. Gamma imaging is a very interesting technique in order to address this problematic. It is based on the superimposition of visible and gamma pictures that are acquired using dedicated devices called gamma cameras.

Gamma cameras currently on the market use either coded masks (RadCam from Radiation Monitoring Devices, Inc., CAR-TOGAM from AREVA CANBERRA) or Compton scattering (POLARIS-H from H3D) for spatial localization. Gamma cameras based on coded masks and commercially available can be considered as first-generation cameras as they use scintillation detectors. CEA LIST Institute designed during last years a second-generation system, called GAMPIX [1-2], and based on the combination of a semiconductor detector and a MURA-type coded mask. The detection system consists in a 1 mm thick CdTe substrate bump-bonded to a Timepix readout chip developed by CERN and composed of 256 × 256 pixels, each 55 µm side, including independent shaping and processing chains. The choice of the coded mask is a trade-off between sensitivity (time required to localize a hot spot) and angular resolution. Figure 1 presents the CEA prototype of GAMPIX gamma camera and its main components.

Compared to previous generation, GAMPIX offers drastic improvements in terms of portability (from 15 kg to less than 2 kg), ease-of-use (with a simplified connectivity and interface) and sensitivity. Table 1 shows the performances of GAMPIX gamma camera in terms of sensitivity with a coded mask of rank 7, 8 mm thick, which allows the best sensitivity.

Table 1: Performances of the GAMPIX gamma camera in terms of sensitivity for coded mask of rank 7, 8 mm thick.

<table>
<thead>
<tr>
<th>Radionuclide</th>
<th>Dose rate (µSv·h⁻¹)</th>
<th>Min. exposure time for detection (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>241Am (60 keV)</td>
<td>0.25</td>
<td>1</td>
</tr>
<tr>
<td>137Cs (662 keV)</td>
<td>2.50</td>
<td>20</td>
</tr>
<tr>
<td>60Co (1.25 MeV)</td>
<td>3.84</td>
<td>60</td>
</tr>
</tbody>
</table>

The GAMPIX gamma camera was validated during several in situ measurements, showing its applicability to all of previously cited fields of use. Figure 2 shows some images obtained during these campaigns.

This second-generation gamma camera is currently under industrialization by AREVA CANBERRA, with the extended support of CEA LIST, under the name of iPIX. This industrial version, presented in Fig. 3, shows improvements in terms of ergonomics with a single connectivity and power supply via PoE (Power over Ethernet), automatic mask recognition and highly simplified software, useable by non-experts. The gamma camera is also expected to be used with a Wi-Fi connection and in battery mode. The commercial system is expected to be released in 2014.

References:
HIGH TEMPERATURE RESISTANT FIBER BRAGG GRATINGS FOR SODIUM-COOLED FAST REACTORS CONTINUOUS MONITORING

RESEARCH TOPICS: INSTRUMENTATION OF SODIUM-COOLED NUCLEAR REACTORS (GEN IV)
G. LAFFONT, R. COTILLARD
SPONSORSHIP: CEA TECNA PROGRAM

Both need for sustainable energy and expected increase in world electricity demand in the coming decades have triggered the development of new generation nuclear reactors. Using the R&D development performed within the research program coordinated by CEA, EDF and AREVA on Sodium-cooled Fast Reactors (SFR), the ASTRID prototype project has been launched in 2010. Standing for Advanced Sodium Technological Reactor for Industrial Demonstration, this prototype is considered as a precursor to assess the technological choices in the scope of an industrial grade SFR reactor. Progress in service instrumentation and continuous monitoring of the nuclear core integrity during reactor operation is a major issue. The very high operating temperature of next generation nuclear cores is a great challenge faced by all new sensing technologies. Sensors able to resist sustained periods at temperature up to 550°C in harsh environments as it is the case within the liquid Sodium-based coolant of SFR reactors require specific developments and lifetime evaluation.

Thermal processing of standard FBGs under high temperature using the so-called regeneration process (see Fig. 1) opens the way for new metrological applications of FBG sensors dedicated to harsh environments, especially under high temperature up to 900°C as it is the case for SFR reactor [1]. The operational and reliable use of regenerated FBGs over long period of time (several hours) for high temperature measurements requires to experimentally assess the evolution of their optical characteristics. The reflectivity of regenerated gratings is much more stable than that of standard gratings. But, to our knowledge, no study has been conducted in order to quantify the evolution of the reflectivity of regenerated gratings at high temperature over long period of time. Such inputs are mandatory in order to evaluate the potential of regenerated FBGs for continuous monitoring of structures and processes operating under high temperature and so requiring sensors lifetime of several years. A long term annealing experiment has been conducted at CEA LIST [2]. Four regenerated FBGs written in a singlemode fiber were inserted in a vertical tubular furnace together with reference thermocouples: distances between gratings were chosen according both to the vertical temperature profile of the tubular furnace and to the annealing temperatures expected for each grating (ranging between 760°C and 890°C).

The reflectivities of the four regenerated FBGs are plotted in Fig. 2. A rapid decrease is observed during the first 100-200 hours of annealing before reflectivity’s stabilization to occur. After burn-in, the amplitudes of the Bragg peaks are much more stable but subsequent evolutions are still observed. This result clearly states the high stability of regenerated FBG reflectivities versus standard FBGs which are erased after only several hours at 890°C. Further studies are being conducted to confirm these results, and especially to determine whether it is possible to extrapolate lifetime estimates and/or to compute master curves to anticipate the ageing behavior of regenerated gratings.

But right now, thanks to the regeneration process, high temperature resistant FBGs can be considered as an efficient solution for real life applications and could now be used for numerous industrial involving severe environments.

References:

ANALYTICAL CALCULATION OF ATOMIC EFFECTS IN ALLOWED β- DECAYS

RESEARCH TOPICS: EVALUATION OF NUCLEAR DATA, BETA SPECTRA CALCULATIONS
X. MOUGEOT, M.-M. BÉ, C. BISCH
SPONSORSHIP: LNE

The Laboratoire National Henri Becquerel (LNHB) is in charge of atomic and nuclear decay data evaluations. Several users, from the fields of nuclear industry, medicine and ionizing radiation metrology, have asked for many years a precise knowledge of β spectra, coupled with well-established uncertainties. For that purpose, a study of these spectra, both experimental and theoretical, was initiated.

Atomic effects play an important role in the shape of β spectra, mainly at low energy. The β spectra of 63Ni and 241Pu were recently measured at LNHB using metallic magnetic calorimeters. These transitions are well suited for testing the atomic effects because of their low maximum energies, and because they can be calculated as allowed transitions, greatly simplifying the calculations. The comparison of these two spectra with classical β calculations has highlighted a significant deviation of the measurements below 8 keV that could not be explained [1].

Our calculations were then improved by accounting for the atomic exchange effect. This effect arises from the creation of a beta electron directly in an orbital which was occupied by an atomic electron in the parent atom. Simultaneously, this atomic electron makes a transition to a continuum state, carrying away the β electron energy minus the binding energy of the orbital. This process depends obviously on the overlap of the continuum and bound wave functions. The calculation is greatly simplified for allowed transitions, because only the s orbitals are reachable by the β electron.

The complete procedure was described in minute detail in [2]. Bound and continuum wave functions have been evaluated analytically from strong assumptions regarding the Coulomb potential. A rough screening correction has been added by determining relativistic effective nuclear charges from tabulated mean radii for each orbital. A good agreement with the measured spectrum was obtained for 63Ni decay (see Fig. 1). For 241Pu decay, the calculation has confirmed that this atomic effect explains a large part of the deviation at low energy. However, agreement with the measured spectrum was not as good as for 63Ni decay. After a review of the possible other effects that could explain the remaining discrepancy at low energy, the rough screening correction seemed to be the best explanation (see Fig. 2).

Indeed, for high-Z radionuclides, one can expect the necessity to take into account the spatial variation of the nuclear charge experienced by the ejected electron to accurately correct for the screening effect.

This involves a precise evaluation of the electron wave functions with a complex Coulomb potential which includes an accurate screened potential. Then, a numerical procedure has to be implemented because analytical solutions of the Dirac equation do not exist anymore in that case.

This work was done for the continuum wave functions and a new screening correction was defined. These preliminary results were presented in [3] and led to good agreement with the measured spectra of 63Ni (Fig. 1) and 241Pu (Fig. 2). A consistent calculation of the bound and continuum wave functions is needed to evaluate precisely these atomic effects on β spectra, which will be the purpose of a future work.

References:
EVALUATION AND PUBLICATION OF RADIONUCLIDE DECAY DATA

RESEARCH TOPICS: RADIONUCLIDE METROLOGY
M.-M. BÉ, C. DULIEU, M. KELLETT, X. MOUGEOT
SPONSORSHIP: LNE
PARTNERSHIP: BIPM, IAEA, DDEP

The monograph “Table of Radionuclides” [1] is published in a series by the Bureau International des Poids et Mesures (BIPM) on behalf of the Consultative Committee for Ionizing Radiation (Comité Consultatif des Rayonnements Ionisants, CCRI).

The purpose of this monograph, number 5 in the series, is to present the recommended values of nuclear and decay data for a wide range of radionuclides. Activity measurements for more than sixty-three of these radionuclides have already been the subject of comparisons under the auspices of Section II (dedicated to the Measurement of radionuclides) of the CCRI. The material for this monograph 5 is now covered in seven volumes which contain the primary recommended data relating to half-lives, decay modes, x-rays, gamma-rays, electron emissions, alpha- and beta-particle transitions and emissions, and their uncertainties.

The data are collated and evaluated by an international working group (Decay Data Evaluation Project, DDEP) led by the Laboratoire National Henri Becquerel of CEA LIST. In a first step, the evaluators have agreed on the methodologies to be used. The CD-ROM included with this monograph contains the evaluators’ comments for each radionuclide in addition to the data tables included in the monograph.

The first volumes were dedicated to nuclides commonly used for the calibration of instruments such as gamma spectrometers or of practical interest in various applications, for example: 51Cr, 88Y, 133Ba, 137Cs, 152Eu, etc.

The second volumes were more specifically dedicated to the evaluation of actinides such as the 235U and 239Pu decay chains or 241Am, 237Np, 239Pu, etc. Moreover, the data are permanently updated and, along the volumes, several nuclides were re-evaluated such as 64Cu, 67Ga, 242Cm, etc.

The seventh volume was published in 2013; it contains twenty three new radionuclide evaluations. Experts from the following organizations have contributed to this volume: CEA LIST (Laboratoire National Henri Becquerel), Lawrence Berkeley National Laboratory (LBNL, USA), Khlopin Radium Institute (Russia), Argonne National Laboratory (USA), Horia Hulubei National Institute of Physics and Nuclear Engineering (Romania), CIEMAT (Spain), National Physical Laboratory (UK), and China Institute of Atomic Energy (China). CEA LIST/LNHB was in charge of the coordination, compilation and publication of results.

The work involved in evaluating nuclear data is on-going and the recommended values are kept up to date on the LNE-LNHB website at http://www.nucleide.org/DDEP_WG/DDEPdata.htm.

The publication of further volumes of Monograph 5 will be continued when necessary to add the data of new radionuclides or the re-evaluated data of formerly studied radionuclides in this permanent format that can be referenced easily.

Although other data sets may still be used when assessing the activity of radionuclide sources, the larger use of those recommended data should help to reduce the uncertainties in activity measurements and lead to more coherent results, for example for comparisons.

References:
DIGITAL PULSE PROCESSING AND OPTIMIZATION OF THE FRONT-END ELECTRONICS FOR NUCLEAR INSTRUMENTATION

RESEARCH TOPICS: DIGITAL PROCESSING, RADIONUCLIDE METROLOGY, FRONT-END ELECTRONICS

C. BOBIN, J. BOUCHARD, C. THIAM, Y. MENESGUEN

SPONSORSHIP: LNE

Current digital technology offers the possibility to renew the nuclear instrumentation installed in National Metrology Institutes. The Laboratoire National Henri Becquerel (LNHB) has first investigated digital systems based on FPGA (Field Programmable Gate Array) circuits to perform counting processing and dead-time management as implemented in home-made analog modules specifically designed for radionuclide metrology [1]. Following the long-term experience acquired with those specialized modules, algorithms developed for digital processing are based on extendable dead times associated with the live-time technique.

The interest of digital nuclear instrumentation is the possibility to reduce the analog part of the electronic chain. In that configuration, the digitization stage is performed as close as possible to the detector in order to register all the useful information for subsequent pulse processing (i.e. including the fast part of original pulses delivered by the nuclear detector). As a result, nuclear functionalities originally performed by a shaping amplifier in a spectrometry chain can be directly programmed in digital systems. For that purpose, a new algorithm was designed for counting and pulse-height analysis using the recursive implementation of the Gaussian filter (and its derivatives) originally proposed by Young and van Vliet [2] for computer vision (image filtering, edge detection, etc.). The algorithm was specifically developed for the processing of signals generated by photons detected in a high-efficiency well-type NaI(Tl) detector and a Silicon Drift Detector (SDD) equipped with a reset-type preamplifier. In both cases, the front-end electronics is especially designed to optimize the coupling between the analog part and the digitization stage (14 bit, 125 MHz, installed in a Stratix III development kit). The new electronic interface was realized to achieve a low-detection threshold; it was also designed to handle the high-dynamic range of signals provided by NaI(Tl) detectors in order to limit the influence of saturated pulses generated for instance by high-energy photons or cosmic rays.

As described in Fig.1., the new algorithm processes digitized pulses according to two channels (slow and fast). In the fast channel, the first derivative of the Gaussian filter is applied for the leading-edge detection of pulses. The low-pass feature of the Gaussian filter attenuates the high-frequency component of signals amplified by the derivative process. The identification of a pulse in the fast channel entails the triggering of the extendable dead-time management, the counting and the pulse-height analysis in the slow channel. The detection of local maxima of the 1st derivative can be also used as a Constant Fraction Discriminator in order to reduce the jitter-effect on the signal triggering. In the slow channel, the Gaussian filter is applied as a low-pass filter for pulse height analysis.

The preliminary results obtained with the new algorithm and the optimized front-end electronics are promising [3]. In the case of the well-type NaI(Tl) detector, a low-energy threshold of about 2 keV was obtained (~6 keV for a classical electronic chain). By avoiding the noise generated by the shaping amplifier, this result can be interpreted as a beneficial effect of the front-end electronics especially designed to optimize the transmission of signals provided by NaI(Tl) detector. This feature is particularly interesting for the zero-energy extrapolation implemented in the 4πγ method. In the case of the SDD connected to a reset-type preamplifier, the FWHM energy estimated on the Mn-Kα peak (127 eV at 5.9 keV) is consistent with the value given by the manufacturer. In addition, the paralysis generated by reset pulses is correctly considered in the dead-time processing.

References:
BETA SPECTROMETRY WITH METALLIC MAGNETIC CALORIMETERS

RESEARCH TOPICS: BETA RAY SPECTROMETRY, CRYOGENIC DETECTORS
M. LOIDL, M. RODRIGUES, C. LE-BRET, X. MOUGEOT
SPONSORSHIP: LNE
PARTNERSHIP: KIP HEIDELBERG

The precise knowledge of the shape of beta spectra is necessary in different fields of nuclear medicine and industry, but also at LNHB for ionizing radiation metrology. The theoretical calculation of spectra is often difficult and needs to be confronted with experimental data. However, the experimental determination of the spectral shape is also a difficult task and published spectra are often inconsistent with one another.

LNHB is developing metallic magnetic calorimeters (MMCs) [1] for beta spectrometry. MMCS are thermal detectors operating at very low temperature (10 – 50 mK) and measuring the energy deposited by each individual particle in a (mostly) metallic absorber as a temperature rise by means of a paramagnetic thermometer in strong thermal contact with the absorber. They offer very high energy resolution and low energy threshold. The beta emitter is enclosed inside the absorber. This results in a solid angle of $\frac{4\pi}{sr}$ and a detection efficiency of nearly 100 % within a wide energy range, and avoids several problems inherent to different techniques of beta spectrometry: energy-dependent detection efficiency, electron backscattering at the detector surface, or electron energy loss in the dead layer in the case of a semiconductor detector.

The potential of MMCS for beta spectrometry was tested with $^{63}$Ni, a low energy ($E_{\text{max}} = 67 \text{ keV}$) pure beta emitter decaying via an allowed transition. This allows calculating the spectrum rather reliably for comparison with experimental spectra. A first series of sources was made by depositing and drying small drops of NiCl$_2$ solution. Both gold and silver absorbers enclosing these sources were used.

The experimental spectra (fig. 1) are discrepant from one another and from theoretical spectra that were calculated with the code BetaShape developed at LNHB [2]. It can be seen, though, that the discrepancy between theory and experiment is much smaller when the atomic exchange effect [3] is taken into account. As expected, the absorber material has no noticeable effect on the spectrum.

We suppose that part of the beta energy deposited in the NiCl$_2$ forming the source is expended in the creation of metastable states like electron-hole pairs that do not recombine on a timescale comparable with the thermal time constant of the detector, and is not converted to heat. This fraction of energy depends on the one hand on the initial energy of the beta particle, on the other hand on the thickness of the NiCl$_2$ deposit. This may explain the distortion of the spectra as well as the discrepancy between the spectra obtained with different sources.

Long-lived electron-hole pairs do not exist in metals, so a second series of $^{63}$Ni sources were made by electroplating, resulting in metallic Ni deposits. Fig. 2 presents the spectrum measured with one of these sources. The agreement with the theoretical spectrum calculated including the exchange effect is excellent, confirming on the one hand the potential of MMCS for beta spectrometry, on the other hand the necessity to take into account the exchange effect.

MMCs have proven to be excellent detectors for beta spectrometry, provided that adequate sources can be fabricated from a given radionuclide. The next step is the development of MMCS for higher-energy beta emitters.

---

References:
SIMULATIONS OF THE LNHB MANGANESE BATH FACILITY

RESEARCH TOPICS: NEUTRON SOURCE CALIBRATION, MANGANESE BATH, M-C SIMULATION

F. OGHEARD, P. CASSETTE

SPONSORSHIP: LNE

PARTNERSHIP: EURADOS

The manganese bath facility of the Laboratoire National Henri Becquerel (LNHB) has been renovated to provide enhanced radiation protection and easier maintenance of sources to be calibrated (Fig. 1). A neutron source is calibrated by measuring the resulting activity of a manganese sulfate solution irradiated by the neutrons: the neutron source is placed at the center of a 1-m diameter spherical tank, filled with MnSO₄ solution; the neutrons interact with the manganese through (n,γ) reactions, creating ⁵⁶Mn, a 2.6 h half-life radionuclide. After having reached the saturation state (when the creation rate of ⁵⁶Mn atoms equals their disintegration rate), the radioactive solution is pumped to an in-line detector in order to measure the resulting ⁵⁶Mn activity. The reference emission rate of the source is the ratio of the total ⁵⁶Mn activity of the bath to the probability of creation of a ⁵⁶Mn atom per neutron. This later factor, the bath efficiency, can be calculated using Monte Carlo simulation codes.

For this purpose, this new facility has been modeled using three Monte-Carlo codes: MCNPX, GEANT4, and FLUKA, in order to determine the correction factors needed in the neutron source calibration process [1]. The most realistic source geometry has been determined, and the most reliable cross sections library has been chosen. The model includes not only the manganese bath but also the various structures of the source container and of the containment building. An example of the model geometry of the source container is given in Fig. 2.

The models calculations were compared and discrepancies between the codes have been pointed out. As an example, the probabilities of creation of ⁵⁶Mn per neutron calculated using the three Monte Carlo codes are given in Table 1, with the associated standard uncertainties.

<table>
<thead>
<tr>
<th>Code</th>
<th>Probability of creation of ⁵⁶Mn per neutron</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCNPX</td>
<td>42.715 (7)</td>
</tr>
<tr>
<td>GEANT4</td>
<td>45.25 (2)</td>
</tr>
<tr>
<td>FLUKA</td>
<td>43.47 (7)</td>
</tr>
</tbody>
</table>

Potential causes of deviations between results were assessed and discussed using additional simplified models. Strong disagreements were found between MCNPX, FLUKA and GEANT4, even with simplified models and the later code was finally rejected for this application, as giving unrealistic results. Significant deviations between MCNPX and FLUKA results were still not well understood and an experimental process was proposed to validate the accuracy of the different codes and their abilities in simulating the neutron capture by the manganese bath. This process is based on relative measurements of two different neutron sources using two manganese baths with different dimensions.

---

References:

Tomography reconstruction from projection data is an inverse problem widely used in the medical imaging and NDT fields. With sufficiently large number of projections over the required angle, analytical algorithms allow fast and accurate reconstructions. However, in the case of limited number of views and/or limited angle, the data available for reconstruction are not complete, the problem becomes more ill-conditioned and the results show significant artifacts. Then, alternative approaches based on a discrete model of the problem are available, and consist in using an iterative algorithm to compute an estimate of the unknown object. All these methods are usually based on a volume representation into a grid of voxels and provide maps of densities. Even if these methods remain simple due to the linearity of the projection operator, they are both time and memory consuming.

In addition, post-processing tools are commonly applied on the reconstructed object to segment the region(s) of interest and extract quantitative measurements. Numerous methods of segmentation with different interpretations of the contours and various minimized energy functional are available.

Recently, novel approaches have been proposed in the field of SPECT that simultaneously perform reconstruction and segmentation. The solution we suggest for X-ray tomography reconstruction combines a new representation of the image with simultaneous reconstruction and segmentation algorithms. The method 1/ reduces the number of unknowns and thus decreases the computation time, 2/ makes the mesh faithful to the object and thus reduces the memory consuming, and 3/makes available a density and a segmented material maps. The representation is based on a mesh composed by triangles, and the density values are affected to their faces. Then, the process alternates three main steps: intensity values updates using an iterative algorithm (reconstruction), Euler-Lagrange’s curve motion formulated by level set method (segmentation), and mesh adaptation to the segmented contours.

Figure 1 shows an example of reconstruction of a numerical object composed by one material. It is a 2D orthogonal slice of a numerical knee model (1a) from which 36 projections have been performed using CIVA RT. FBP reconstruction (1b) lets appear artefacts due to the lack of data, while image 1d is more homogeneous. Image 1d shows how the mesh is coarse to represent the background and the inner region of the knee, and how it is thin all around the contour of the knee. The mesh is composed by 515 triangles to be compared to an image of 256x256 pixels (65536 pixels) or more.

Figure 2 shows an example of reconstruction of a numerical object composed by several materials. It is a typical Shepp-Logan phantom (2a) which projections have also been performed using CIVA RT. The object is here reconstructed and represented using 9545 triangles at the end of the process.

Patent in progress, nov. 2012, BD13742, Procédé de reconstruction volumique adaptative en tomographie par rayons X.
Metal implants such as prosthetic devices, dental fillings and surgical clips produce artefacts on the computed tomography (CT) images. The metal artefacts arise during the reconstruction process, and they appear as dark and white streaks and bands in the reconstructed image. The most important effects of metallic objects on X-ray CT scanning are: 1) beam hardening due to the metal object in combination with the broad X-ray spectrum in most CT scanners, 2) poor signal-to-noise ratio from photon starvation, 3) scatter, and 4) edge-gradient effects. Metal artefacts severely limit the CT assessment of soft tissue and skeletal structures surrounding metallic objects. The need for improved metal artefact reduction in computed tomography grows with the increasing importance of CT imaging.

Several image processing solutions have been provided to remove artefacts caused by metal implants. In this work, we suggest to repair sinograms using inpainting that consists in restoring missing image information based upon the still available cues from masked sub-regions of the image. The inpainting/interpolation method used here is based on an appropriate sparse representation dictionary of the image, combining to a statistical point of view that describes inpainting as an estimation problem with incomplete or missing data [1].

In the context of corrective distal radius osteotomy, based on pre- and intraoperative imaging, simulated CT data have been performed using CIVA RX.

Figure 1 shows an example of result (c) of the inpainting method applied on a raw sinogram (a) where two pins appear. It is compared (d) to the corresponding sinogram simulated without any pins.

Figure 2 shows the impact of the correction on the reconstruction. Image (a) shows the reconstruction using the raw sinogram Fig.1a to be compared to image (c) reconstructed from sinogram Fig.1c. The image of reference is (b) that has been reconstructed from a simulated sinogram without any pins.

References:
STRUCTURAL NOISE FILTERING IN ULTRASONIC IMAGING BY DECOMPOSITION OF THE TIME REVERSAL OPERATOR

RESEARCH TOPICS: ULTRASONIC PHASED-ARRAYS - ADAPTIVE INSPECTION METHOD
E. LOPEZ VILLAVERDE, S. BANNOUF, S. ROBERT, C. PRADA
PARTNERSHIP: INSTITUT LANGEVIN/M2M

In ultrasonic imaging, the Total Focusing Method (TFM) is one of the methods that provide an optimal focusing and spatial resolution. In NDT applications, the imaging method is also interesting to enhance the characterization of extended defects, such as cracks.

The TFM algorithm is applied to the full array-response matrix $K(t)$ describing the transmit-receive process of a transducer array. For a $N$-elements array, this matrix is built by recording the $N \times N$ inter-element responses $k_{ij}(t)$ corresponding to the signal received by an element $n°i$ when an element $n°j$ is excited by an electric pulse. As only one element is used per transmission, the main drawback of the TFM imaging is a lower signal-to-noise ratio (SNR) compared to the more conventional imaging methods.

This problem has been overcome by filtering the structural noise with the DORT method (the French acronym for Decomposition of the Time Reversal Operator). This method takes place in the frequency domain and consists in the analysis of the singular values and vectors of the transfer matrix $K(\omega)^{[1,2]}$.

The noise filtering principle consists in identifying and separating the signal subspace from the noise sub-space for each frequency of the transducer bandwidth. The signal sub-space identification is based on cross-correlations of the singular vectors with a reference one (e.g., the singular vector at the central frequency). In the present case where the medium contains only one point-like defect, the signal sub-space is represented by only one singular value (red curve in Fig. 2a). The phase of the associated singular vector provides a constant delay law (Fig. 2b) that focuses on the defect whatever the frequency.

Once the signal sub-space has been identified, a filtered transfer matrix $K_s(\omega)$ is built and an inverse Fourier transform is applied to return in the time domain. Then, the TFM imaging algorithm is applied to the final filtered matrix $K_s(t)$ to obtain an image with a reduced structural noise.

The TFM image obtained after noise filtering is displayed in Fig. 3b. Compared to the initial image (Fig. 1a), the SNR is improved of 30 dB.

Note that the principle of the DORT filtering has been also applied to crack-type defects located near the back-wall of a component. In such a case, the difficulty lies in the identification of the sub-space associated to the back-wall. First promising results have been obtained in a non-noisy steel component and works are in progress.

Figure 1. Austenitic-ferritic steel component and TFM imaging of an artificial defect (hole).

Figure 2. Singular values in the transducer bandwidth (a) and focusing law extracted from the analysis of the associated singular vectors (b).

Figure 3: Singular value associated to the defect (a) and new TFM image with high SNR (b).

References:
ULTRASONIC FIELD RADIATED BY EMAT INTO FERROMAGNETIC MEDIA

RESEARCH TOPICS: ULTRASONIC TRANSDUCTION FOR NDT – MODELING – SIMULATION
C. ROUGE, A. LHÉMERY
PARTNERSHIP: CETIM

INTRODUCTION

ElectroMagnetic Acoustic Transducers (EMAT) can radiate ultrasonic waves without mechanical contact with the piece, allowing the testing at high scanning rate of pieces at high temperature. There are made of a permanent magnet and a coil (Fig. 1), which produce a static and a dynamic magnetic fields.

Three transduction mechanisms are involved in pieces made of a ferromagnetic medium, resulting in three body forces generated by complex interactions of the electromagnetic fields (static and dynamic) with the elastic structure of the medium: the Lorentz, the magnetic and the magnetostrictive forces. In non-magnetic conductive media, only the first is created.

The simulation platform CIVA could deal with the Lorentz force but not with the two others, assuming high bias field and low current intensity. The present work aimed at developing a model dealing with all the three forces and for arbitrary EMAT configuration.

GENERAL MODELING APPROACH

The inputs of the model are the electromagnetic and elastic properties of the piece in which the ultrasonic field is generated, the description of the EMAT (shape, orientation and strength of the magnet, geometry of the coil and time-dependent intensity in it). The model is decomposed in four steps:

i) the static and dynamic magnetic fields and eddy-currents due to the EMAT are computed;

ii) body forces associated to the three transduction mechanisms are evaluated;

iii) body forces are transformed into surface stresses;

iv) surface stresses constitute the input of ultrasonic radiation models.

Interestingly, steps i) and iv) already exist in CIVA in two different modules: CIVA-ET (electromagnetic testing) and CIVA-UT (ultrasonic testing). The two original contributions in this work [1] concern steps ii) published in [2] and iii), published in [3].

EMAT TRANSDUCTION IN FERROMAGNETIC MEDIA - STEP II)

A general model has been developed to deal with the various transduction mechanisms, under the assumption of weak coupling of electromagnetic and elastodynamic phenomena, that is to say, assuming that the elastodynamic fields generated do not modify the electromagnetic fields which generate them. The model successfully predicts theoretical results of the literature obtained in simple classical configurations. For high currents and low bias field, transduction mechanisms are far more complex; in these non-classical cases, which may potentially be of great practical interest, our model successfully predicts results experimentally known of harmonic generation. Harmonic contributions may be of higher amplitude than that due to the fundamental excitation frequency in some circumstances.

TRANSFORMATION OF BODY FORCE INTO EQUIVALENT SURFACE STRESSES – STEP III)

Transductions happen in the volume of the piece in a region limited by the skin-depth of the induced eddy-currents. This region is the source of the ultrasonic radiation. To avoid a costly integration of sources over this 3D region, a mathematical method has been developed to transform body forces into distributions of surface stresses, under the assumption that the skin-depth is shorter than the ultrasonic wavelength. Existing models of CIVA-UT for bulk or guided wave radiation —step iv)— rely on such a surface (2D) source description: they can be readily used once the transformation has been operated. Figure 2 shows examples of ultrasonic fields radiated by the same EMAT in two different ferromagnetic materials.

At CETIM, our partner in this study, experiments are in progress to validate the theoretical models developed. There are still works to be done to extend the approach to surfaces of arbitrary shape —notably, step iii)—, to study the behavior of EMAT working as receiver, to apply the approach to model magneto-strictive transducers.

References:

Figure 1. Schematic representation of an EMAT.
Figure 2. Ultrasonic fields radiated by an EMAT.
INTRODUCTION

Acoustic Emission (AE) is one of the non-destructive methods used in the industry especially for examining pressure vessels (often, mandatorily) and potentially large structures (dams, bridges). A defect present in a structure under stress growths at microscopic scale and emits elastic waves which can propagate over long range. If several wave sensors are positioned on the structure, signals they deliver from the same event are used to detect the presence of the flaw and locate it by triangulation, assuming the speed of waves emitted is known. Practitioners believe that the method could be improved, testing configurations optimized, if the various phenomena involved (at various space-time scales, Fig. 1) were modeled and if tools for simulating them were available.

An ANR-founded project MACSIM [1] and a PhD thesis [2,3] were carried out in parallel to develop models of AE testing and a software demonstrator of its simulation. Project consortium included practitioners of EA (CETIM, DCNS, EADS, EXTENDE), who are CIVA users for other NDT methods, and specialists of modeling and simulation (CEA, UTC), Extende being part of MACSIM notably in testing proposed solutions.

GENERAL MODELING APPROACH

Predicting crack growth in a piece under a given stress is out of the scope of AE simulation. Thus, crack position, dimension and evolution are inputs in our model. The model must deal with phenomena arising at various scales as shown in Fig. 1 and be implemented as a tool running in a PC. The solution consists in chaining several models, each dedicated to one phenomenon at its own space and time scales:

i) elastodynamic states associated to crack growth [µm,µs] (the crack-opening displacement or COD) are computed for modes I, II and III of crack evolution by superposition of static calculations;

ii) the COD becomes source of elastic waves; a model predicting the way it radiates was developed in the moment tensor formalism;

iii) waves propagating in the structure [m, ms] are modeled as Rayleigh or as guided waves, using existing CIVA tools, computation being restricted to directions of sensor positions to minimize computation;

iv) sensor responses to incoming (vector-) waves are computed, this including effects caused by the thickness influences strongly the signal waveform.

To demonstrate the interest of the overall approach, the model has been implemented within CIVA, but restricted to the case of planar parts. Depending on the part thickness, waves emitted by the defect propagate as Rayleigh or as guided waves; the model automatically switches from one to the other. In this demonstrator, the dynamically varying defect size, its position and orientation relatively to sensors, sensor characteristics (size, sensitivity) and part thickness are the free parameters to play with. The three crack evolution modes I, II and III are considered in parallel.

More works are to be done to propose a simulation tool adapted to industrial use but the overall approach and the tools developed constitute a relevant basis to further extensions.

References:

EXPERIMENTAL STUDY FOR THE VALIDATION OF CIVA ULTRASONIC TESTING (UT) PREDICTIONS

RESEARCH TOPICS: ULTRASONIC MODELS
R. RAILLON AND G. TOULLELAN
PARTNERSHIP: EXTENDE

A long-term validation work is being done at CEA in order to precisely validate the models implemented in the CIVA simulation software and consequently to quantify the level of reliability of the predictions provided by the code. The validation process is based on comparisons with experiments [1]. To accurately achieve these comparisons, experimental, simulation and comparison procedures were defined and followed. They are described below for the inspection of a planar block containing back-wall breaking notches with a T45 probe.

An experimental B-scan, showing the “classical” corner and top edge echoes of a back-wall breaking notch is displayed Figure 2.

EXPERIMENTAL PROCEDURE

Specific mock-ups and reflectors were designed and realized to ensure a slow variation of one parameter at a time (the height of the 11 notches varies slowly from 0.5 to 10mm while their depth and extension are constant) and their characteristics were well controlled (shape and dimension of the reflectors and mock-ups, material homogeneity and isotropy). The attenuation and velocity of longitudinal and transversal waves were estimated using successive back-wall echoes and the material homogeneity was experimentally checked. The probe displacement was accurately mastered and its calibration performed over Ø2mm Side Drilled Holes at different depths (calibration block). The profile and top edge shape of the artificial notch are well mastered and provided by the manufacturer. The measurement uncertainty is evaluated by reproducibility measurements.

SIMULATION PROCEDURE [2]

CIVA model: the model to be used to simulate the experimental inspection depends on the physical phenomena involved in the inspection. In the case of the vertical back-wall breaking notches the appropriate model in CIVA is the “KIRCHHOFF+GTD” model which allows computing both the specular corner echo with the KIRCHHOFF approximation and the top edge diffraction echoes with a model based on the Geometrical Theory of Diffraction.

CIVA inputs, values and checks: as seen above the values of the specimen, reflectors and probe parameters used as inputs for the simulation in CIVA are accurately mastered and known from manufacturer and experimentally measured when possible. In addition to the knowledge of these values it is important to check if the influence of all the inspection parameters is taken into account by the model. For example in the “KIRCHHOFF+GTD” model the notch aperture is supposed to be almost zero. But, if the notch aperture is not an influential parameter for the corner echo, we know from literature that it is one for the diffraction echo amplitude and so this echo will not be studied in our example.

MEASURED/SIMULATED RESULTS COMPARISON PROCEDURE

The amplitudes of the echoes are compared but also their shape (A-scans) and time of flight.

These comparison results are discussed considering the physical basis and domain of applicability of the model and the approximations made for their implementation. For example the gap observed Figure 3 for the smallest notch heights illustrates the known limitation of the KIRCHHOFF model in case of small notches dimension relatively to the wavelength. Accurately achieving these discussions allow drawing conclusions on the model validity and limitations.

Figure 1. Steel specimen with 11 back-wall breaking notches of different heights.

Figure 2. Experimental B-scan for a 5 mm height notch.

Figure 3. Comparison of the measured and CIVA amplitudes and A-scans of the corner echoes.

References:

SIMULATION OF GUIDED WAVES INSPECTIONS BASED ON A COUPLING OF MODAL AND FINITE ELEMENTS METHODS

RESEARCH TOPICS: GUIDED WAVES, SCATTERING, FINITE ELEMENTS
V. BARONIAN, K. JEZZINE AND F. LE BOURDAIS

A Hybrid numerical method coupling modal description and finite element computation in the vicinity of guide perturbations has been developed several years ago with POEMS laboratory [1]. This model allows simulating complete NDT inspections (generation, scattering, reception) by guided waves of complex structure. Simulations for the control of critical component constituting the main vessel of the prototype ASTRID (Advanced Sodium Technological Reactor for Industrial Demonstration) are presented.

GENERATION OF GUIDED WAVES THROUGH THE JUNCTION #1

The motivation of this study is to simulate the control of the integrity of welded structures (junction of steel plates) supporting the core of the reactor. Specifically cracks-like defects located in the vicinity of the different welds have to be detected. The inspection is performed from outside the main vessel with transducers positioned along its walls since the inside of the reactor is not accessible. Figure 1 illustrates the geometry of the structure and the inspection configuration.

The issues of this study are to optimize the guided waves generation in the inspection branch, and to study the sensitivity of guided waves with cracks interaction.

Guided waves have been focused in the inspection branch in order to maximize the energy transmitted to the junction #2. Figure 2 represents the elastodynamic field in the time domain, generated through the junction #1 by the transducer, designed to excite T-30° and L-66° bulk waves. The A1 mode is mainly excited at the end of the branch 3a, presenting the advantage of limiting losses along the propagation between the two junctions.

INTERACTION OF GUIDED WAVES CRACKS IN JUNCTION #2

Result of the simulation of guided waves interaction with crack labeled C2 is shown (results with other cracks are given in [2]), as well as the reception on the probe. The main conversion occurs between the A1 and S1 (same propagation properties) modes. Figure 3 corresponds to the pulse-echo signal in terms of the average of the normal displacement component emitted and received on the transducer. The contribution of the modes A1 and S1 are isolated in the received signal.

Perspectives concern the extension of the numerical method to properly take into account the effect of sodium fluid as surrounding media.

References:
MAXWELL’S EQUATIONS IN COVARIANT FORM FOR SIMULATING THE INSPECTION OF A CONDUCTIVE SLAB OF COMPLEX SHAPE

RESEARCH TOPICS: EDDY CURRENT MODELING, MODAL APPROACH, MAXWELL
F. CAIRE, D. PRÉMEL, G. GRANET
SPONSORSHIP: CEA LIST, SIMPOSIUM EUROPEAN PROJECT
PARTNERSHIP: ARCELOR MITTAL

Among the numerous processes used in industry for the preparation of steels, the continuous casting is one of the most critical. During this manufacturing process, the hot metallic liquid is cooled, solidified and finally cut into slabs. Mechanical properties of steel are modified since many internal stresses and cracks may appear into the slab. Moreover, the specimen presents a wavy surface at this stage and the development of fast simulation tools is a crucial step toward the characterization of the interaction between the surface waviness and any crack occurring into the slab.

Eddy current modelling of the inspection of such a conducting slab of complex shape by any 3D eddy current probe is addressed by the curvilinear coordinate method (CCM) usually used in the high frequency domain: we have extended this method to the low-frequency regime in order to tackle some Eddy Current configuration. More specifically, it is based on a judicious change of coordinates associated to the rough surface considered and allowing us to write boundary conditions over the complex shape in an analytical form. The solutions of Maxwell equation are expanded in eigen-functions depending on the analytical expression of the interface.

This approach does not require any mesh and leads to a fast semi-analytical model since the geometry of the interface can be represented by a finite number of basis functions. In this numerical model, the Fourier basis is implemented in order to handle the local smooth shape of the wavy surface. Here, we focus on the computation of the quasi-static field induced in the slab by any 3D EC probe and the response of the probe (impedance). This is the preliminary work before considering the global problem of modelling the response of the crack in the slab.

An experimental validation of the fast numerical model has been proposed in the framework of the European SIMPOSIUM project (standing for Simulation Platform for Non Destructive Evaluation of Structures and Materials) project: Fig.1 depicts the configuration considered: an air-core probe scanning a wavy surface.

Figure 2 depicts the comparison between simulated data and experimental data. These results demonstrate the validity of the numerical model.

The figure 3 depicts another configuration of interest: the complex part constituted by two non-parallel complex interfaces.

Figures 4 and 5 depict a comparison between CCM and Finite Element (FEM) data (field components along the top interface) obtained in this configuration.

Future work will be dedicated to non-homogeneous layers with a conductivity varying along the depth.

References:
THE TRIPLE SINGULARITY: A SEMI-NUMERICAL MODEL FOR THE NEAR-CRITICAL ANGLE SCATTERING

RESEARCH TOPICS: ULTRASONIC SCATTERING MODELS
M. DARMON, S. CHATILLON, P. CALMON
PARTNERSHIP: SOUND MATHEMATICS LTD

This modelling study carried out in collaboration with Sound Mathematics Ltd aims at improving the existing models [1] used to simulate the scattering of waves from an embedded crack near critical incidence.

Numerous phenomena in the fields of physics and mathematics as seemingly different as seismology, ultrasounds, crystallography, photonics, relativistic quantum mechanics and analytical number theory are described by integrals with integrands that contain three coalescing criticalities, a pole, phase stationary point and branch point. Evaluating such integrals is a challenge addressed in this paper. A novel numerical method based on the regularised composite Simpson’s rule is proposed, and its efficacy is demonstrated by revisiting the scattering of an elastic plane wave by a stress-free half-plane crack embedded in an isotropic solid. This scattering problem has a well-known analytical solution expressed in an integral form. In this problem, the reflected, diffracted and head waves can be viewed as contributions of poles, phase stationary points and branch points, respectively. The proposed method is used in a fast and accurate semi-numerical scheme, which allows for description of the non-classical diffraction effects that take place near the critical angle of total reflection. These include the far field spikes in diffraction patterns and near field interference ripples.

Indeed we show in [1] how combining the proposed numerical method with approximate solutions of the problem of scattering by scatterers with sharp edges can be used to extend the limits of validity of the latter. Two solutions of such nature are widely used in applications, the Kirchhoff approximation and GTD (the Geometrical Theory of Diffraction). The Kirchhoff approximation can be used to describe the amplitudes of the bulk waves diffracted by the edge. Nonetheless at small observation angles to the relevant interface GTD provides a much better recipe than Kirchhoff. But GTD is valid only in the geometrical regions where diffracted waves do not interfere with other ones.

Let us now turn to the critical zones ($\Theta = 57^\circ$ or $360-57^\circ$ in Figure 1) where the head waves interfere with the bulk diffracted waves only, that is, the integral contains coalescing branch point and stationary point, but the geometrical poles ($\Theta = 0^\circ$ or $360^\circ$) lie away from both of these criticalities. At smaller distances ($d=10$) to the diffracting edge the critical spikes are smoothed. At larger distances, somewhat surprisingly, the spikes obtained with the numerical model turn out to be the same as the critical spikes in GTD coefficients. Thus, in the very far field, GTD provides an accurate description of bulk diffracted waves in the critical zones, so that the critical spikes in GTD diffraction coefficients are physical in nature.

The conclusion appeals to physical intuition, because the head waves decrease in amplitude faster than the bulk diffracted waves and one would not expect the latter to be influenced by the former in the far field.

The fact that at smaller distances to edge the scattering coefficients exhibit no spikes must be due to destructive interference between the head waves and diffracted waves - with the head waves more prominent close to edge and decaying away from edge slightly faster than the bulk diffracted waves.

Let us now consider the critical zones ($\Theta = 76^\circ$ or $360-76^\circ$ in Figure 2 example) where the head waves interfere with both the bulk diffracted waves and reflected waves. In such zones the integral contains coalescing branch point, stationary point and pole. In presence of reflected waves, GTD is invalid and it is more suitable to use a combination of GTD and Kirchhoff called PTD (Physical Theory of Diffraction) and then apply the numerical method to PTD. In this triple singularity configuration (Figure 2), the conclusions are the same than for the double singularity one (Figure 1), replacing GTD by PTD.

The proposed mixed PTD/Simpson scheme can be applied to simulate not only the bulk waves diffracted from edges but also the head waves. Its efficacy has been demonstrated by investigating the critical regions where the head waves interfere with bulk diffracted waves and even reflected waves. An implementation of the proposed mixed PTD/Simpson scheme is envisaged in CIVA (software platform developed at CEA/LIST and widely used in NDE) in order to improve simulation of the ultrasonic response from large cracks near critical angle.

Figure 1. Scattering numerical coefficients vs GTD ones versus observation angle (a) $d=10$, (b) $d=500$.

Figure 2. Scattering numerical coefficients vs PTD ones versus observation angle (a) $d=10$, (b) $d=500$.

References:
HEAD WAVES RAY TRACING ON IRREGULAR SURFACES FOR TOFD ULTRASONIC INSPECTION

RESEARCH TOPICS: ULTRASONIC WAVE PROPAGATION
A. FERRAND, M. DARMON, S. CHATILLON
PARTNERSHIP: I2M (UNIVERSITÉ DE BORDEAUX 1)

HIGHLIGHTS
– This paper provides a modeling study of head waves near irregular surfaces in NDT.
– Head wave propagation near such complex surfaces implies bulk mechanisms.
– A generic algorithm of ray tracing between interface points (GIRT) is developed.
– GIRT, based on Generalized Fermat’s Principle, models all waves propagating near complex surfaces.
– The head wave fronts computed by GIRT are in good agreement with FEM simulations.

The TOFD (Time Of Flight Diffraction) technique is a classical ultrasonic inspection method used in ultrasonic non-destructive evaluation (NDE). This inspection technique is based on an arrangement of two probes of opposite beam directions and allows a precise positioning and a quantitative evaluation of the size of cracks contained in the inspected material thanks to their edges diffraction echoes. Among the typical phenomena arising for such an arrangement, lateral waves also called head waves, since they are chronologically the first waves reaching the receiver, are notably observed. Figure 1 shows the typical waves occurring in TOFD inspection over a planar specimen containing an embedded crack-like flaw:
– The lateral wave (1)
– Diffraction echoes (2-2’) are scattered by the edges of the flaw, and caught by the receiving probe.
– The backwall echo (3), due to the specular reflection of the beam on the component backwall.

Head wave propagation on planar surfaces in TOFD configurations is well known since it propagates along the specimen surface. However, realistic inspection configurations often involve components with irregular surfaces, like excavated specimens.

The previously developed head wave simulation can only be applied for a planar specimen. A PhD thesis has thus been carried out with the aim of modeling lateral waves on irregular surfaces.

In [1], complete interpretation and simulation of the head wave propagation mechanisms for irregular surfaces are proposed. This approach leads to the development of a generic algorithm of ray tracing between interface points (called Generic Interface Ray Tracing/GIRT) which is able to compute the travel path of head waves in specimens inspected by TOFD technique. Based on an adaptation to the NDE domain of seismic ray tracing algorithms, GIRT is a new method for solving the two-points ray tracing problem, which is the calculation of the true ray between two defined points: the source and observation points.

The GIRT algorithm is adapted to find not only the head wave ray path (first arrival) but also all complex later arrival waves including mode or nature conversion by adding constraints on the searched ray path. Secondly, GIRT can take into account all types and natures of waves propagating in NDE. Indeed the propagation simulation is available for both longitudinal and transverse waves and accounts for modes conversions. Surface waves as Rayleigh waves and head waves are taken into account as well as diffractions from surfaces in any kind of waves. Finally, GIRT approach is extended to flaw scattering modeling since the flaws contained in the specimen are also meshed by the GIRT algorithm in order to model rays diffracted from flaws. Consequently the GIRT algorithm is generic for application to NDE as it can deal with any ultrasonic wave propagating near irregular surfaces or flaws of any geometry (CAD defined).

By modeling the complex wave interactions with the surfaces thanks to this ray approach, GIRT is thus able to correctly predict the head wave time of flight and to interpret the propagation of complex waves at the vicinity of an irregular interface.

Indeed, comparisons between GIRT and FEM simulations of the head wave fronts near the receiver surface show good agreement (Figure 2) for an excavated specimen (whose entry surface is in red) and validate GIRT algorithm. The head wave GIRT simulated rays (in white) partially propagate in the specimen bulk.

For corrugated surfaces, a bulk propagation mechanism has thus to be taken into account for the head wave modeling. As valid head waves rays are determined by GIRT, ray models for head wave amplitude simulation will be investigated thereafter. GIRT would then lead to a complete modeling of the head wave propagation on irregular interfaces.

Figure 1. Illustration of TOFD inspection technique over a planar specimen and related echoes for an embedded crack.

Figure 2: FEM snapshot (in color code) and GIRT simulated head wave front (in black) received by the transducer. At right: zoom on the receiver.

SIMULATION OF ULTRASONIC INSPECTIONS OF WELDS USING A PARAXIAL RAY TRACING METHOD

RESEARCH TOPICS: ULTRASONIC MODELS, RAY TRACING
A. GARDAHAUT, K. JEZZINE AND N. LEYMARIE

On-site inspections of bimetallic or austenitic welds can be very difficult to interpret owing to their internal structure. Skewing and splitting of the ultrasonic beam may occur due to the anisotropic and inhomogeneous properties of the welding material. The understanding of such phenomena requires the use of accurate numerical models. We propose a modeling approach based on a Kinematic Ray Tracing system, to evaluate the ray trajectories and the travel time, and a Dynamic Ray Tracing system to compute the amplitude along the ray and in its vicinity. This method has been numerically validated against a finite element code in 2D, and applied to an actual mock-up of a bimetallic weld [1].

DYNAMIC RAY TRACING MODEL FOR A SMOOTH DESCRIPTION OF WELD

The DRT model has been applied to inhomogeneous media. This model is based on the evaluation of the ray trajectories and the travel-time, at the same time as the ray amplitude, during the propagation. To evaluate the ray-paths and travel-time, we have to solve the eikonal equation. By deriving this equation, written in Hamiltonian form, a differential ray tracing system called the axial ray system is expressed. This system is composed of two ordinary coupled differential equations describing the variation of the position x and the slowness p with respect to the travel-time. On Figure 1, the trajectory of a ray through a V-weld, evaluated with the dynamic ray tracing model, is shown.

To describe the conservation of energy inside a ray tube and compute the amplitude of a ray, the transport equation has to be solved along a ray in an anisotropic inhomogeneous medium. This leads to the so-called paraxial ray system, from which the amplitude of the wave-field is computed.

VALIDATION OF THE METHOD BY COMPARISON WITH FE CALCULATION

The ultrasonic wave-field computed with the DRT model in a V-weld has been compared to results obtained with a 2D hybrid code [2] involving a FE computation inside the weld. Results are shown in figure 2. The comparison of the maximum particle velocity obtained with the DRT model applied on a smooth cartography of the crystallographic orientation shows an excellent agreement with the hybrid code results.

APPLICATION ON A WELD MOCK-UP

Finally, the method has been applied in 3D to a realistic weld mock-up. To obtain a cartography of grain orientation, an image processing technique has been developed and applied on the original macrograph of the weld (Figure 3). This cartography has been subsequently used as input data for the ultrasonic propagation model.

The simulated results have shown a very good agreement with experimental data when the flaws are located outside the weld in the base metal. Nevertheless, discrepancies have been observed when the notch is located at the weld seam, possibly because of the weld description being not accurate enough in this area. Future works on this topic deal with the improvement of computation time and numerical precision by using higher-order numerical methods such as the common fourth-order Runge-Kutta method.

References:
GENERIC SIMULATION OF ULTRASONIC ECHOES FROM PLANAR CRACKS BY COMBINING THE KIRCHHOFF AND GTD MODELS

ANNUAL SCIENTIFIC REPORT
2012-2013

Non Destructive Evaluation often aims at detecting cracks in structures. As a result, there is a significant interest in the simulation of NDE towards the prediction of echoes from such defects. The ultrasonic testing simulation module of the CIVA software offers ways of computing crack echoes, including semi-analytical models and a finite element method. The semi-analytical models have the advantages of allowing fast simulations compared to numerical models. However, the validity domains of these models tend to be limited by their approximations.

In the CIVA 10 version, two semi-analytical models could be used in the case of cracks: Kirchhoff and the Geometrical Theory of Diffraction (GTD). They are based on different approximations and have complementary validity domains. Choosing between the two models required expertise. Moreover, in some ambiguous cases, no model was entirely satisfactory. The PTD model [1,2], recently developed and made available in CIVA 11, aims at overcoming these limitations.

COMBINING CRACK SCATTERING MODELS

The echo coming from a crack depends linearly of the displacement that occurs at its surface when it is exposed to an ultrasonic field. Scattering models differ in the way they obtain this displacement [2]. The Kirchhoff model makes an approximation that is valid along the surface of the crack but fails at its edges, whereas the GTD model relies on a canonical solution that is valid at the edge but fails to account for the surface of the crack. It results in the two models being suitable for different types of echoes.

The Physical Theory of Diffraction (PTD) was developed as a way to combine the advantages of these two models. Its principle can be summarized by the following expression:

$$PTD = Kirchhoff + GTD - \text{Correction}.$$ 

It sums the two models and subtracts a corrective term. This correction is another scattering model which has the remarkable feature of accumulating the shortcomings of both Kirchhoff and GTD: it fails at predicting edge phenomena in the same way as Kirchhoff, and has the same problem as GTD in accounting for crack surface. As a consequence, this correction will cancel Kirchhoff in the calculation of edge echoes and GTD in the calculation of surface echoes. In both cases, the PTD result will coincide with the result of the remaining model, which will be the valid one.

COMPARISON OF RESULTS

This approach, summing two usual approximate models and subtracting an ad-hoc corrective term, remains fast compared to numerical approaches such as finite elements.

The following plot compares the result of the PTD model with the Kirchhoff and GTD ones and with a finite element method. It corresponds to an inspection when a probe rotates around a cylinder containing a crack, which allows seeing it from several directions.

The finite element computation of this example took approximately 18 minutes, whereas the three semi-analytical computations took each less than 1 second.

PTD (green) is in good agreement with the finite element method (black) for all angles. The other two models prove to be inaccurate for some angles: Kirchhoff (red) for low angles, which correspond to edge diffraction, and GTD (blue) for high angle, which correspond to surface echoes.

These results illustrate the improvement due to the new PTD model: instead of having to choose between Kirchhoff or GTD with their limited validity domains, the CIVA user can use PTD and have a valid result over the entire range of angles, in a reasonable computation time.

Figure 1. Right: simulated echo amplitudes in the inspection of Figure left versus the incident angle on the crack, 90° being the normal incidence.

References:
HOW TO DETERMINE POD CURVES FROM SIMULATION RESULTS WITH CONFIDENCE?

RESEARCH TOPICS: NDT PERFORMANCE DEMONSTRATION
N. DOMINGUEZ, C. REBOUD, A. DUBOIS, AND F. JENSON

Performances and reliability of NDT operations are often quantified through the determination of Probability of Detection (POD) which can then be used in lifecycle management approaches like the damage tolerance design. Recent progresses in NDT models and POD methodology using uncertainty propagation through simulation codes have demonstrated the relevance of simulation to establish POD curves. Besides obvious cost reductions, an advantage of using simulation for POD is that it potentially gives access to very large datasets, which results in reducing the width of confidence bands. On the other hand a major difficulty in this approach is often to describe the statistics of uncertain parameters, which intuitively affects the “confidence” one can have in the obtained POD results. A new approach of “confidence” for POD estimation from simulations has thus been proposed [1]. It consists in computing a confidence band based on the level of knowledge that we have on the statistics of the uncertain parameters rather than on the number of available data. This approach is also a way to introduce “conservativity” in POD with simulation, which is generally appreciated in industrial practices.

CONFIDENCE APPROACH FOR POD USING SIMULATION

The aim of the confidence band is to cover for the “lack of knowledge” in the estimation process. In the classical context of POD estimation using experimental data, the “lack of knowledge” comes from the limited number of data and the confidence band characterizes the POD estimation precision. In the context of POD using Uncertainty Propagation through simulation codes (POD-UP), the “lack of knowledge” comes more from the way to describe the uncertainties on the inputs of the simulation than from the number of produced data. For instance, choosing a probability distribution and setting the parameters describing this distribution results from an engineering judgment which is sometimes a difficult task and may be subject to some arbitrary. Therefore the confidence for POD-UP has to somehow represent this “lack of knowledge”.

The “classical” scenario of POD-UP is to:
1. Define uncertainties on the identified uncertain parameters,
2. Compute the associated POD curve.

In this scenario, some of the distribution parameters are well known and mastered and can be considered as “sure”. Some other distribution parameters may be difficult to appreciate and their values result, at some level, from an arbitrary decision.

For these “not sure” distribution parameters, one could have chosen other values (“neighbor” values) for the “not sure” values, and obtained other POD curves - one POD curve by POD scenario (figure 1). The proposed confidence calculation method consists in using the scattering on the output POD curves to calculate a confidence band at % by leaving % of the computed POD curves on the left of the confidence curve. The POD curve with % confidence is the set of points shown on figure 1. Practical implementation of this method requires description of the degree of knowledge on the distribution parameters of the POD study. To manage that step we have introduced a scale of “certainty indices” which apply on the distribution parameters and describe the “level of certainty” that one has on these values (figure 2).

References:
A BAYESIAN APPROACH FOR THE DETERMINATION OF POD CURVES FROM EMPirical DATA MERGED WITH SIMULATION RESULTS

Inspection reliability is one of the key issues in ensuring safety of critical structural components. Among the various methods dedicated to NDE performance evaluation, probabilistic approaches have aroused interest due to their ability to naturally account for uncertainties related to the large number of factors influencing inspection results. Within this approach, the NDE performances are expressed using probabilistic criteria such as the Probability Of Detection (POD). POD curves relate the detectability of a flaw to its size. Following a general procedure, POD curves are estimated from NDI reliability data that come mostly from dedicated round-robin inspection programs. The whole process is expensive and a time consuming. In recent years, it has been proposed to replace some of the required empirical data by simulated inspection results obtained with physics-based models [1]. In this study, a way to merge both types of data is proposed. The approach relies on a Bayesian updating algorithm which allows us to update the knowledge or information that we have on the POD curve parameters using both simulated and experimental data.

The underlying idea of this work [2] consists in both minimizing the amount of empirical data required to determine a POD curve and mitigating the risk introduced by the use of numerical simulation results in the process. A convenient way to achieve this task is to work within a “Bayesian” framework in which the Bayes’ rule will be used to update the probability estimate for a hypothesis, as additional evidence is gathered. The posterior probability is determined from two antecedents, a prior probability and a “likelihood function” derived from a probability model for the data to be observed. Here, the data to be observed are inspection outcomes and the probability to observe a specific outcome follows a Bernoulli law. The likelihood function of the POD curve parameters given the observed data indicates how likely chosen values of the parameters are, in light of the observed outcome.

The selected application case has been described in several papers [1]. It corresponds to a High Frequency Eddy currents Testing (HFET). An experimental dataset of 345 values was available for this application case. Inspections were also simulated with the software CIVA. This has been done following a Monte-Carlo approach which has been extensively described in previous papers [1].

SIMULATED AND EXPERIMENTAL DATA FOR A HFET INSPECTION OF CRACKS IN TITANIUM PLATES

The Bayesian updating scheme has been applied starting from non-informative prior distributions (uniform distributions) for the POD curve parameters. The distributions are then updated according to the scheme described previously and using the simulated dataset (figure 1). In a second step, a Bayesian update was performed using the experimental data set. Here, the prior distributions are those obtained previously, i.e. after performing a first Bayesian update with the simulated data set. These results show that adding experimental data in the Bayesian estimation process does not modify strongly the median POD curve. Thus, the POD curve determined using simulated data was close to the empirical POD curve.

Figure 1. Bayesian updating scheme.

Figure 2. POD curves estimated from the median of the prior distribution obtained using simulated data (dashed line) and posterior distribution updated using experimental data (solid line).

References:
This work aims the effective simulation of complex configurations of nondestructive testing by electromagnetic methods. The complexity lies in the geometry, but also in the coexistence of different models (static, quasi-static or dynamic). In particular, many sensors (typically made up of a coil, a ferrite core and an electric shielding) involve this type of configuration, for which we do not have analytical solver and therefore have to use a numerical solver.

Surface integral equations reduce the three-dimensional problem to an equivalent surface problem by taking into consideration the transmission conditions at the interfaces between homogeneous media and the radiation properties in each medium. Nevertheless, the commonly used surface integral formulation of the Maxwell problem is unstable at low frequency or in the presence of high-contrast environments. In practice, a specific formulation is retained for each model, and difficulties still arise in highly conductive and permeable media. The specificity of this work is to seek a generic integral equation that can be adapted to each class of media to facilitate the cohabitation of the subdomains.

The limits of the commonly used formulation are induced by a scale effect between the contributions of the system that results in the prominence of digital noise. A Helmholtz decomposition of the approximation space separates the potentials at the origins of this problem and pushes significantly the limits of the model, even if the resulting system still suffers from bad conditioning.

An investigated solution, inspired from applications on dielectric materials, is to rescale the decomposed system via a normalization matrix. However, this normalization was derived from an asymptotic expansion of Green’s function that is no longer valid in the presence of conductive objects because of the high contrast between the air and the medium. The introduction of weighting coefficients associated with each medium overcomes this difficulty by splitting contributions from dielectric and conductive parts. But it brings up a twisted identity term that is not well-discretized by the Galerkin approach, and which requires the use of a specific dual test space.

Several formulations have been proposed and evaluated the first two years of this study. In particular, two formulations were selected and tested to treat the frequency-conductivity-permeability range of electromagnetic testing applications [1]. The study continues with the search of a method of domain decomposition by physical sub-sets (such as the electric and magnetic current densities) to reduce the size and improve the conditioning of the system [2].

This work is already used to address the simulation of steam tubes testing.
MODELLING OF FERRITE CORED COILS WITH COMPLEX SHAPE FOR ELECTROMAGNETIC TESTING

RESEARCH TOPICS: EDDY CURRENTS, INTEGRAL EQUATIONS, COUPLING
E. DEMALDENT, A. SKARLATOS, C. REBOUD, A. VIGNERON
SPONSORSHIP: CEA

A typical eddy current testing configuration consists in displacing a probe (e.g., a coil) on the surface of a conductive part (the inspected workpiece) to detect the presence of a cut which leads to a variation of the signal. The presence of a ferrite core in the probe greatly improves the electromagnetic coupling between the probe and the conductive part and therefore the signal-to-noise ratio. Numerical simulation is a precious tool as it helps the interpretation and prediction of signals or the estimation a priori of the efficiency of a configuration (design of sensor). However, the presence of ferrite cores prevents the use of conventional semi-analytical calculation tools and complicates significantly the electromagnetic problem considered.

As an alternative approach to its complete numerical resolution, the complete problem can be solved as two interacting sub-problems, corresponding to probe and conductive domains, respectively. Therefore, each sub-problem is modeled independently of the other, using the most effective technique, while their magnetic coupling is evaluated by successive iterations.

The initialization process of the coupling is to calculate the magnetic field radiated by the probe in air at the location of the workpiece interface, by computing the core's response to the field induced by the coils. Hence no interaction between the probe and the conductive part is present at this step. As the conductivity of the core is assumed to be zero, this particular sub-problem is quasi-static but can be solved with a magneto-static formulation. To this end a high-order boundary element method was used to solve the single layer surface integral equation [1].

We then calculate the workpiece's response to this field, which is the magnetic field reflected by it at the surface of the ferrite core. When the shape of the workpiece is canonical, this calculation can be performed by a fast and accurate modal approach. We recalculate the field radiated by the probe at the surface of the workpiece, but this time by taking into account its previous response in the incident field.

We then iterate until convergence. In practice, it takes a few iterations to converge (less than 10 iterations with a stopping criterion set at 1/1000). After convergence of the iterative process, the primary electromagnetic field is obtained in the conductive part.

This coupling procedure was developed, initially, to treat tilted cylindrical ferrite cores with an axisymmetric finite integration technique solver. Solving surface integral equations by the boundary element method was later considered to handle 3D ferrite cores with complex shape. The advantage of using a boundary element method resides, in addition to the implied boundary conditions, in a significant decrease in the number of unknowns. We first applied this procedure to model +Point [2] and Rototest [3] industrial probes.

References:
Biomechanical risk factors assessment of a work activity is usually based on the study of a human operator’s postures and forces while performing the work task. Hence, assessing the ergonomics of a future work-station at the design stage requires that an operator performs the work on a prototype or a similar equipment. An alternative solution has emerged through the use of digital human models (DHM) for ergonomics analysis. Yet, using industrial DHM software packages available for ergonomic assessment is usually a complex and time-consuming task.

A challenging aim therefore consists in developing an easy-to-use DHM capable of computing dynamic, realistic movements and internal characteristics (position, velocities, accelerations and torques) in quasi-real time, based on a simple description of the future work task, in order to achieve reliable ergonomics assessments of various work task scenarios at an early stage of the design process.

We have developed such a dynamic DHM [1] automatically controlled in force and acceleration, inspired by human motor control and based on robotics and physics simulation. More precisely, we implemented a modified minimum jerk criterion with via-points to perform trajectories which take into account Hyck-Hyman’s law, Fitt’s law and kinematic invariance (2/3 power law). (See [2] for mathematical details)

In our simulation framework, the DHM motion is dictated by real-world Newtonian physical and mechanical simulation, along with automatic control of applied forces and torques.

Our controller handles multiple simultaneous tasks (balance, contacts, manipulation) in real time along with human-like feedforward force and impedance control.

An experimental insert-fitting activity has been simulated (see Figure 2) and assessed based on the OCRA ergonomic index. Following an assessment process described in Figure 1, a comparison with experimental human data showed consistent results: joint torques, DHM movements and their related OCRA assessment were realistic and coherent with human-like behavior and performance (Figure 3 – more details are given in [2]).

The main interest of our DHM is that it requires minimal information for a simulation: a starting point, an intermediate point for obstacle avoidance and an end point, along with the applied force for insert clipping. Moreover, changing the subject’s anthropometry and the scenario does not require new trajectory specification nor additional tuning.

Figure 1: Proposed methodology

Figure 2: Simulation Environment

Figure 3: Results – similar figures with Human and DHM

References:
A prime functionality of a virtual character is to perform manipulation tasks. The choice of postures can impact the possibility of fulfilling a task successfully. For example, the virtual character may have to lean forward to push an object, but lean backward to pull it; the feet may have to be separated from each other to be able to generate manipulation forces that are sufficiently strong. These questions suggest that before performing a task, it is important and beneficial for the virtual character to choose postures that are optimal for the task.

In the context of computer animation where motion capture has become an essential technique, an operator's postures can be taken as references for the virtual character. Captured motions are lifelike, but they need to be adjusted to handle manipulation forces and to deal with disturbances. This is because the operator usually does not really manipulate objects and thus cannot sense interaction forces between the virtual character and the virtual environment. Consequently, his postures can be inappropriate for the virtual character to balance the interaction forces or to improve task performances.

We have developed a generic approach that can automatically find optimal postures for a wide variety of manipulation tasks. For each manipulation task, a constrained optimization problem is solved off-line to find a sequence of optimal postures associated with a desired manipulation path, in the neighborhood of a given initial posture. The optimization problem is formulated based on a simplified model of the character's states (Figure 1), including interaction forces with the environment and the kinematic relations among key frames. Once a solution is found, one can use a motion controller to make the character adjust the contact positions, then to perform object manipulation by following the desired motions of the key frames. Our approach considers quasi-static cases where dynamic effects can be ignored; therefore, we use the quasi-static controller described in [1]. Our approach is summarized in Figure 2.

The main contributions of our approach are as follows:

- It is the first generic posture optimization approach that couples geometric and kinematic constraints with force and moment constraints.
- It can improve task performance by looking ahead; and it can be used to evaluate a task before actually executing it.
- It deals with the redundancy of poses, and helps to make contact positions as robust as possible. The structure of our posture optimization problem allows us to take precaution against mechanical interactions and possible perturbations. By adopting the optimized postures, the risk of failures either due to poor postures or due to perturbations can be greatly decreased.
**PASSIVE HIERARCHICAL CONTROL BASED ON WRENCH BOUNDS**

**RESEARCH TOPICS:** DIGITAL HUMAN MODEL, HUMANOID WHOLE-BODY CONTROL

MINXING LIU (CEA LIST), ALAIN MICAELLI (CEA LIST), PAUL EVRARD (CEA LIST), ADRIEN ESCANDE (CEA LIST), CLAUDE ANDRIOT (CEA LIST)

Interactive, digital humans, seen as a form of simulated intelligent robots, usually need to interact with human operators and with virtual environments. They should react to operator’s instructions and simulated events. Our work focuses on real-time interactions where an operator guides the digital human using motion capture techniques and can interact in an unpredictable way, such as during reaching or manipulation tasks in training environments or in industry design.

There are two main difficulties in this research topic. First, the operator’s posture can be inappropriate for the digital human, because the operator usually cannot sense forces that are applied on the digital human during task execution. Consider the scenario where a digital human pushes an obstacle; the operator just sends out the intention of pushing by reaching out his hands, but the digital human may have to lean toward the obstacle to push it while the operator cannot, because the virtual obstacle does not exist in the world of the operator. Hence, captured motions should be adjusted to be more suitable to handle external contact forces during interactions with the environment. Second, while multiple tasks are performed simultaneously, some tasks can be incompatible with one or another. Therefore, task conflicts have to be handled. To address these problems, we investigate a two-level prioritized control framework, where multi-objective control is combined with motion capture techniques to retain the advantages of each. The digital human is considered as a mechanical system which is influenced by multiple wrenches. The motion of each task frame is guided by its task wrench. Multiple tasks, as well as task priorities, are handled by regulating task wrenches. Wrench bounds are imposed on lower priority task wrenches to ensure that they will not drive a higher priority task frame out of its admissible domain.

Figure 1 shows the controller’s architecture.

The contribution of this work is the extension of the multi-objective control framework described in [1] with a prioritized control approach based on wrench bounds. This approach allows inequality constraints on a higher priority task, and takes into consideration the passivity of the system, which is the main difference with state-of-the-art hierarchical controllers, as shown in [2]. Moreover, our approach is suitable for interactive applications such as virtual reality. With the application of wrench bounds, the operator can provide any reference motion without being worried about the balance of the digital human. The digital human can perform a wide variety of tasks, such as reaching and manipulation, and keep balance while trying to increase its workspace (Figure 2).

---

**References:**


HUMANOID PUSH RECOVERY CONTROL IN CASE OF MULTIPLE NON-COPLANAR CONTACTS

RESEARCH TOPICS: HUMANOID CONTROL, PUSH RECOVERY, LOCOMOTION

DARINE MANSOUR (CEA LIST), ALAIN MICAELLI (CEA LIST), PIERRE LEMERLE (INRS)

SPONSORSHIP: INRS

PARTNERSHIP: INRS

We present a method for humanoid push recovery in the general context of multiple non-coplanar contacts [1] as most of state-of-art push recovery algorithms only consider coplanar feet contacts.

The method consists of a controller that minimizes the kinetic energy of a perturbed whole body humanoid system, while controlling the support change to achieve the stabilization (or push recovery) of the system. The controller uses a simple model based approach to determine the necessity of support change and in this case to approximate a new contact position that allows stabilizing the system.

Our approach can be explained as follows. We consider that the whole body model, initially at rest, is subject to horizontal perturbations during a short time duration. We use the times (t1, tB, tBR, tE, Reflex and Step times – see Figure 1), to characterize the push recovery process. For the whole body model, the application of Fpert (as perturbation) begins at tI and ends at tB, with (tB ≠ tI) and we consider short time disturbances, such that tBR ≥ tB.

The whole body model is controlled by a “whole body controller” developed by CEA LIST [2] until it recovers its stability at full stop. The recovery method consists of combining a simple model based approach with a whole body controller as follows:
1) Between tI and tBR, the controller minimizes the kinetic energies (EKCOM – for CoM (Center of Mass) translation, EKINT – for global rotation) of the whole body model.
2) At tBR, the whole body model and its control are suspended and offline operations are conducted using the simple model based approach:
   - The simple model is initialized with the same CoM state and contact configuration of the whole body model at tBR. Since we consider short time horizontal perturbations, the vertical CoM velocity is negligible at tBR, which is consistent with the simple model.
   - A fall indicator determines the necessity of support change for stabilization.
3) The whole body model and its control are resumed at tBR, after the offline operations. In case of necessary support change, it is controlled between tBR and tE. The control pursues until a full stop is reached.

The method has been tested on a simulated humanoid robot and it succeeded in stabilizing the robot for coplanar and non-coplanar environments. Figure 2 shows a non-coplanar Push Recovery clip.

References:
A HIERARCHICAL FRAMEWORK FOR REALIZING DYNAMICALLY-STABLE MOTIONS OF HUMANOID ROBOT IN OBSTACLE-CLUTTERED ENVIRONMENTS

RESEARCH TOPICS: DIGITAL HUMAN MODEL, ERGONOMICS
ZHAOPENG QIU (CEA LIST), ADRIEN ESCande (CEA LIST), ALAIN Micaelli (CEA LIST), THOMAS ROBERT (IFSTTAR)
SPONSORSHIP: IFSTTAR
PARTNERSHIP: IFSTTAR

Realization of a humanoid’s dynamic motions in a cluttered environment involves multiple problematics including motion planning, balance maintenance, collision avoidance, motion control, inverse kinematics (IK), etc. Recent progress in researches on these problematics brings more and more exciting achievements, but up to now, there is not yet an universal and efficient approach for realizing humanoid motions in rather cluttered environments. This study [1] aims to explore an efficient approach to this problem.

Consider a Digital Human aiming at realizing a dynamic multi-step motion via a series of support (contact or grasp) configurations. The scheme of the hierarchical framework in our study is shown in Figure 1. It generates and realizes the humanoid’s motion in a cluttered environment via three steps:

1) At the global level, by pre-defining a sequence of support configurations, a global robust CoM (Center of Mass) trajectory as well as the most adequate timing information is firstly generated by an optimization based method. The humanoid robot maintains its dynamic balance during the motion by tracking this trajectory. (See [2] for specific details)

2) At the local level, the whole-body collision-free motion that tracks the imposed CoM trajectory is generated piece wisely for all the transition phases between stances. A local sampling-based method is associated with a flying end-effector (a hand or a foot) for planning locally its trajectory.

3) At the control level, the generated trajectories (CoM, end-effector, joints) serve as control references so that the humanoid can realize the motion by virtue of dynamic controller.

This framework has been tested for a simplified car-ingress scenario. The motion execution at the control level is realized using XDE® software developed by CEA-LIST. Results compared to motion capture data show quite good similarity. Future studies will focus on motion naturalness based on joint postures during the task.

Figure 2 shows a clip of the dynamic simulation.

Figure 2: Simulation Clip for Car-ingress scenario.

References:
CONTROL OF RADIATIONS FOR HEALTH
COMPARISON OF ABSORBED-DOSE-TO-WATER UNITS IN DIFFERENT BEAM SIZES AND BEAM ENERGIES BETWEEN LNHB AND PTB

A work package of the Euramet project ‘External Beam Cancer Therapy’ was dedicated to the realization of references in terms of absorbed dose to water under Intensity Modulated Radiation Therapy conditions. Two Primary Standard Dosimetry Laboratories (PSDL), Physikalisch Technische Bundesanstalt (PTB, Germany) and Laboratoire National Henri Becquerel) (LNHB, France), have established absorbed-dose-to-water references for different photon beam energies and beam sizes on their own linear accelerators. To look at the agreement between the absorbed-dose-to-water references of PTB and LNHB is of interest for both laboratories, especially as they do not rely on the same kind of primary standard instruments, a graphite calorimeter for LNHB and a water calorimeter for PTB [1].

The LNHB reference of absorbed dose to water is based on graphite calorimetry. The ratio between absorbed dose to graphite and absorbed dose to water is calculated with Monte Carlo codes (EGSnrc and PENELOPE):$$D_w = D_{core} \left( \frac{D_w}{D_{core}} \right)_{MC} \Pi_k$$with Dw, the absorbed dose to water at the reference point and Dcore, the mean absorbed dose in the core of the graphite calorimeter. The subscript “MC” denotes values calculated by Monte Carlo and Πk corresponds to the product of correction factors for graphite core impurities and for the conversion of [Dw]MC calculated in a water finite volume to Dw at a point. The reference point is at 10 cm depth along the beam axis in a water phantom of 30 × 30 × 30 cm³. References in terms of absorbed dose to water have been determined for the 6-MV flattening-filter free, the 6-MV and the 12-MV beams of the LNHB linear accelerator Saturne 43 GEMS and for beam sizes of 10 cm × 10 cm, 4 cm × 4 cm and 2 cm × 2 cm. The standard uncertainties (k = 1) of the calibration coefficients of the reference ionization chambers lie between 0.33 % and 0.46 %.

During the project, it has been shown by LNHB and ENEA–INMRI (Italian PSDL) that the usual beam specifiers (quantity related to the beam energetic spectrum) measuring the beam attenuation in water on the beam axis under different conditions, TPR20,10 and %dd(10x), are inadequate as they are very sensitive to the beam size: for large fields, a significant part of the measured signals on the beam axis is due to photons scattered from outside the beam axis to the beam axis. Therefore, it has been decided to use the calculated water-to-air stopping power ratio (Sw,air = (S/ρw)/(S/ρair)) as beam quality specifier, because the comparison is done with air-filled ionization chambers and because this is the parameter contributing to the ionization chamber calibration coefficient that shows the largest variation with beam quality. The stopping power S is the mean energy lost by the electrons per length unit.

Figure 1 presents the comparison results as a function of Sw,air with an Exradin A1SL ionization chamber. Except for the LNHB point corresponding to the 6-MV 4 cm × 4 cm field, the points as a function of Sw,air are very well aligned compared with their uncertainties. PTB calibration coefficients are all higher than LNHB ones. All the comparison results, based respectively on water calorimetry and on graphite calorimetry, are within 1.5 standard deviations.

Figure 1. PTB and LNHB calibration coefficients for the Exradin A1SL no 100986 for different field sizes (10x10 and 3x3 cm² for PTB at 6 and 10 MV and 10x10, 4x4 and 2x2 cm² for LNHB at 6 and 12 MV) as a function of Sw,air. Sw,air increases with the field size. Uncertainties correspond to one standard deviation.

References:
**ABSORBED DOSE TO WATER STANDARDS ESTABLISHED BY WATER CALORIMETRY AT THE LNE-LNHB FOR MEDIUM ENERGY X-RAYS**

**RESEARCH TOPICS:** ESTABLISHMENT OF NEW ABSORBED DOSE STANDARDS  
B. RAPP, M. DENOZIÈRE, J DAURES, J-M BORDY  
**SPONSORSHIP:** LNE

Medium-energy x-rays, whose tube high voltage is between 80 and 300 kV, are used in radiotherapy for low-depth dose delivery (from millimeters to a few centimeters). These beams are used in some skin cancer treatments; they are also used for palliative care. The other field of application is imaging.

Nowadays, the national standards available for medium-energy x-ray beams are mostly established in terms of air kerma. To determine the absorbed dose to water (quantity of interest in radiotherapy) in their beams, medical physicists have then to use transfer dosimetry techniques based on air-kerma calibrated ionization chambers, which are described in international radiotherapy protocols (like IAEA and AAPM protocols). The protocol IAEA TRS-398, based on absorbed-dose-to-water calibration coefficients of ionization chambers, is rarely used, because of the lack of standards available today in national metrology laboratories.

This is why, in order to be able to apply dosimetry protocols based on absorbed dose to water, LNHB has set up new standards of absorbed dose to water using the water calorimetry technique for six medium-energy x-ray beams (from 80 to 300 kV).

The basic principle of water calorimetry is to determine the absorbed dose to water by measuring the temperature rise under irradiation knowing the specific heat capacity of water. A new water calorimeter [1] (see Fig.1) was designed and set up at LNHB to perform measurements directly at a depth of 2 cm, which is the reference depth recommended by protocols for medium-energy x-ray beams. To make it possible, the cylindrical quartz vessel containing the temperature probes is inserted into the front face of the water phantom. This is the first calorimeter able to realize measurements at such a low depth. The quartz vessel contains a small volume of ultra-pure water saturated with N2 gas (to control the heat defect of water). The calorimeter is designed to operate at 4°C (temperature of the maximum density of water, in order to minimize problems due to water convection). The absorbed dose rate to water was measured for the six medium-energy x-ray beams. The absorbed dose to water obtained has a relative combined standard uncertainty between 0.49 % and 0.72 % depending on the x-ray beam quality.

Four protocols based on air-kerma standards were applied to two ionization chambers for the six medium-energy x-ray beams. The comparison performed between protocols shows a good agreement, whatever the beam or the chamber used [2]. In the same way, the comparison between absorbed dose rates to water obtained by water calorimetry and by application of the air-kerma based protocols were also found in good agreement. A maximum deviation of 2.1 % was registered, with an uncertainty below 0.8 % (k = 1) for water calorimetry and around 2.5 % (k = 1) when applying protocols. So the results are consistent, but it is shown that a significant reduction of uncertainties on absorbed dose rates to water is obtained using water calorimetry.

Then, the application by medical physicists of protocols based on absorbed dose to water, like IAEA TRS-398, will be possible. This should bring a significant reduction of uncertainties, a simplification of the procedure and thus a limitation of the potential mistakes in the determination of the absorbed dose to water.

**References:**


INTRODUCTION

The National Physical Laboratory (NPL), the Physikalisch-Technische Bundesanstalt (PTB) and the Laboratoire National Henri Becquerel (LNHB) have been involved in the EURAMET project “External Beam Cancer Therapy”. The electron paramagnetic resonance (EPR)/alanine dosimetric method was chosen for performing measurements in small photon fields used in radiotherapy [1], [2]. In this context, these laboratories compared the results of their alanine dosimetric systems (detector, modus operandi, etc) at radiotherapy dose levels to check their consistency. This comparison (EURAMET. RI(I)-S7) was performed using two photon beams (60Co and one Linac beam) in each laboratory. Each laboratory had both to irradiate the dosimeters of the two other participants and to measure its dosimeters once irradiated by them. The Bureau International des Poids et Mesures (BIPM), as a neutral organization, was charged of collecting the results and, only after receiving all of them, to distribute them to the three participants, to ensure that the comparison was “blind”.

IRRADIATION GEOMETRIES

The alanine dosimeters of “measuring” laboratories were irradiated at doses known by the “irradiating” laboratory and given in terms of absorbed dose to water. These irradiations were made under the reference conditions of each laboratory in a 60Co beam and in an accelerator beam (10 MV or 12 MV) in a water phantom of 30×30×30 cm³, or larger (Fig.1.), using a square field of 10×10 cm² (at the reference depth). Irradiations were performed at doses within 10 % of the nominal values of 5 Gy and 10 Gy.

DOSE MEASUREMENTS

Each “measuring” participant read out its own dosimeters and assessed the doses using its current protocol (calibration curve, positioning device, …), the aim being to compare the whole process.

Measurements were made at NPL and PTB using a Bruker EMX spectrometer with a ST4102 and a ER 4119 HS resonator respectively, and at LNHB using a Bruker Elexsys spectrometer with a ER 4119 HS resonator. The pellets of the dosimeter are placed in a quartz tube introduced in the spectrometer. All the participants measured the amplitude of the central peak of the resonance spectrum. This quantity is proportional to the number of free radicals created by the irradiation, and then to the absorbed energy.

RESULTS

The results are summarized in Table 1.

CONCLUSION

The maximum deviation in the ratio of measured to delivered doses is less than 1 % (except result 5, rejected after the identification of an experimental problem due to climatic conditions). The results demonstrate the effectiveness and consistency of the EPR/alanine dosimetry systems operated by national metrology institutes as a method for assuring therapy-level doses with the required accuracy. The three different approaches to position the dosimeters in the phantom were shown to be equally effective and convenient to use.

![Figure 1. Irradiation of dosimeters at LNHB](image_url)

<table>
<thead>
<tr>
<th>Beam</th>
<th>Irradiation</th>
<th>Analysis</th>
<th>Mean ratio</th>
<th>u (ratio)</th>
<th>Analysis</th>
<th>Mean ratio</th>
<th>u (ratio)</th>
</tr>
</thead>
<tbody>
<tr>
<td>60Co</td>
<td>LNE-LNHB</td>
<td>1 NPL</td>
<td>0.9999</td>
<td>0.0091</td>
<td>2 PTB</td>
<td>1.0087</td>
<td>0.0062</td>
</tr>
<tr>
<td>60Co</td>
<td>LNE-LNHB</td>
<td>3 NPL</td>
<td>1.0043</td>
<td>0.0098</td>
<td>4 PTB</td>
<td>0.9992</td>
<td>0.0073</td>
</tr>
<tr>
<td>60Co</td>
<td>PTB</td>
<td>5 LNHB</td>
<td>1.0348</td>
<td>0.0080</td>
<td>6 NPL</td>
<td>0.9962</td>
<td>0.0079</td>
</tr>
<tr>
<td>60Co</td>
<td>PTB</td>
<td>7 LNHB, 2nd</td>
<td>1.0099</td>
<td>0.0068</td>
<td>8 NPL</td>
<td>0.9976</td>
<td>0.0129</td>
</tr>
<tr>
<td>12 MVX</td>
<td>LNE-LNHB</td>
<td>10 LNHB</td>
<td>1.0094</td>
<td>0.0109</td>
<td>11 PTB</td>
<td>1.0027</td>
<td>0.0080</td>
</tr>
<tr>
<td>10 MVX</td>
<td>LNE-LNHB</td>
<td>12 LNHB</td>
<td>0.9883</td>
<td>0.0079</td>
<td>13 NPL</td>
<td>0.9991</td>
<td>0.0090</td>
</tr>
</tbody>
</table>

References:


CHARACTERISATION OF THE X-RAY TUBE SPECTRA USING SEMI-COCONDUCTOR DETECTORS

RESEARCH TOPICS: IMPACT OF THE SPECTRUM VARIATIONS ON THE DOSE STANDARDS
J. PLAGNARD, M. DENOZIÈRE, I. AUBINEAU-LANIÈCE, S. DELOULE
SPONSORSHIP: LNE

In the field of low- and medium-energy photon dosimetry (below 300 keV), a good knowledge of the energy spectra emitted by X-ray tubes is crucial. In order to ensure the quality of the metrological chain between the national and accredited calibration laboratories, it is essential that each of them uses exactly the same beam quality. Therefore, international norms which describe the characteristics of several reference beams (i.e. ISO 4037-1: relative to X and gamma radiation standards for calibrating dosimeters and flow meters) have been developed. Due to the strengthening of controls of radiation systems used in industrial and medical sectors, the Laboratoire National Henri Becquerel (LNHB) is increasingly involved in calibration activities in the low- and medium-energy range. The direct measurement of the spectra emitted by those systems is becoming an interesting solution to verify the beam quality. A research program was then initiated at LNHB to develop systems able to measure the spectra emitted by an X-ray tube using semiconductor detectors [1].

The first system developed by LNHB was tested using the X-ray tube used at the radiobiology department of the Gustave-Roussy Institute (IGR France). The measurements of spectra were carried out using a CdTe detector. To reduce the photons high flux rate, a tungsten collimator having a small 100-µm aperture was fixed on the entrance window of the detector. In order to obtain a good alignment between the collimator and the beam axis, a specific unique device, including two automatic rotation stages has been built (Fig.1). A dedicated Labview® program was developed to control the detector position, to find the position corresponding to the highest count rate and to store spectra automatically.

Due to interaction processes in the detector and the surrounding structures, various distortions are present in the measured spectra. Two algorithms were developed to correct for the pile-up distortions due to the high count rate and for the photon escape phenomenon which takes place in the detector (CdTe or Ge detectors).

Several spectra were measured with different filtrations and corrected using both algorithms. At the same time, the programs XCOMP5 and SpeKcalc V1.0 were used to calculate the spectra emitted by the X-ray tube. The quality of the calculated data depends on the knowledge of the internal components of the X-ray tube. Our work showed that the experimental spectra can reveal some important information, such as the thickness of several unknown materials (e.g. the thin layer of barium used to trap the residual gases inside the X-ray tube).

The comparison between the calculated and measured spectra shows a good agreement (Fig.2). It has been demonstrated that the few discrepancies, which can be seen on the different spectrum shapes, have a negligible effect on the mean energy of the beam.

Other spectrum measurements were carried out for several X-ray tubes used at LNHB using the CdTe detector and also a Ge detector. Again, a good agreement has been observed between the germanium and CdTe corrected spectra, which demonstrates the robustness of our method.

The developed system allows the accurate spectrometric characterization of unknown photon fields. This method avoids the knowledge of the tube voltage, of filtrations or of half-value layers (HVL). This information is not always available with a good precision. It can also be used outside the direct beam, for example in medical applications (radiology) where the scattered field coming from the patient has to be assessed in order to know the energy distribution of the photons contributing to the operator exposure.

References:
Radiotherapy modalities are now more and more complex and accurate dose calculations are essential for treatment planning. However, in presence of small field sizes and light densities, commercial treatment planning systems (TPS) often fail to predict accurate dose distribution. To remedy this, we have developed a new Monte Carlo (MC) system called PENSSART for PENELOPE Simulation for the Safety in Radiotherapy.

The PENSSART system uses the 2006 release of the PENELOPE code adapted to allow for the transport of particles in voxelized geometries. PENSSART is in itself a TPS that uses PENELOPE to perform dose calculations in voxelized geometries. The PenSSaRT system is divided into three modules (Fig. 1). The dose calculation module is the core of the PenSSaRT system as it includes the MC dose computation engine itself which is based on PENELOPE. This module was designed to perform MC dose calculations within voxelized geometries and needs as input data a description of the patient anatomy or phantom geometry as well as a description of the radiation source. The patient module allows the implementation of complex geometries such as phantoms and patients. The radiation source module is also based on PENELOPE and allows the simulation of different kinds of radiation sources, going from simple sources, such as monoenergetic beams, to more complex sources like the one resulting from the complete modeling of a linac treatment head, including the multileaf collimator.

As PENELOPE has already been widely benchmarked against other codes, a special attention was paid to the benchmark of the major functionalities introduced in the PENSSART platform, i.e. the patient anatomy description and the beam settings implementation [1].

Physical anthropomorphic phantoms have been used to compare measured and simulated radiation doses in the cases of breast and lung treatments. Measurements were performed with Optically Stimulated Luminescence (OSL) dosimeters while the PENSSART system was used to determine dose distributions. Differences between OSL measurements and PENSSART estimations (Table 1) are of magnitude similar to that of the statistical uncertainties of the MC simulations (about 3%). On the contrary deviations of TPS calculations from OSL measurements are larger especially in lung and vary from 21 to 71%. We also note that the differences between TPS calculations and PenSSaRT estimations are close to the differences observed between OSL and TPS values.

Table 1. Comparison between simulated and measured dose values.

<table>
<thead>
<tr>
<th></th>
<th>OSL</th>
<th>TPS</th>
<th>PENSSART</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung</td>
<td>6.89</td>
<td>2.00</td>
<td>6.61</td>
</tr>
<tr>
<td>Bone</td>
<td>43.4</td>
<td>34.3</td>
<td>44.7</td>
</tr>
<tr>
<td>Bone</td>
<td>3.07</td>
<td>2.00</td>
<td>3.01</td>
</tr>
</tbody>
</table>

The results showed that the system did not introduce any bias in the dose calculation due to an error in beam settings or in geometry model. Moreover the system also allows determining the limits of a TPS algorithm in the cases of breast and lung treatments. In a near future, we plan to extend the functionalities of this tool by incorporating dynamic motions [2] to allow the simulation of leaves movements during the treatment (intensity modulated radiotherapy and arctherapy).

References:
The significant rise of medical imaging exams in the past few years has led to an increase of the collective dose due to irradiation. The medical physics community agreed that this increase must be accompanied by a strong understanding of the radiation dose and its associated risks. Most of already existing tools do not take the person biological specificities into account and thus only provide common dose index and effective dose rather than actual organ doses. We proposed to replace the non-individualized indices commonly used to estimate the radiation dose with an accurate individualized radiation dose evaluation obtained by Monte Carlo (MC) simulations.

The PENRADIO system is based on PENS-SART, a system initially developed for the safety in radiotherapy. The system uses the 2006 release of the PENELOPE code adapted to allow for the transport of particles in voxelized geometries. Different kinds of radiations sources are supported by the system but for radiologic applications, an energy spectrum provides accurate enough dosimetric results.

In conventional (CR) and interventional (IR) radiology, information provided by the manufacturer in the technical notes is sufficient to model the X-ray tube in the MC tool. The CR and IR tube models were adjusted and validated using half-value layers (HVL). Physical anthropomorphic phantoms (CIRS ATOM dosimetry phantoms) and their DICOM images were then used to compare measured and simulated radiation doses in the case of a hip CR and in the case of a cardiac IR. Measurements were performed with Optically Stimulated Luminescence (OSL) dosimeters inserted in the phantoms while the PENRADIO software was used to determine the dose in some organs of interest [1, 2]. MC dose maps are reported in Fig. 1.

We obtained a deviation of less than 20% between simulated and measured dose values.

In computed tomography (CT), the technical notes do not provide enough information to model the X-ray tube and filtrations, especially because of the particular shape of the bowtie filters. To overcome this problem, we have adapted a method presented in the literature [2]. The equivalent inherent filtration was first determined thanks to measured HVL and an iterative process (Fig. 2). The equivalent bowtie filter shapes were then determined using a similar methodology based on experimental profile measurements. The results we obtained for the spectra and the bowtie filters are in agreement with our expectations.

The first results obtained with the PENRADIO software are encouraging. The validation of the program is part of the ongoing work for several phantoms and examination procedures in CR and IR as well as in CT exams. We expect the same accuracy in the results that we have obtained up to now.

Figure 1. On the left, male patient in a hip conventional radiography examination at 77 kV. On the right, female patient in cardiac interventional radiography examination at 68 kV.

Figure 2. Flowchart for the experimental determination of the inherent filtration.
HIGH-RESOLUTION PORTAL IMAGE PREDICTION FOR RADIOTHERAPY TREATMENT VERIFICATION

RESEARCH TOPICS: RADIOTHERAPY, MONTE CARLO SIMULATIONS, DENOISING, TREATMENT VERIFICATION

D. LAZARO, E. BARAT, C. LE LOIREC, T. DAUTREMER, T. MONTAGU, L. GUERIN, A. BATALLA

PARTNERSHIP: MEDICAL PHYSICS DEPARTMENT, F. BACLESSE CENTRE, CAEN

The need of treatment verification in external radiotherapy has become increasingly important these past years, driven by the complexity of current delivery techniques such as intensity modulated radiation therapy (IMRT). Electronic portal imaging devices (EPIDs) are very valuable tools both for patient positioning and dosimetric verification of complex treatment plans. Verifications can be performed by comparing the image acquired with the EPID before or during the treatment with a reference image. This reference image can be calculated very accurately by Monte Carlo (MC) simulations, but this approach is still computationally intensive. This work investigated the possibility of computing the reference EPID image with a resolution similar to that of the acquired image, while keeping computing times compatible with clinical use. To this end, a new methodology combining MC simulations with a new denoising algorithm was developed [1].

This methodology proposes to split the reference image computation into two stages. First, the EPID image is simulated using the MC code PENELOPE. An accurate MC model of the OptiVue1000 EPID (Siemens), integrating a non-uniform backscatter modelling, was developed and validated against measurements. The resulting simulated image could be very noisy, and needs to be smoothed in a second step. The denoising algorithm must have given properties: it must be able to smooth highly pixelated and noisy images, while preserving all the fine details brought in the image by the MC simulation. Two denoising algorithms already employed in radiotherapy were tested, namely IRON (Iterative Reduction Of Noise), and LASG (Locally Adaptive Savitzky-Golay filtering). A third denoising algorithm, based on a nonparametric Bayesian framework and called DPGLM (for Dirichlet Process Generalized Linear Model), was also developed. Smoothing capabilities and computing performances of the IRON, LASG and DPGLM methods were compared for EPID images computed for different statistical uncertainties (up to 10%) in different irradiating configurations.

A profile drawn through the EPID image denoised with IRON, LASG and DPGLM is shown on Fig. 1. Results clearly demonstrate that DPGLM outperforms both IRON and LASG by providing better smoothing performances and a better robustness with respect to noise. Additionally, no parameter tuning is required by DPGLM, which makes the denoising step very generic and easy to handle for any EPID image. Concerning the computation time, the denoising of 1024 × 1024 images takes about 1 h 30, 2 h and 5 min using DPGLM, IRON, and LASG, respectively.

This paper shows the feasibility to predict within a few hours and with the same resolution as real images accurate EPID images, combining MC simulations with the DPGLM denoising algorithm. Figure 2 shows an example of beam verification for a prostate treatment plan.

![Figure 1. Profiles drawn through the acquired EPID image (black line), the MC reference image undenoised (cyan crosses), the MC reference image denoised with IRON (green), LASG (blue) and DPGLM (red), for a statistical uncertainty of 5.7%.](image1)

![Figure 2. Comparison of the acquired EPID image before treatment with the reference image provided by the developed methodology.](image2)

References: