The Faculty of ICT's Exhibition Event

FICTeX 2017

13 & 14 JULY 2017

10TH ANNIVERSARY EDITION

UNIVERSITY OF MALTA
Faculty of Information and
Communication Technology

RakeTech

THE BEST FIRST JOB YOU'LL EVER HAVE

Visit **www.raketech.com** for our current vacancies

KPMG CRIMSONWING

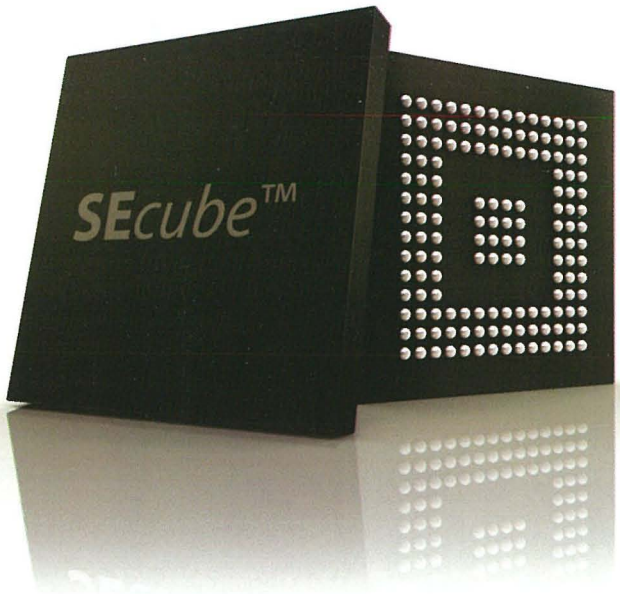
Take a step in the right direction

KPMG Crimsonwing is an international business & IT solutions provider that supports organisations in business transformations enabled by technology. Our work focuses on Enterprise Resource Planning (ERP), Customer Relationship Management (CRM), eCommerce and IT integration.

Working with us

- Business travel opportunities
- Large international clients
- Annual salary reviews
- Long service awards
- Performance-related bonuses
- Working on cutting-edge technologies
- Health insurance
- Life insurance
- Ample employee car parking
- Further studies encouraged
- Career development
- Outstanding training opportunities

We are always looking for talented ICT graduates. If interested in working with us, send us your CV at careers@crimsonwing.com



Open Hardware Platforms Developed in Malta to deliver seamless Trusted systems to the world

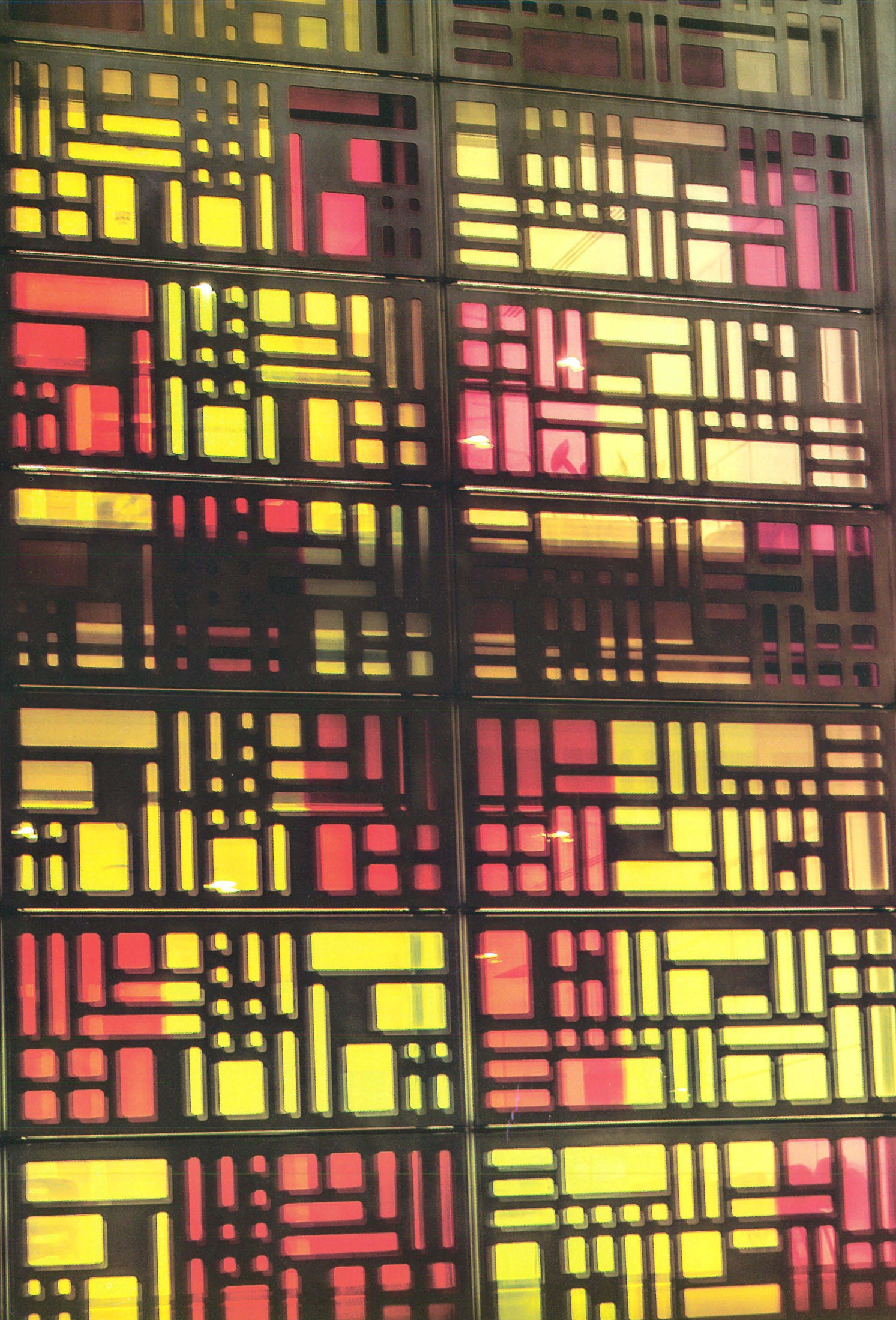
Blu5 Labs Limited - Ta' Xbiex - Malta
info@blu5labs.eu

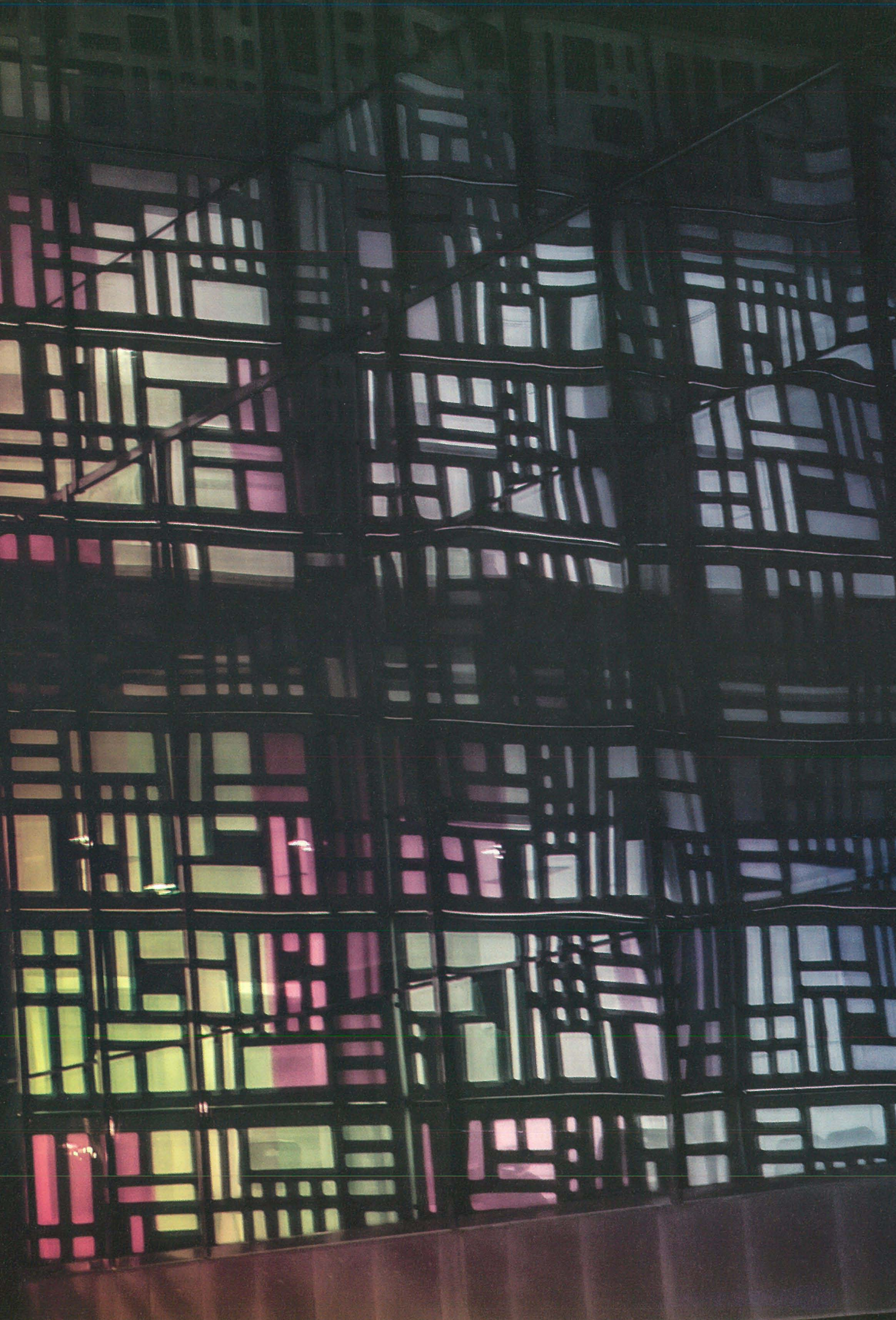
B.Sc. (Hons) | B.Sc. IT (Hons)

Final Year Project Exhibition 2016/7

ORGANISING COMMITTEE

Chairperson	Professor Ernest Cachia
Co-Chairperson	Ms Stephanie Abood
Members	Ms Anabel Abela Ing. Maria Abela Scicluna Mr Paul Bartolo Mr Rene' Barun Mr Anthony Buhagiar Mr Hose' Borg Ms Shirley Borg Mr Austin Camilleri Mr Patrick Catania Mr Keith Cauchi Mr Kevin Cortis Ing. Francarl Galea Ms Elaine Grech Mr Albert Sacco Mr Vincent Sammut Ms Francelle Scicluna Mr Raymond Vella Ms Ruth Vella Caruana Mr Mark Anthony Xuereb
Front Cover designed by	Mr Anthony Sultana
Photography	Rene Monseigneur
Printing	Printwell Ltd.





Contents

Message from the Dean

Dear Friends,

On behalf of all the staff at the Faculty of Information and Communication Technology (ICT), it is my pleasure to welcome you to the 17th Edition of the Annual Faculty of ICT Exhibition - FICTeX'17

As is invariably the case, we are proud to have come to that time of the year that celebrates, amongst other things, our final year student community's project accomplishment, as well as the outstanding academic achievement of students in our veracious degree programmes, this last through the Dean's List Awards. This year's FICTeX also marks 10 years since the establishment of our Faculty. It has been a challenging and yet interesting 10 years for the Faculty. 10 years in which the Faculty of ICT has gone from basically an idea based on innovative and forward-looking perspectives of people like Professor Juanito Camilleri, ex-Rector of the University, the developing national and international economic, social and professional landscapes, as well as the desire of academic staff to form an effective entity to participate in today and tomorrow's teaching and research needs and realities, to a Faculty that offers five well-established undergraduate degree programmes covering the whole spectrum of ICT servicing students from all over the University and several postgraduate degree programmes housing 115 postgraduate students, as well as a joint Master's degree with foreign European universities. From a faculty that no one had heard about to a faculty with a multitude of national, regional and international research collaborations on several levels, ever-growing numbers of exchange students, and 11 industry partners to whom we are always grateful for the support they offer us.

Notwithstanding the strides that this young faculty managed to make in this relatively short period, this event remains a time when we give due credit to various aspects of what our Final Year students have been able to accomplish in the relatively short duration of a few months, basing themselves on their undergraduate education and its application, and their own creativity, skills and capabilities. It is also a time where both the Faculty and our esteemed job-market partners show their recognition and appreciation of the effort that our Final Year students put into their Final Year Project through the award of various prizes during FICTeX.

As already mentioned, an important constituent of this event is The Faculty of ICT Dean's List Awards ceremony. The Dean's List Awards recognise truly outstanding effort within our undergraduate student community and we therefore deem it a worthy element to "exhibit" along with our students' Final Year Projects. I just wish to remind you all of the fact, that the criteria used by the Faculty to determine which students to include in the Dean's List are extremely stringent and demanding, resulting in a varying number of students eligible for this award each year. Therefore, I wholeheartedly congratulate this year's eight students who have made it into the list. A veritable well done!

Therefore, and I would like the coming statement to lie at the heart of this forward, on behalf of the Faculty of ICT, I would like to congratulate all our final-year students on their work, their achievement, and their dedication. I would also like to wish our out-going students success, fulfilment and happiness in their professional careers, whether they choose to pursue it in industry or continue deepening their studies in academia. These two career paths should never be exclusive of each other. We also very much augur, for those of our students who choose to take up their merited place in industry and commerce, that they do not irrevocably turn the page on their Faculty, and hopefully go on supporting their Faculty by placing themselves at the much needed, and at times unclear and unappreciated, boundary that interfaces academia with industry and that defines the area in which true innovation and social value can flourish. I would also like to congratulate all the academic staff in the Faculty for their contribution to science in the form of their research activities and bode that this will find its way into the teaching and supervision activities they engage in.

We also see this event as an ideal platform for people not necessarily involved in ICT to gain an insight into this far-reaching and pervasive field. This especially applies to the parents of the students directly involved in this exhibition. My warmest congratulations go out to them for the achievements of their children. Something indeed to be proud of.

Finally, I also wish to use this year's forward message to let lose, as it were, a few words with regards to the increasingly relevant issue of discernment between academic and vocational education. The first thing that one should note is that they are both forms of education. Rightfully so. While strongly believing that both forms of education are critically vital for any nation's prosperity, I also strongly believe that confusing the two or misplacing expectations from each can be equally detrimental. Academic education focuses on the acquisition of scientific knowledge and broad coverage of fundamental concepts applicable to all aspects of a specific domain of human endeavour. It cultivates discipline of thought to allow analytical processes to flourish for the benefit of understanding and progress. One would expect persons coming from such an educational background to exhibit traits of in-depth understanding leading to long-term growth and adaptability to prevailing process and technology, as well as a propensity towards trend-setting and forward vision. Academic education results in various degrees of qualification. Vocational education, on the other hand, focuses on the imparting of knowledge in the acquisition and use of skills and technologies. Vocational education is crucial in imparting of applicative and practical knowledge and skills pertaining to tangible environments and established procedure focused towards

a specific need or a specific technology. Persons opting for this type of education are invaluable practitioners and should be at the forefront of implementation. Vocational education results in various degrees of certification. A possible way to put it, as Tony Cowley from the University of Sheffield said way back in 1998, it boils down to "the knowing" as distinguished from "the doing". One without the other does not really make any sense – as does confusing the two.

When asked what defines our graduates, I like to reply - their ability to orientate themselves with relative ease in today's ICT industry, their ability to drive, rather than be driven by, technology, their ability to challenge the status quo, and their motivation to grow and strive for improvement in whichever context of human enterprise they might find themselves. This, we feel, is a national asset that must never be diluted or lost.

Allow me to thank those administrative persons in our Faculty who, through their effort, the organisation of all infrastructural and logistic aspects of this year's FICTeX, from this booklet you hold in your hands to the setup you will see in our various laboratories, would not have been possible. I would also like to thank all the excellent academic staff in all our five departments: Artificial Intelligence, Communications and Computer Engineering, Computer Information Systems, Computer Science, and Microelectronics and Nanoelectronics, for the invaluable expertise they lend in the tutoring of the students – whose culminating handiwork you can witness during this exhibition and for any other assistance they may have offered towards this event. A very big thank you is also due to the Faculty's hard-working technical, support and building staff for the sterling support they always wholeheartedly offer, sometimes even beyond the proverbial "call of duty".

I also wish to convey my thanks to both the Faculty of Engineering and the Faculty of Science, with whom we enjoy a close relationship and with whom we share a common history, for the productive atmosphere of collegiality and support we always find. I also wish to thank all the other University Faculties who work with us towards improving the quality of our collective services. I would very much wish to thank our industry partners who have seen the mutual benefits that can be gained by maintaining closer ties and discussion with the Faculty and who provide much needed material support in the form of service-for-sponsorship agreements.

Last but surely not least, I would like to thank all our University support offices for their help in handling any administrative and financial procedures, our Marketing, Communications & Alumni Office for providing exposure for this event, and especially the University Rectorate, through Professor Alfred Vella, for the constant support and trust that is shown towards the Faculty of ICT throughout its relatively short but exciting history. The future remains rife with dynamicity and promise.

Professor Ernest Cachia

Dean

Faculty of ICT

Acknowledgements

The Faculty of Information and Communication Technology gratefully acknowledges the following firms and organisations for supporting this year's exhibition:

- **Ascent Software**
- **Betsson Group**
- **Bit8**
- **Blu5labs**
- **CCBill**
- **Ernst &Young**
- **ICE Malta, Institute of Computer Education Limited**
- **IEEE Malta**
- **Information Systems Limited**
- **KPMG Crimsonwing**
- **Numencapital**
- **Parallels**
- **RakeTech**

OUR GOLD SPONSORS

Faculty of ICT
Members of Staff

DEPARTMENT OF COMMUNICATIONS AND COMPUTER ENGINEERING

Associate Professors	Professor Ing. Victor Buttigieg, B.Elec.Eng.(Hons.), M.Sc. (Manc.), Ph.D.(Manc.), M.I.E.E.E.
	Professor Ing. Carl J. Debono, B.Eng.(Hons.), Ph.D.(Pavia), M.I.E.E.E., M.I.E.E. (Head of Department)
	Professor Ing. Adrian Muscat, B.Eng. (Hons.), M.Sc. (Brad.), Ph.D.(Lond.), M.I.E.E.E.
	Professor Ing. Saviour Zammit, B.Elec.Eng.(Hons.),M.Sc.(Aston),Ph.D.(Aston), M.I.E.E.E. (Pro-Rector for Research and Innovation)
Senior Lecturers	Dr Johann A. Briffa, B.Eng. (Hons)(Melit.),M.Phil.(Melit.),Ph.D.(Oakland)
	Dr Ing. Reuben A. Farrugia, B.Eng.(Hons.),Ph.D.,M.I.E.E.E.
Lecturers	Dr Ing. Trevor Spiteri, B.Eng.(Hons.), M.Sc., Ph.D.(Bris.), M.I.E.E.E., M.I.E.T.
	Dr Gianluca Valentino, B.Sc.(Hons.)(Melit.), Ph.D. (Melit.),M.I.E.E.E.
Assistant Lecturer	Ing. Etienne-Victor Depasquale, B.Elec.Eng.(Hons.), M.Sc.(Eng.), M.I.E.E.E.
Visiting Assistant Lecturers	Ing. Brian E. Cauchi, B.Sc.IT (Hons.), M.Sc.
	Ing. Antoine Sciberras, B.Eng.(Hons.),PG.Dip.(Brunel)
	Ing. Leslie Spiteri, B.Elec.Eng.(Hons.), M.Sc., M.I.E.E.E.
	Ing. Martin Zammit, B.Elec. Eng. (Hons.)
Research Support Officers	Ms Leanne Attard, B.Eng.(Hons.), M.Sc.
	Ms Gabriella Azzopardi, B.Sc.(Hons.), M.Sc.
	Dr Ahmed Bannour, Ph.D.
Administrative & Technical Staff	Ms. Anabel Abela, (Administrative Assistant)
	Mr Albert Sacco, (Senior Laboratory Officer)
	Ing. Maria Abela-Scicluna, B.Eng.(Hons.)(Melit.) (Systems Engineer), M.Sc. ICT (Melit.)

Research Areas

Computer Networks and Telecommunications

- 5G Technology
- Adaptive and Intelligent Techniques in Wireless Systems
- Automation in Network Management
- Cognitive Radio Systems
- Error Correction Codes
- Image Coding for Novel Camera Architectures
- Internet of things
- Machine-to-Machine Communications
- Multimedia Communications
- Multi-view video coding and transmission
- Network Coding
- Software-Defined Networking
- Telecommunications and Network Modelling
- Video Coding
- Wireless Sensor Networks and Telematics

Signal processing and Pattern Recognition

- Automated Assessment in Intelligent Tutoring Systems
- Biometrics
- Computer Vision
- Gesture recognition
- Image Processing
- Integrated Vision and Language
- Medical Image Processing and Coding
- Multimedia Retrieval and Indexing
- Multimedia Security and Forensics
- Shape Analysis and Understanding

Computer Systems Engineering

- Digital Games Platforms
- Human Machine Interfaces
- Implementation on Massively Parallel Systems (e.g. GPUs)
- Reconfigurable Hardware

DEPARTMENT OF COMPUTER SCIENCE

Professor

Professor Gordon J. Pace,
B.Sc., M.Sc. (Oxon.), D.Phil. (Oxon.)

Senior Lecturers

Dr Adrian Francalanza,
B.Sc.I.T. (Hons.), M.Sc., D.Phil.(Sussex)

Dr Kevin Vella,
B.Sc., Ph.D. (Kent)
(Head of Department)

Lecturers

Dr Keith Bugeja,
B.A.(Hons), M.IT, Ph.D.(Warw.)

Dr Christian Colombo,
B.Sc.I.T. (Hons.), M.Sc. Ph.D. (Melit.)

Dr Joshua Ellul,
B.Sc.I.T. (Hons.), M.Sc. (Kent) , Ph.D. (Soton)

Dr Mark Micallef,
B.Sc.(Hons.), Ph.D. (Melit.)

Dr Sandro Spina,
B.Sc.I.T.(Hons), M.Sc. (Melit), Ph.D.(Warw.)

Dr Mark J. Vella,
B.Sc.I.T.(Hons.), M.Sc. Ph.D. (Strath.)

Research Support Officers

Mr Luke Chircop,
B.Sc (Hons.) (Melit.), M.Sc (Research Support Officer I)

Mr Shaun Paul Azzopardi,
B.Sc (Hons.) (Melit.), M.Sc (Research Support Officer I)

Administrative Staff

Mr Kevin Cortis,
B.A.(Hons) Graphic Design & Interactive Media
(Clerk)

Research Areas

- Concurrency
- Computer Graphics
- Compensations
- Compilers
- Model Checking and Hardware/Software Verification
- Operating Systems
- Semantics of Programming Languages
- High Performance Computing and Grid Computing
- Runtime Verification
- Software Development Process Improvement and Agile Processes
- Software Engineering
- Software Testing
- Security

DEPARTMENT OF MICROELECTRONICS AND NANOELECTRONICS

Professor	Professor Ing. Joseph Micallef, B.Sc.(Eng.)(Hons.),M.Sc.(Sur.),Ph.D.(Sur.), M.I.E.E.E.
Associate Professor	Professor Ing. Edward Gatt, B.Eng.(Hons.),M.Phil.,Ph.D.(Sur.),M.I.E.E.E. (Head of Department)
	Professor Ivan Grech, B.Eng.(Hons.),M.Sc.,Ph.D.(Sur.),M.I.E.E.E.
Senior Lecturers	Dr Ing. Owen Casha, B. Eng.(Hons.) (Melit.), Ph.D. (Melit.), M.I.E.E.E. (Faculty Representative on Senate)
	Dr Ing. Nicholas Sammut, B.Eng. (Hons.) (Melit.), M.Ent. (Melit.), Ph.D. (Melit.), M.I.E.E.E. (Deputy Dean)
Research Support Officers	Mr Russell Farrugia, B.Eng. (Hons)(Melit.), M.Sc.(Melit.) (Research Support Officer II)
	Mr Barnaby Portelli, B.Eng. (Hons)(Melit.), M.Sc.(Melit.) (Research Support Officer II)
Administrative & Technical Staff	Mr Keith Cauchi, B.A. (Hons.) Youth & Community Studies (Melit.) (Executive Officer)
	Ing. Francarl Galea, B.Eng. (Hons.),M.Sc.(Eng.) (Systems Engineer)

Research Areas

- Analogue and Mixed Mode ASIC Design
- RF CMOS Circuits
- Embedded Systems
- Biotechnology Chips
- Micro-Electro-Mechanical Systems (MEMS)
- Quantum Nanostructures
- System-in-Package (SiP)
- System-on-Chip (SoC)
- Accelerator Technology

DEPARTMENT OF ARTIFICIAL INTELLIGENCE

Associate Professors	Professor Alexiei Dingli, B.Sc.I.T. (Hons.) (Melit.), Ph.D. (Sheffield),M.B.A (Grenoble) (Head of Department)
	Professor Matthew Montebello, B.Ed.(Hons.), M.Sc.,M.A.(Ulster),Ph.D.(Cardiff)
Senior Lecturers	Dr Christopher Staff, B.A.(Hons.)(Sussex), D.Phil. (Sussex)
Visiting Senior Lecturer	Mr Michael Rosner, M.A. (Oxon.),Dip.Comp.Sci.(Cantab.)

Lecturers	Dr Charlie Abela, B.Sc. I.T. (Hons)(Melit.), M.Sc. (Comp.Sci.)(Melit.),Ph.D.(Melit)
	Dr Joel Azzopardi, B.Sc. (Hons.) (Melit.), Ph.D. (Melit.)
	Dr George Azzopardi, B.Sc. I.T. (Hons.)(Lond.), M.Sc. (Lond.), Ph.D. (Groningen)
	Dr Claudia Borg , B.Sc. I.T. (Hons.) (Melit), M.Sc. (Melit.), Ph.D. (Melit)
	Dr Vanessa Camilleri, B.Ed. (Hons.) M.IT, Ph.D. (Cov)
Assistant Lecturer	Mr Kristian Guillaumier, B.Sc. I.T. (Hons.), M.Sc.
Administrative Staff	Ms Francelle Scicluna, (Administrative Assistant) B.A. W.H.R (Hons.) (Melit.)

Research Areas

- Title: Language Technology for Intelligent Document Archive Management
  Area: Linked and open data
- Title: Crowdsourcing in Education
  Area: ICT in Education
- Title: Medical image analysis and Brain-inspired computer vision
  Area: Intelligent Image Processing
- Title: MyOcean Follow-On, MEDESS4MS, and Calypso 2 projects
  Area: Down-stream services
- Title: GBL4ESL
  Task: Creation of digital resources for educators using a Game Based Learning Toolkit
- Title: Smart animal breeding with advanced machine learning techniques
  Area: Predictive analysis, automatic determination of important features
- Title: Real-time face analysis in the wild
  Area: Computer vision
- Title: Integrated Vision and Language Processing Models
  Area: Computer Vision/NLP
- Title: Maltese Language Resource Server (MLRS)
  Area: Natural Language Processing
  Task: Research and creation of language processing tools for Maltese

An updated list of concrete areas in which we have expertise to share/offer

- AI, Machine Learning, Adaptive Hypertext and Personalisation
- Pattern Recognition and Image Processing
- Web Science, Big Data, Information Retrieval & Extraction, IoT
- Agent Technology and Ambient Intelligence
- Natural Language Processing/Human Language Technology
- Document Clustering and Scientific Data Handling and Analysis
- Intelligent Interfaces, Mobile Technologies and Game AI
- Optimization Algorithms

Areas which we are interested in starting/rekindling

Area	Health
Interest	Computer aided diagnosis for the detection of diabetic retinopathy from retinal images
Area	Health / Internet of Things
Interest	Automated decision-support for personalised self-management to prevent recurrence of low-back pain
Area	Security
Interest	Forensic analysis based on video surveillance cameras
Area	Big Data
Interest	Identification of patterns using multi-modal data and all kinds of media
Area	Banking
Interest	Detection of abnormal credit-card-based spending behaviour
Area	Brain-inspired vision
Interests	Understanding the role of feedback circuits in the visual cortex and use it to design more effective computer vision algorithms
Area	Human Robot Interaction
Interest	Multi-modal communication
Area	Digital Archive Management
Interest	Information Extraction (Text Mining) from Digital Archives
Area	Higher Education
Interest	Applying ICT to enhance e-learning

DEPARTMENT OF COMPUTER INFORMATION SYSTEMS

Associate Professor

Professor Ernest Cachia,
M.Sc.(Kiev), Ph.D.(Sheff.)
(Head of Department /Dean)

Senior Lecturers

Dr John Abela,
B.Sc.(Hons.), M.Sc., Ph.D.(New Brunswick), I.E.E.E., A.C.M.

Dr Lalit Garg,
B.Eng.(Barkt), PG Dip. I.T.(IIITM), Ph.D.(Ulster)

Dr Colin Layfield,
B.Sc. (Calgary), M.Sc.(Calgary), Ph.D.(Leeds)

Visiting Senior Lecturer

Dr Vitezslav Nezval,
M.Sc.(V.U.T.Brno),Ph.D.(V.A.Brno)

Mr Rodney Naudi,
B.Sc., M.Sc.(Eng.)(Sheff.)

Lecturers

Dr Peter A. Xuereb,
B.Sc.(Eng.)(Hons)(Lond.), ACGI,M.Phil.(Cantab.), Ph.D.(Cantab.)

Dr Joseph Vella,
B.Sc., Ph.D.(Sheffield)

Dr Christopher Porter,
B.Sc.(Bus.&Comp.), M.Sc. , A.C.M., Ph.D.(UCL)

Dr Conrad Attard,
B.Sc.(Bus.&Comp.), M.Sc., Ph.D.(Sheffield)

Visiting Assistant Lecturers

Ing. Saviour Baldacchino,
B.Elec.Eng.(Hons.), Dip.Mgt., A.C.M., I.E.T.

Mr Norman Cutajar,
M.Sc. Systems Engineering

Assitant Lecturers	**Mr Michel Camilleri,** **B.Sc., M.Sc., Dip.Math.&Comp.**
	Mr Clyde Meli, **B.Sc., M.Phil.**
	Mr Joseph Bonello, B.Sc.(Hons)IT(Melit.), M.ICT(Melit.)
Associate Academic	**Mr Anthony Spiteri Staines,** **B.Sc., M.Sc., A.I.M.I.S., M.B.C.S.**
Administrative Staff	**Ms Shirley Borg, (Administrative Assistant)**

Research Areas

- Object Oriented Platforms, Languages and Techniques in Distributed Environments
- System Development including Real-Time scheduling, stochastic modelling, and Petri Nets
- Modern Software Engineering (based on Conceptual Modelling and Agile development)
- Database Management Systems, Data Modelling including Spatial-temporal Modelling
- Data Mining and Data Warehousing
- Software Project Management
- IT Strategic Management including E-strategy
- Services and Security (including Electronic Identities and Spam Detection)
- Quality Assurance and Risk Management of IT Frameworks
- Applicative Genetic Algorithms and Genetic Programming
- 3D Graphics Modelling Technologies
- Mobile Computing and Technologies
- IT Psychology and Semantic Technologies (Web and Applications)
- Cloud Computing Solutions and Technologies
- HCI, machine learning, latent semantic analysis, scheduling, timetabling, optimisation
- Business Intelligence
- Information Systems and Business Applications Development
- Enterprise Resource Planning
- Bioinformatics
- IT Audit and IT Forensics
- Health & Social Care Modelling
- Missing Data Analysis
- Biomedical Informatics
- Traffic analysis and sustainable transportation

FACULTY OFFICE

Ms Stephanie Abood, Dip.Soc.Stud.(Industrial Relations) (Melit.) (Administrative Officer/ Faculty Officer)

Mr Rene' Barun, BA (Hons.) Philosophy (Melit), (Clerk)

Mr Hose' Borg, (Clerk)

Mr Anthony Buhagiar, (Administrative Officer)

Ms Ruth Vella Caruana, Dip.Marketing (Melit) (Administrative Assistant)

Mr Vincent Sammut, Dip.Crim.(Melit.) (Administrative Assistant)

Ms Elaine Grech, H.Dip.Administration & Management (Melit.) (Executive Officer)

SUPPORT STAFF

Mr Patrick Catania A.I.M.I.S. (IT Officer III)

Mr Paul Bartolo (Senior Beadle)

Mr Austin Camilleri (Beadle)

Mr Raymond Vella (Technical Officer II)

PROJECT NO. 39

Investigating the use of the Map-Reduce Paradigm for Genome Alignment

SARA ANN ABDILLA

SUPERVISED BY
DR JEAN-PAUL EBEJER

Genome alignment has many critical applications in multiple areas, from medical diagnosis to forensic investigation [1]. Genomes, particularly human ones, consist of a large amount of data – approximately 3 billion base-pairs (or 3 GB); each base being one of four possible nucleotides: adenine, cytosine, guanine and thymine, often referred to by the letters A, C, G and T respectively [2]. Due to limits in DNA sequencing technologies, the whole genome cannot currently be read entirely at one go. These technologies instead output millions to billions of short DNA reads (fragments of the genome) [3] from which the whole genome is reconstructed through software. All of these reads are then aligned to the reference genome as depicted in Figure 1. This alignment is not a perfect exact match and may contain differences due to sequencing errors, the natural variation between species [2] and the differences between DNA sequences of individuals.

The motivation for this project stems from the Malta Human Genome Project (MHGP) where researchers at the University of Malta aim to sequence 1% of the Maltese population. This project requires analysis tools for its alignment technology. Our aim is to investigate the best strategy for the alignment of reads (from a sample) to the reference genome using the distributed Map-Reduce paradigm to run this alignment in the cloud to attempt to achieve speed up over current methods.

This project was designed around the investigation and evaluation of two genome compressors – a 32-bit and a 64-bit encoder – and seven Map-Reduce enabled read-aligners – naive matching using the Hamming distance metric, naive matching using the Edit distance metric, Boyer-Moore, *k*-mer index, Smith-Waterman, FM index and Burrows-Wheeler. The implementation was developed using the Python programming language. Also, the *Apache Hadoop* Map-Reduce framework was used for local and cloud executions (Amazon Web Services being selected as the cloud computing infrastructure provider).

The results and evaluation showed that the 32-bit encoder performed similarly to the *Gzip* compression tool. The latter is however slightly faster, so it was selected to be used by the Map-Reduce aligners. The results and evaluation of the aligners showed that most produced accurate alignments and were executed in a reasonable amount of time – the *k*-mer index aligner clearly executing the fastest. Evaluating these against the existing, popular alignment tool *BWA* showed that, in terms of accuracy, all the aligners return similar results; our aligners providing a more readable and concise output. In terms of run-time, *BWA* outperforms most of the Map-Reduce aligners we implemented with the exception of the *k*-mer index aligner which showed comparable timings. However, *BWA* is limited to execute on a multi-threaded single node while our approach is more scalable as it allows execution on multiple nodes.

Future work on this area may involve optimisations for the 32-bit encoder and for the *k*-mer index aligner (for instance, by increasing the cluster size to outperform *BWA*), the use of quality scores to measure the sequencing accuracy of the resultant alignments and the construction of more analysis tools for the MHGP such as genome visualisation tools and genome browsers.

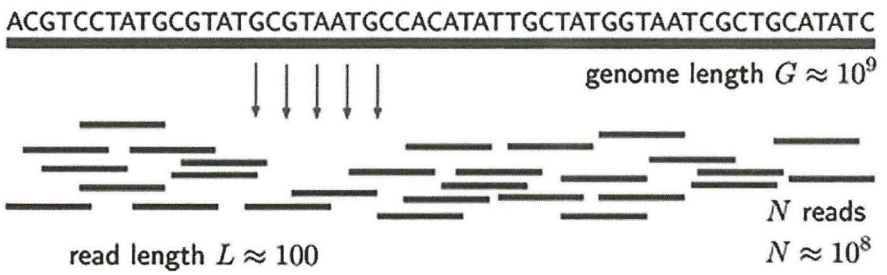


Figure 1: Illustration of DNA sequencing reads against a genome [4] – Nowadays it is common to find a larger number of reads N due to improvements in sequencing technologies.

REFERENCES

[1] The Hidden History of the Maltese Genome. Think Magazine, April 2016.

[2] Arthur M. Lesk. Introduction to Genomics. Oxford University Press, Great Claredon Street, Oxford, OX2 6DP, UK, 2 edition, 2012.

[3] David W. Mount. Bioinformatics: Sequence and Genome Analysis. Cold Spring Harbor Laboratory Press, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA, 2001.

[4] Abolfazl Motahari, Guy Bresler, and David Tse. Information Theory of DNA Shotgun Sequencing. Technical report, October 2013.

PROJECT NO. 9

Interactive High-fidelity Rendering in the Cloud

ANDREW ABELA

SUPERVISED BY
DR KEITH BUGEJA AND DR SANDRO SPINA

High-fidelity rendering requires a substantial amount of computational power to render even moderately complex scenes. This is exacerbated by the need for interactivity, which can only be achieved on powerful high-end systems or computer clusters. Cloud systems provide elasticity of computational systems and a seemingly infinite number of resources that can be made available to a client device. This work investigates the use of the cloud in providing interactive high-fidelity rendering to client devices irrespective of their computational power. A novel framework for rendering as a service is presented and evaluated.

Figure 1 shows the architecture of the proposed framework. The client-server approach was adopted; the client requests to render a particular scene while the server is responsible for handling the client parameters and render the scene accordingly. The figure also illustrates the different components of the proposed system: the façade serves as an entry point into the system, the resource manager keeps a record of all allocations and deallocations amongst the various computation resources, the resource manager daemon (RMD) is a delegate of the resource manager and responsible for executing the commands issued by the resource manager, the renderer (R) partitions the frame into a number of tiles and gathers the results from the workers and the worker (W) renders a particular tile and sends the results to the renderer.

Results show that the framework is scalable and elastic. The proposed system achieved quasi-linear speed-up as the number of workers increased and the experiments demonstrate an upper-bound of 30 seconds to the elastic response of the system. Furthermore, the proposed framework incurred minimal realisation penalties; it has been shown to sustain rendering of path-traced images at 9 frames per second on a commodity hardware infrastructure.

The framework also supports scheduling policies for resource management amongst multiple connected clients; the policies are based on different job characteristics, from interactive jobs to batch-based offline rendering. The implemented policies were inspired by operating systems as resource managers which, for example, schedule processes amongst the available CPUs.

Results show that shortest job first scheduling policy minimises average waiting time, while the priority scheduling policy modulates job waiting time by the assigned priority. As expected, round robin scheduling minimises response time for each individual job, while priority-based round robin shows improved response times for high-priority jobs.

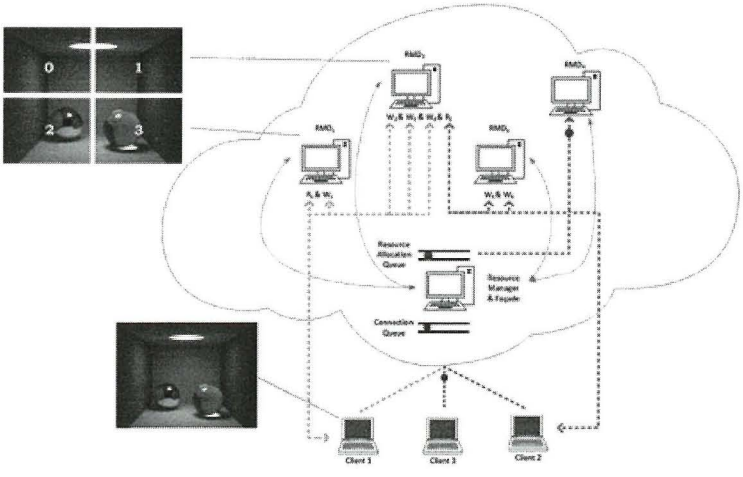


Figure 1: The image shows 3 clients, all of which sent their rendering request.

PROJECT NO. 14

Machine Learning Methods for Handling Missing Data

DYLAN LUKE ABELA

SUPERVISED BY
DR LALIT GARG

WHAT IS MISSING DATA?

Missing data is that type of data that was left empty during a situation. Some examples demonstrating this are:

- A questionnaire was given to someone and one or more questions were left empty
- When a person has to enter his/her name or any personal data and he/she keeps this empty because of reasons

Unfortunately these missing data create a lot of problems for researchers.

MULTIPLE IMPUTATION (MI)

This method works by creating multiple copies of the dataset and then, the missing data is replaced by valid elements chosen at random. It will then, calculate the mean of each of the datasets and it will use the following formula [1]:

where **M** s $\sqrt{\frac{1}{M}\sum_{k=1}^{M} s_k^2 + \left(1+\frac{1}{M}\right)\left(\frac{1}{M-1}\right)\sum_{k=1}^{M}(a_k-\bar{a})^2}$ ıe standard error [2] and **a** is the parameter estimate [3] [4]. Then, after an answer is obtained from the formula, it will replace all the missing data in the dataset under test.

DIRECT MAXIMUM LIKELIHOOD (DML)

This method uses the following equation [5]:

where **y** is the number of occurrences, **g** is the probability of that element to occur and **total** is the total number of

$$\Pr(x = y \mid p = g) = \binom{\text{Total}}{x} (g)^{y} (1 - g)^{\text{total}-y}$$

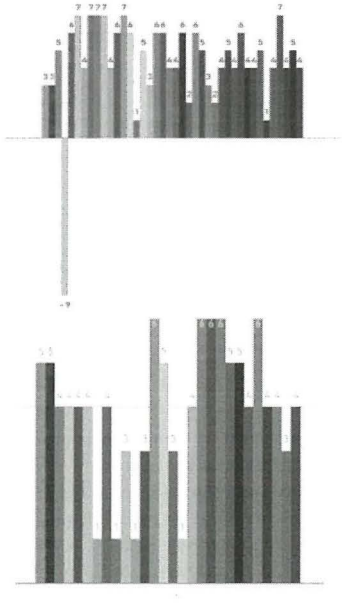
occurrences in the dataset. Then after calculating for each and every occurrence, the maximum number among all the answers will be picked and its respective element will replace all the missing data that occur in the dataset.

OBSERVED DATA BECOMING MISSING A RANDOM (MAR)

Let's say that a student forget to write his/her name on the examination paper. That creates quite a problem. So researchers will analyse the observed data and try to figure out that missing data and do their best to replace it with the correct observed data. This can be done by having multiple extra copies that can be used for double checking. Taking the example of the student forgetting to write his/her name on the paper, teachers can check previous work of the students and try to match the handwriting.

GRAPHS ABOUT THE DATASETS

The graph on **top** shows what a particular person answered in each given question. There is a block in the graph going to the negative side of the y-axis. This person has some missing data and it is easily spotted. The graph at the **bottom** shows the same concept but this time, there is not missing data found, meaning this person did not leave any questions unanswered. Each graph is colour coded to represent each question.



REFERENCES

[1] Allison, P.D., 2012, April. Handling missing data by maximum likelihood. In SAS global forum (Vol. 23).

[2] http://ncalculators.com/math-worksheets/calculate-standard-error.htm

[3] J. V. Beck and K. J. Arnold, *Parameter Estimation in Engineering and Science*. James Beck, 1977.

[4] https://www.mathworks.com/discovery/parameter-estimation.html

[5] N. M. Kiefer, "Discrete parameter variation: Efficient estimation of a switching regression model," *Econometrica: Journal of the Econometric Society*, pp. 427-434, 1978.

Automatic Symbol Localisation and Recognition from Music Sheets

JULIAN ANTHONY ABELA

SUPERVISED BY
DR GEORGE AZZOPARDI AND DR ALEXANDRA BONNICI

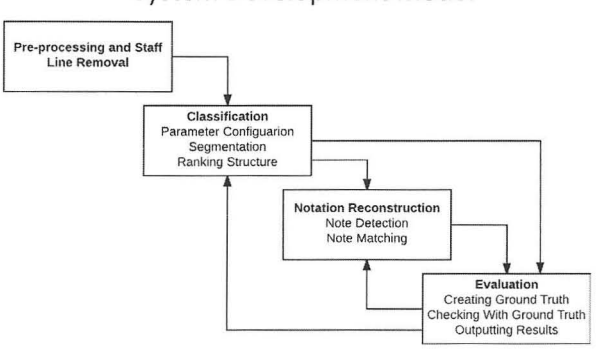
Optical music recognition is the computerised process of reading images of music sheets and transforming them into a digital form, the result of which has many useful applications in the music industry [1, 2]. The focus of this project was on the classification phase of this process, specifically on the recognition of symbols which are lacking in previous research.

The first step towards achieving this was the pre-processing of the images with the use of adaptive binarisation [3] and staff line removal [4]. Next, the classification of music symbols was implemented using two different methods, the COSFIRE approach [5] and template matching [6], to determine which of them are better suited for this application. These image recognition methods were trained on two different fonts, one of which was the font found in the images, in order to evaluate how well these methods are able to specialise and generalise. The recognition of each symbol was also configured differently according to their different shapes and sizes. The segmentation phase was subsequently implemented, which split up the various symbols from the rest of the image using segmentation of connected components [7].

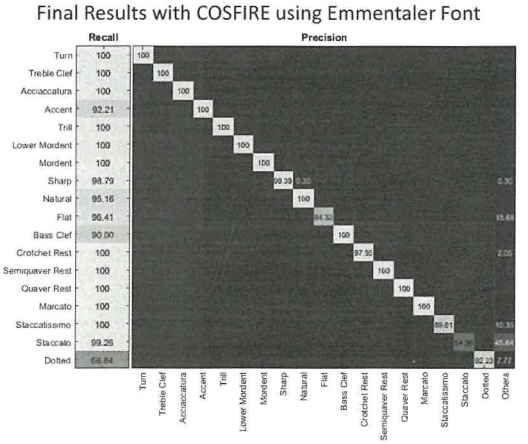
A ranking structure which represents the order in which the symbols are detected was established, with the aim of minimising the amount of false positives between the symbols themselves. The last step consisted of the matching of the symbols with their respective notes, not only to improve the accuracy, but to also further the notation process of optical music recognition.

Both image recognition methods obtained high results on almost all the symbols recognised, with many of them obtaining close to or even exactly 100\% precision and recall. It was concluded that the COSFIRE approach performs better than the template matching in most cases, both when using the same or a different font than found on the music sheets. The two techniques which were applied, the ranking structure and note matching, were also successful in improving the recognition results by a noticeable margin, ranging between a 1\% to 10\% increase in the average precision each time.

System Development Model



Final Results with COSFIRE using Emmentaler Font



REFERENCES

[1] V. Charvillat, A. Tonazzini, L. Van Gool and N. Nikolaidis, "Image and Video Processing for Cultural Heritage", *EURASIP Journal on Image and Video Processing*, vol. 2009, pp. 1-3, 2009.

[2] E. Pollastri and G. Simoncelli, "Classification of melodies by composer with hidden Markov models", *Web Delivering of Music, 2001. Proceedings. First International Conference*, pp. 88-95, 2001.

[3] J. Sauvola and M. Pietikäinen, "Adaptive document image binarization", *Pattern Recognition*, vol. 33, no. 2, pp. 225-236, 2000.

[4] A. Dutta, U. Pal, A. Fornes and J. Llados, "An efficient staff removal approach from printed musical documents", *Pattern Recognition (ICPR), 2010 20th International Conference*, pp. 1965-1968, 2010.

[5] G. Azzopardi and N. Petkov, "Trainable COSFIRE Filters for Keypoint Detection and Pattern Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 2, pp. 490-503, 2013.

[6] J.P. Lewis, "Fast template matching", *Vision interface*, vol. 95, no. 120123, pp. 15-19, 1995.

[7] L. Shapiro and R. Haralick, "Computer and robot vision", *Reading: Addison-Wesley*, vol. 8, 1992.

PROJECT NO. 63

PenTaGon: Personal Task Data Generator

JULIAN GEORGE AGIUS

SUPERVISED BY
DR CHARLIE ABELA AND DR CHRIS STAFF

Personal Information Management (PIM) is the area of research concerned with understanding how people acquire, maintain, retrieve and use information items [1]. In order to develop tools which aid users to manage their information more effectively, PIM researchers require large amounts of high quality user data [2]. Collecting user data is a very time consuming process and also raises privacy and confidentiality concerns [3]. To overcome these problems we propose PenTaGon, a personal task data generator which provides synthetic web browsing session data. PenTaGon utilises a variety of web search behaviour patterns [4], [5] and strategies [6] based on the web search experience and the domain-related knowledge of the user it attempts to mimic. The prototype requires a user specified information seeking task as input. Each task has a name and is divided into sub-tasks [4], [7]. Each sub-task is considered as an independent query session [5] consisting of a number of user-defined queries. Figure 1 summarises the processes carried out by PenTaGon. PenTaGon focuses on the latter stages of the information retrieval process described in Figure 2, the process of finding and interacting with web documents in a realistic manner. The system uses a web user behaviour model develop by Baeza-Yates et al. [8] to select a number of results from the search results page after submitting a query to the Google search engine. The level of web and domain experience of the artificial user will determine the patterns used to access and interact with the selected web documents. The system represents browsing sessions as graphs, where nodes symbolise a web page and links symbolise a switch between one web page and another. Preliminary findings show that the synthetic browsing sessions of an artificial user who is considered to be a web and domain expert are significantly more complex than those of a web and domain novice. Furthermore, the degree distributions of the browsing sessions follow the power law distribution, reinforcing the results achieved during this initial attempt of this type of web task data generator.

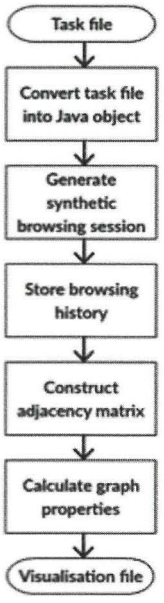


Figure 1: High-level overview of PenTaGon

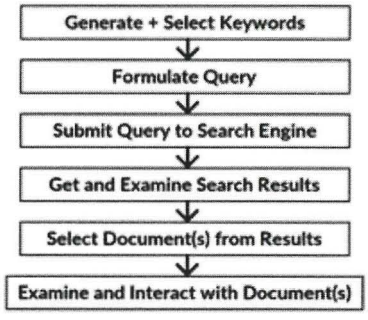


Figure 2: The information retrieval process as described by [4]

REFERENCES

[1] W. Jones, "Personal Information Management," *Annu. Rev. Inf. Sci. Technol.*, vol. 41, no. 1, pp. 453–504, 2007.

[2] D. Kelly, "Evaluating personal information management behaviors and tools," *Commun. ACM*, vol. 49, no. 1, pp. 84–86, 2006.

[3] P. Christen and A. Pudjijono, "Accurate synthetic generation of realistic personal information," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 5476 LNAI, pp. 507–514, 2009.

[4] C. Holscher and G. Strube, "Web search behavior of internet experts and newbies," *Comput. Networks*, vol. 33, no. 1, pp. 337–346, 2000.

[5] C. J. and C. L. C. and S. Wiedenbeck, "Patterns of information seeking on the Web: A qualitative Study of domain expertise and Web expertise," *IT Soc.*, vol. 1, pp. 64–89, 2003.

[6] A. Thatcher, "Web search strategies: The influence of Web experience and task type," *Inf. Process. Manag.*, vol. 44, no. 3, pp. 1308–1329, 2008.

[7] T. Joachims, L. Granka, B. Pan, H. Hembrooke, F. Radlinski, and G. Gay, "Evaluating the accuracy of implicit feedback from clicks and query reformulations in Web search," *ACM Trans. Inf. Syst.*, vol. 25, no. 2, p. 7–es, 2007.

[8] R. Baeza-Yates, C. Hurtado, M. Mendoza, and G. Dupret, "Modeling user search behavior," *Web Congr. 2005. LA-WEB 2005. Third Lat. Am.*, p. 10 pp., 2005.

PROJECT NO. 64

Biometric Analysis for Personal Identification Using Retinal Fundus Images

ADRIAN APAP

SUPERVISED BY
DR GEORGE AZZOPARDI

In 1935, Dr Carleton Simon and Dr Isodore Goldstein [1], proposed a new method for biometric identification using the blood vessel structure of the retina. However, this method was not researched further until recent years because of the technological limitations in manufacturing fundus cameras. Fundus cameras are flash enabled cameras with a microscope attached to them [2], which are used to obtain retinal images, Figure 1.

In this work, we develop a personal identification system based on the COSFIRE filter algorithm [3]. Figure 2, shows the two main stages of the implemented COSFIRE based system. Where the first stage consists of enrolling the identities in a database by creating a unique pattern, using features extracted from the retinal image. The second stage consists of the identification process. In this stage, the system attempts to return the identity of a retinal image using the enrolled patterns created during the enrolment stage. We also implement a similar, personal identification system based on the SIFT algorithm [4]. We implement the second system to compare the results obtained by the COSFIRE based system.

To evaluate our system, we use images obtained from the DRIVE [5] and STARE [6] datasets, which are both freely available online. One of the main problems when using retinal images for personal identification, are the transformations that occur in the retinal images due to head or eye movement. To address this problem, we create several datasets from the DRIVE [5] and STARE [6] datasets by applying multiple combinations of rotations and Gaussian noise to the retinal images.

Using the created datasets, we evaluated both systems. Where the COSFIRE based system, was able to produce a higher accuracy score than the SIFT based system in all the experiments that were performed. The COSFIRE based system produced a **100%** accuracy score when evaluated using the dataset of noisy images, and the dataset of rotated images. When using the datasets containing combinations of rotations and noise the COSFIRE based system produced an average accuracy score of **91.31%**. From our experiments we also showed that unlike the SIFT based system, the COSFIRE based system was successfully able to extract the retinal bifurcations from the noise and textured background.

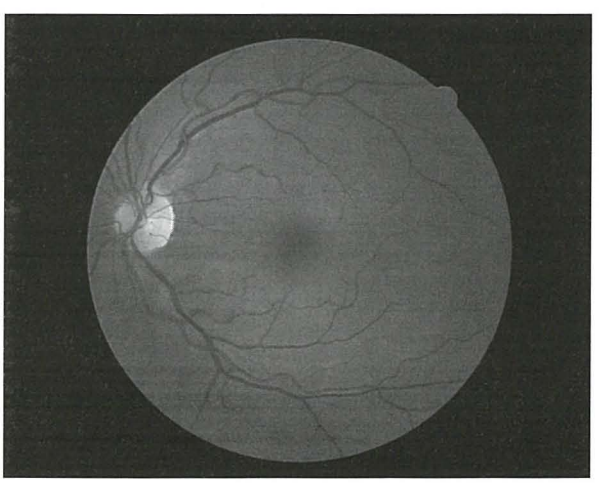


Figure 1: Retinal fundus image [5].

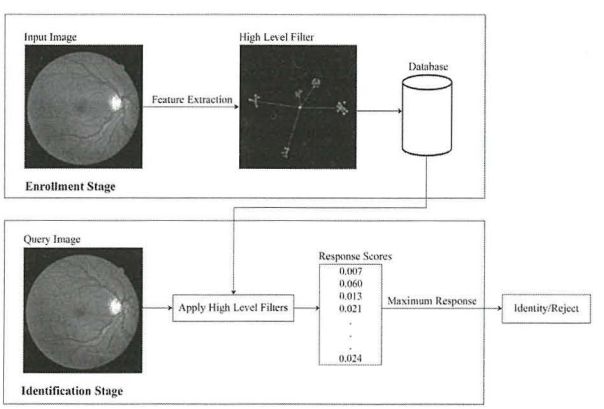


Figure 2: COSFIRE based implementation

REFERENCES

[1] C. Simon and I. Goldstein, "A new scientific method of identification," New York state journal of medicine, vol. 35, no. 18, pp. 901-906, 1935.

[2] R. Taylor, L. Lovelock, W. Tunbridge, K. Alberti, R. G. Brackenridge, P. Stephenson and E. Young, "Comparison of non-mydriatic retinal photography with ophthalmoscopy in 2159 patients: mobile retinal camera study.," Bmj, vol. 301, no. 6763, pp. 1243-1247, 1990.

[3] G. Azzopardi and N. Petkov, "Trainable COSFIRE filters for keypoint detection and pattern recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 2, pp. 490-503, 2013.

[4] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," International journal of computer vision, vol. 60, no. 2, pp. 91-110, 2004.

[5] "DRIVE: Digital Retinal Images for Vessel Extraction," 2016. [Online]. Available: http://www.isi.uu.nl/Research/Databases/DRIVE/. [Accessed 2016].

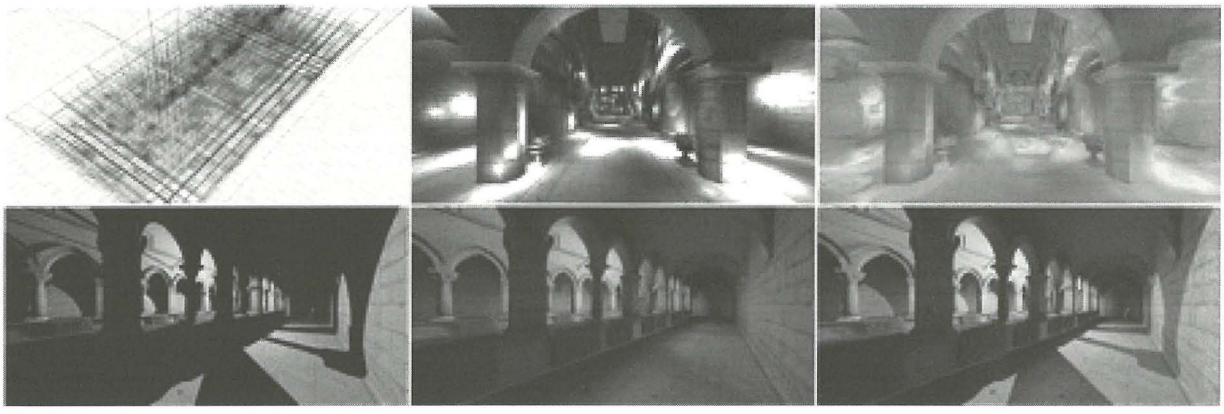
[6] "The STARE (STructured Analysis of the Retina) Project," 2016. [Online]. Available: http://cecas.clemson.edu/~ahoover/stare/. [Accessed 2016].

PROJECT NO.
40

Efficient Rendering Of Shadow-Casting For Multiple Dynamic Point Light Sources

JEAN KARL APAP

SUPERVISED BY
DR SANDRO SPINA AND DR KEITH BUGEJA



Rendering is the process of synthesising an image from a set of models by drawing and colouring them. The inclusion of shadows in this image, although significantly contributing towards the computational costs of rendering, greatly increases visual quality, immersion and aids in the discrimination of depth, helping users' understanding of a scene [LB00]. A straightforward but expensive approach is that of computing the shadow contributions for each of these light sources. To reduce rendering costs, clustering of these light sources can be used to create new light sources, representing groups of light sources in the scene, that generate perceptually identical output. In real-time rendering, where a user navigates a scene interactively, the perceptual contribution of a light source to the output frame may vary significantly over time; a clustering algorithm should factor this in, together with the limited computational budget constraints that come with interactive rendering.This dissertation presents an interactive light prioritisation and clustering technique (SLIC), based on the Barnes-Hut [BH86] n-body algorithm, which determines and renders shadows that are deemed to be mostly relevant from an observer's point of view according to a given computational budget. A Global Illumination [Dor95] method is also presented which implements realistic lighting in a scene and which is compatible with SLIC. Light probability and colour maps are used to generate light sources in a scene according to a probability distribution and given a certain colour. A city map is provided to generate a simple city. The clustering algorithm can be easily configured to work with the presented maps to provide an improvement in performance.

REFERENCES

[BH86] Josh Barnes and Piet Hut. A hierarchical o (n log n) force-calculation algorithm. 1986.

[Dor95] Julie Dorsey. Radiosity and global illumination. *The Visual Computer*, 11(7):397-398, 1995.

[LB00] Michael S Langer and Heinrich H Bülthoff. Depth discrimination from shading under diffuse lighting. *Perception*, 2000.

PROJECT NO. 10

Crowdsourcing Information on Points of Interest

RYAN AQUILINA

SUPERVISED BY
DR CONRAD ATTARD

Crowdsourcing has come a long way from the time Jeff Howe first defined it in 2006 [1]. The recent growth in smartphone users has extended the reach of crowdsourcing to location based content [2]–[4]. This project aims to leverage the crowd's wisdom by gathering information on points of interest. This FYP will take advantage of smartphone sensors to create a tool that uses the GPS receiver to retrieve location, photographic data and textual input from participants to gather information on points of interest, this has been designed on par with participatory sensing applications found in literature [5]–[8]. The system's data sets have been used to train four classification models to classify the reports produced by the participants. The Naive Bayes classifier was found to be the most suitable scoring an accuracy of **95.4%**. Two controlled experiments have been conducted in a laborotary and on site (Valletta) with 30 participants to evaluate the prototype's usability and effectiveness in gathering such data. The researcher conducted the experiment with three individuals at a time; this has been designed so that particpants can suggest answers to factual questions using the suggestion system implemented in the crowdsourcing tool. A total of 90 reports and 1710 data items were produced by the participants, scoring an average validity score of **79.87%** on factual data sets. The score was calculated by processing the reports under two quality control methods found in literature specifically **Majority Rule** and **Expert Review** methods [9]. The validity score was calculated by assigning one point per valid answer in a report. Using the chi-square test and Kruskal-Wallis test on these scores, it was found that valid answers on a particular point of interest does not depend on the amount of reports created, rather it depends on the extent of knowledge among the public on them. A System Usability Scale (SUS) questionnaire was given to every participant at the end of the experiment. The prototype's SUS study scored **82.4**, that is, it scored an A, and surpassed the average score of 68. The score obtained means that the participants found the application usable and satisfactory in terms of user experience. Findings are further supported by a Cronbach alpha [10] score of 0.651 which demonstrates that the participants SUS questionnaire answers are reliable.

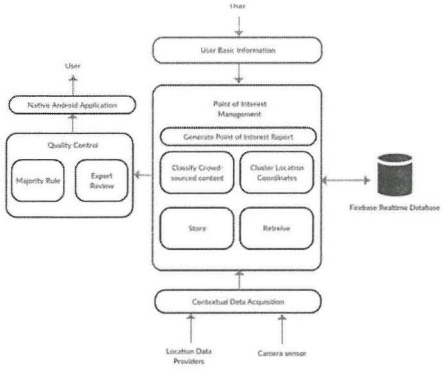


Figure 1. System Block Diagram

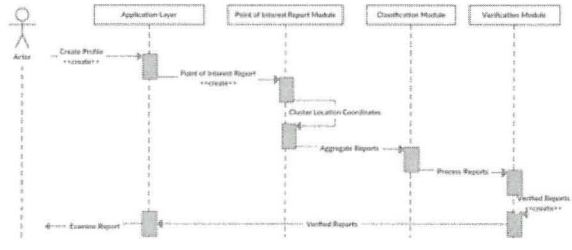


Figure 2. Project Sequence Diagram

REFERENCES

[1] J. Howe, "The Rise of Crowdsourcing | WIRED," 2006. [Online]. Available: https://www.wired.com/2006/06/crowds/. [Accessed: 29-Apr-2017].

[2] G. Chatzimilioudis, A. Konstantinidis, C. Laoudias, and D. Zeinalipour-Yazti, "Crowdsourcing with smartphones," *IEEE Internet Comput.*, vol. 16, no. 5, pp. 36–44, 2012.

[3] M. Stevens and E. D'Hondt, "Crowdsourcing of pollution data using smartphones," Proc. Work. *Ubiquitous Crowdsourcing, UbiComp2010*, pp. 1–4, 2010.

[4] S. Geerts, "Discovering crowdsourcing: theory, classification and directions for use," *Tech. Univ.* Eindhoven, 2009.

[5] D. Christin, A. Reinhardt, S. S. Kanhere, and M. Hollick, "A survey on privacy in mobile participatory sensing applications," *J. Syst. Softw.*, vol. 84, no. 11, pp. 1928–1946, 2011.

[6] R. K. Rana, C. T. Chou, S. S. Kanhere, N. Bulusu, and W. Hu, "Ear-phone: an end-to-end participatory urban noise mapping system," *Proc. 9th ACM/IEEE Int. Conf. Inf. Process. Sens. Networks - IPSN '10*, p. 105, 2010.

[7] J. Burke, D. Estrin, M. Hansen, N. Ramanathan, S. Reddy, and M. B. Srivastava, "Participatory sensing," *Work. World-Sensor-Web Mob. Device Centric Sens. Networks Appl.*, pp. 117–134, 2006.

[8] M. (Ucla) Mun, S. (Ucla) Reddy, K. Shilton, and N. Yau, "PEIR, the personal environmental impact report, as a platform for participatory sensing systems research," *MobiSys*, pp. 55–68, 2009.

[9] M. Allahbakhsh, B. Benatallah, A. Ignjatovic, H. R. Motahari-Nezhad, E. Bertino, and S. Dustdar, "Quality control in crowdsourcing systems: Issues and directions," *IEEE Internet Comput.*, vol. 17, no. 2, pp. 76–81, 2013.

[10] J. D. Brown, "The Cronbach alpha reliability estimate," *JALT Test. Eval. SIG Newsl.*, vol. 6, no. February, pp. 17–18, 2002

PROJECT NO. 11

The Impact of Information Architecture Techniques on Website Usability

ANDREA ATTARD

SUPERVISED BY
DR CHRIS PORTER

Improving the usability of a website may have drastic effects on customer experience, effectiveness and overall user satisfaction. This may in turn lead to better returns on investment for organizations investing in an online presence or for those building an intranet for internal use. However, the time and money required to run usability tests may sometimes be perceived to be too high to justify the benefits. One low-cost approach to improve usability is to re-design a website's Information Architecture: matching the site's structure (e.g. information hierarchies) to the mental model of the website's core users [1].

This study aims to analyse the impact that an inexpensive Information Architecture technique, named card sorting [2], has on the usability of a website.

Two websites were chosen as case studies: The University of Malta website, and the Vodafone Malta website. Personas were selected – representing the core users of each website – and participants under each persona were invited to participate in a card sorting study. The results arising from the card sorts were then used to treat the two websites selected for this study. The Modified-Delphi Method [3] was used to further reduce the cognitive cost of the technique, while providing an easier way to reach consensus.

User performance on treated and untreated website versions was compared through an A/B usability test as part of a blind experiment. Eye-tracking technologies were used to get a better understanding of users' actions while solving the assigned tasks. The data collected was then evaluated to determine if there were any statistically significant improvements in completion rates, time-on-task and participant satisfaction. The latter was measured via standardized metrics generated through the System Usability Scale (SUS), Net Promoter Score (NPS), and Single Ease Questions (SEQ).

The results show strong evidence to support the claim that card sorting has a significant impact on usability, and that it therefore is a valid technique that returns substantial benefits with minimal investment. Additionally, qualitative insights gained throughout the course of the study are also presented, as well as issues encountered by users.

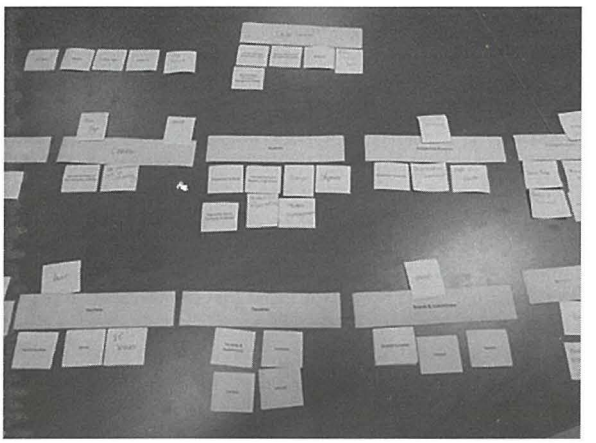


Figure 1: The Modified Delphi card sorting method

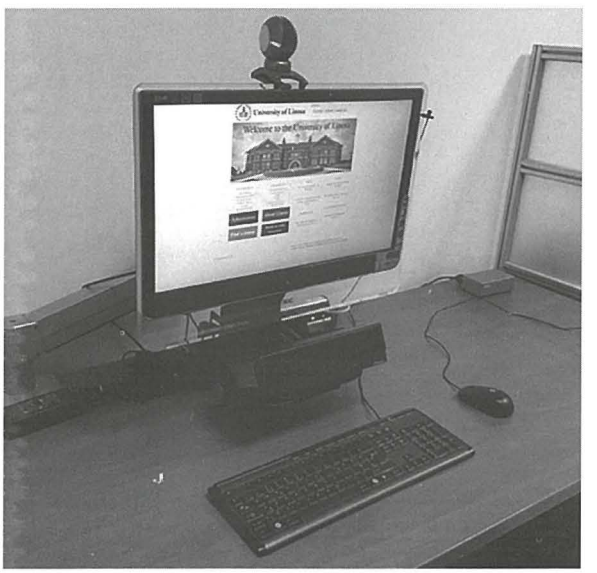


Figure 2: The website usability experiment setup

REFERENCES

[1] Peter Morville and Louis Rosenfeld, *Information Architecture for the World Wide Web*. Sebastopol, CA: O'Reilly Media, Inc., 2006.
[2] Donna Spencer. (2004, April) Card sorting: a definitive guide. [Online]. http://boxesandarrows.com/card-sorting-a-definitive-guide/
[3] Celeste Lyn Paul, "A Modified Delphi Approach to a New Card Sorting Methodology," *Journal of Usability Studies*, vol. 4, no. 1, pp. 7-30, 2008.

Smart Home Automation System

PROJECT NO. 70

JENNY ATTARD

SUPERVISED BY
PROF. MATTHEW MONTEBELLO

Ambient Intelligence is currently a well-known aspect being used in several different areas of study, this allows Artificial Intelligent techniques to perform actions which are usually done by humans. In this thesis an intelligent smart home automation system is designed, structured, implemented and tested in a real life environment to determine whether it could improve the standard of living.

Rather than using the more common machine learning components such as Neural Networks, naïve Bayes classifier [1] Decision Trees and SVMs which require training and acting phases, Pattern Recognition is used to determine the behavior of multiple users living in such a household. This allows the A.I. component to learn, unlearn and act simultaneously. For this reason no distinct training phase is required but all the user has to do is use the system. Additionally the AI mechanism should be as independent as possible of the individual devices, their type and features.

Although distributed topologies are not very common [2] in smart houses, the latter increases the scalability and simplifies the processing mechanism as data processing is split into blocks and shared across different nodes. Moreover, it also offers fault tolerance when compared to the most common topology which is centralized. In the latter case, the central processing component acts as a single point of failure and causes the whole system to stop working whenever a problem is encountered. For this reason a distributed topology is used to achieve a higher standard of living.

The A.I. mechanism designed and implemented for the proof of concept is divided in 3 components: firstly the *Environment Awareness* scouts what devices are available on the network, secondly the *Learning Component* recognizes users' patterns and finally the *Action Component* which instructs the devices to perform a task. Moreover this uses a distributed concept where multiple instances of the A.I. engine running on multiple physical devices collaborate and share data.

The AI mechanism used in this prototype was evaluated using F-measure and an almost perfect score was reached when configuration of the prototype was done properly and the users used the prototype correctly. This study concluded that such a system could indeed improve the standard of living.

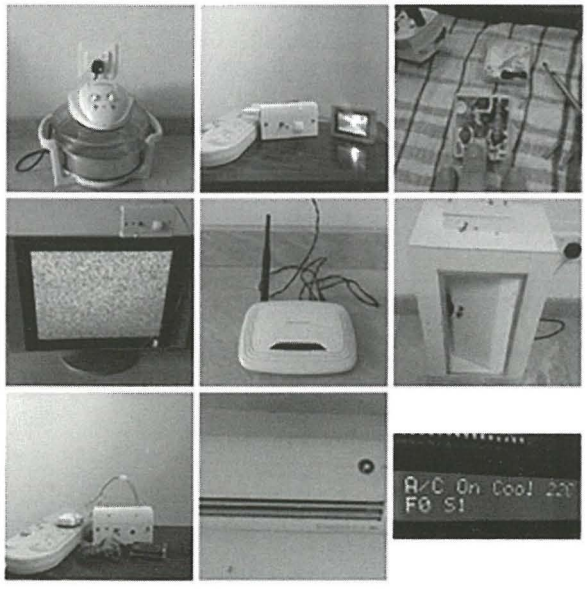


Figure 1: The devices used in order for real testing to take place

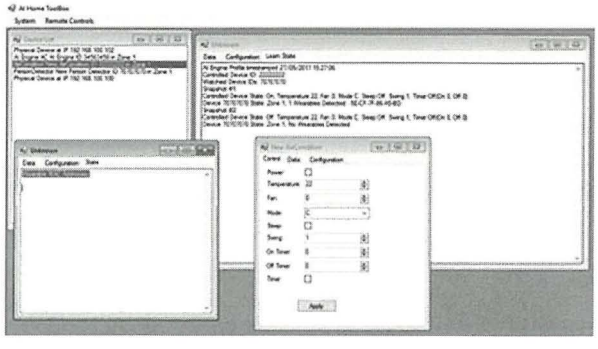


Figure 2: Air condition is switched on when the wearable device is within range

REFERENCES

[1] D. J. Cook, "Learning Setting-Generalized Activity Models for Smart Spaces," pp. 1-13, 2011.

[2] P. Lynggaard, "Artificial Intelligence and Internet of Things in a "smart Home" Context", Aalborg University, 2014

PROJECT NO. 68

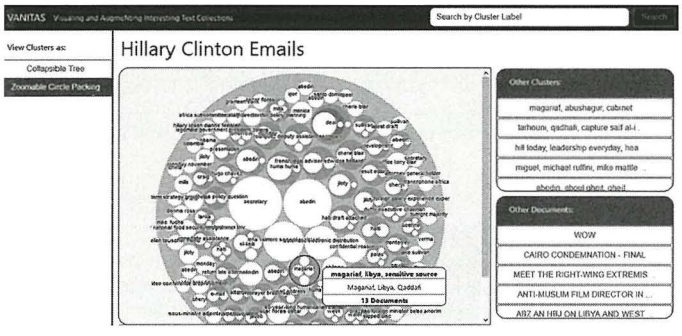
VANITAS – Visualising and Augmenting Interesting Text Collections

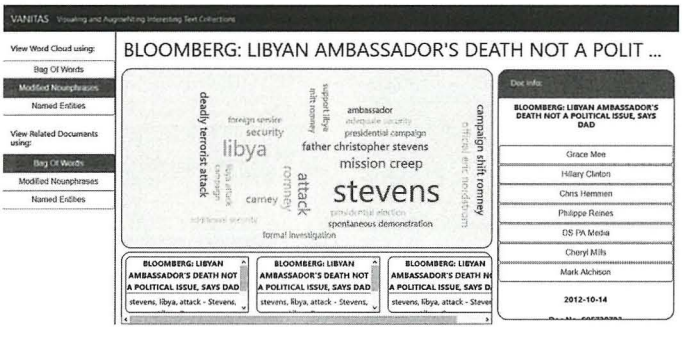
AYRTON SENNA AZZOPARDI

SUPERVISED BY
DR JOEL AZZOPARDI AND DR CHARLIE ABELA

With the evolution of internet technologies, data available on the World Wide Web is constantly being created at a rate too fast for humans to process. Despite being extremely large, web data is mostly accessed in textual format and thus unstructured. As a result, internet users are forced to mine through large amounts of textual data every day. For this reason there is a growing need for automatic and efficient tools that are able to extract and visualise meaningful information so that users are able to observe this knowledge at a simple glance.

The aim of this dissertation was to research and develop computational text-analytical methods as well as interactive visualisation techniques, whereby users can obtain an overview, of interesting document collections without needing to read the whole text, whilst being able to 'drilldown' in the information. In view of this, this project presents VANITAS, a document visualisation tool that can be applied on any document collection, provided that the documents are in English. It makes use of a number of computational text-analytical methods, aiming to extract meaningful relevant information. In fact, this tool summarises documents by utilising a number of different





representations (Bag-Of-Words, keyphrases and named entities) that are then used to group related documents into clusters. Additionally, VANITAS is capable of extracting and gathering knowledge regarding 'identifiers of interest', by augmenting information from other sources (Wikipedia). Finally, VANITAS presents this information using various visualisation techniques that comprise its user interface.

Whilst evaluating VANITAS, our implemented keyword phrase extraction algorithm was found to perform quite satisfactory when considering a few salient keyphrases. Furthermore, the utilised No-K-Means [1] clustering algorithm proved to be sufficient at creating fine-grained clusters surpassing state-of-the-art algorithms, despite not knowing the number of clusters beforehand. Finally, a user study showed that VANITAS is useful in gaining an overview of document collections and for this reason has real-world application value to end-users.

REFERENCES

[1] J. Azzopardi and C. Staff. Incremental clustering of news reports. *Algorithms*,5(3):364-378, 2012.

PROJECT NO. 12

Detection of Fast Flux Botnets Networks

FRANK PAUL AZZOPARDI

SUPERVISED BY
MR CLYDE MELI

Botnets are mainly large collections of devices running a specific malware. This malware typically awaits instructions from one or more Command and Control (C&C) servers which it then follows. Therefore, hiding and protecting C&Cs is crucial. This is in fact the motivation behind the use of Fast Flux techniques which use a large pool of compromised IP addresses to form a proxy layer to conceal the C&C's IPs. However, research has shown that the concept of Fast Flux has evolved considerably both in its form and pace; Domain Flux [1], Single-IP Flux [2], NS Flux [3], Double Flux [4], DNS Tunnelling [5] and even riding the wave of the IoT revolution to make detection of C&C more difficult [6] and [7].

The aim of this project is to analyse various Flux techniques and extract features that indicate Flux behaviour. An Anti-Fast Flux software artefact is proposed as figure 2 as a proof of concept, with the primary aim to detect domains or IPs that exhibit such behaviour. This acts as proxy between the client/network and the external DNS server to passively analyse traffic and look for anomaly features extracted from the literature. Also, it automatically blacklists any unsolicited IPs that attempt to connect with Telnet honeypot which could be a result of Internet of Things (IoT)-based botnets. The system uses a scoring mechanism to collate different types of Flux detections and assigns scores according the exhibited features while blacklisting any domains that exceeds the score of 100. The system successfully classifies fast flux domains with 74.6% accuracy taking into consideration DNS changes since it processes live data. An as SMTP proxy can be also incorporated to cross check to increase accuracy rates.

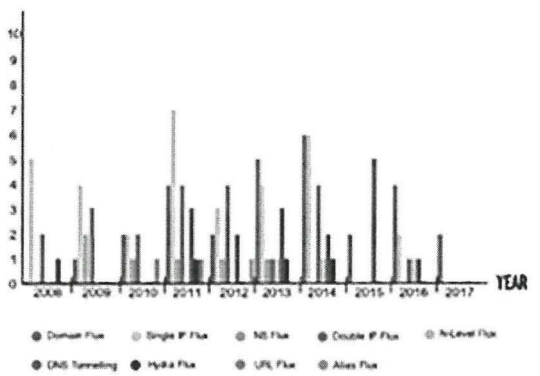


Figure 1: Flux techniques Evolution

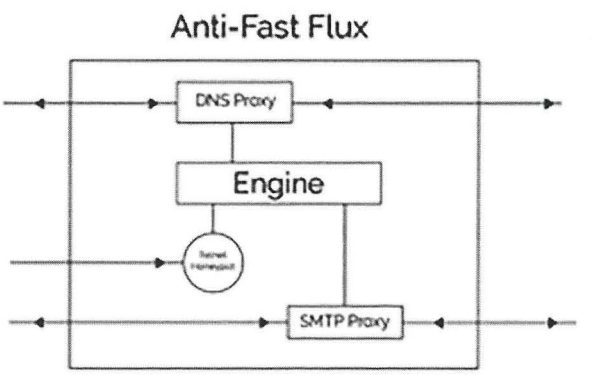


Figure 2

REFERENCES

[1] S. Yadav, A.K.K. Reddy, and S. Ranjan, "Detecting Algorithmically Generated Domain-Flux Attacks with DNS Traffic Analysis," *IEEEACM Trans. Netw.*, vol. 20, no. 5, pp. 1663–1677, 2012.

[2] W. Xu, X. Wang, and H. Xie, "New Trends in FastFlux Networks," presented at the Black Hat Conference, 2013.

[3] A. Caglayan, M. Toothaker, D. Drapeau, D. Burke, and G. Eaton, "Real-Time Detection of Fast Flux Service Networks," in *2009 Cybersecurity Applications & Technology Conference for Homeland Security*, 2009, pp. 285–292.

[4] Riden, "HOW FAST-FLUX SERVICE NETWORKS WORK | The Honeynet Project." [Online]. Available: https://www.honeynet.org/node/132.

[5] OpenDNS, *The Role of DNS in Botnet Command & Control*. 2012.

[6] K. Angrishi, "Turning Internet of Things(IoT) into Internet of Vulnerabilities (IoV) : IoT Botnets," *Eprint ArXiv*, 2017.

[7] "How the Grinch Stole IoT," *Netformation*. [Online]. Available: http://www.netformation.com/level-3-pov/how-the-grinch-stole-iot.

PROJECT NO. 13

Advanced AI for Enemy Management in Digital Games

GEORGI HRISTOV BESHOVSKI

SUPERVISED BY
DR PETER XUEREB AND DR ANTONIOS LIAPIS

The video game industry suffers from increasing costs of production. From multimillion dollar projects to single developers, new techniques for increasing game value are needed. Players in many products become too powerful in the final stages of a game, which decreases late game playability and player satisfaction. A set of rules implemented in the design stage of the game can be monitored by an AI controller and can boost the game playability duration without financial overheads. This is achieved through the AI controller, which, through observation, deduces which enemies the player is weak against and spawns more of them.

The problem consists of 3 aspects:

- Creating a rule balanced game, where each and every 'power' comes with an advantage and a disadvantage.
- Creating an AI module which can be integrated into the game and control the spawning of enemies on the scene.
- Establishing that the AI is providing a greater challenge than a mechanism which spawns the waves of enemies at random.

CREATING A RULE BALANCED GAME

The figure shows the element relations between the different 'elements' at work in the prototype game. The blue arrows indicate an element which deals twice as much damage to the subsequent element whereas the red arrows deal only half the damage to the previous element.

There are 2 cases where these rules are implemented: when the player hits an enemy and when the player is hit by an enemy. For example, if the player chooses 'earth' and strikes a 'lightning' monster, the monster will receive 2X damage whereas if the player is hit by the monster he will receive 0.5X damage.

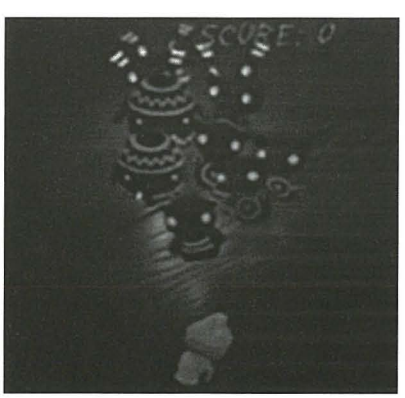
CREATING AN AI MODULE

Damage Based

This was the initial option: an AI controller which used health lost as a measuring function to define which wave is better suited to kill the player. It was found to be insufficient.

Stats Based

This was the final AI which built upon the foundation of the Damage Based AI. It tracked the health lost but also started tracking how many attacks were needed to clear the group of enemies, how many seconds it took to clear them and how much distance the player covered. It used these four aspects incrementally to determine the optimal wave to kill the player.



REFERENCES

[1] R. Hunicke and V. Chapman, "AI for Dynamic DiNculty Adjustment in Games", 2016
[2] O. Missura, "Dynamic DiNculty Adjustment", olana.missura@uni-bonn.de, 2016.
[3] "Artificial Intelligence: A Modern Approach", Aima.cs.berkeley.edu, 2016
[4] M. Silva, V. Silva and L. Chaimowicz, "Dynamic DiNculty Adjustment Through an Adaptive A", 2016

PROJECT NO. 41

Heart Beat Rate Calculation from Facial Video Recording on Smartphones

NICHOLAS BONELLO

SUPERVISED BY
PROF. ING. CARL JAMES DEBONO

Heart rate estimation is an important tool that could single-handedly determine the overall health of a person [2]. However, measuring the heart rate accurately requires some form of specialised equipment that often tend to be too costly for the general public and therefore ends up getting disregarded.

In this final year project, an Android application capable of evaluating a person's heart rate through a video recording of his/her facial region was developed. As blood flows into the blood vessels close to the skin the facial region becomes temporarily redder, however human vision does not have the spatio-temporal capabilities required to be able to detect these minor colour changes. The developed Android application attempts to amplify these aforementioned near invisible colour changes over time so that the user can clearly visualise them. By clearly amplifying the colour changes in a manner that does not also amplify noise, the application will also be able to accurately count the number of times blood flows in and out of the facial region and thus extrapolating these results to evaluate a beats per minute (BPM) value.

A recently proposed algorithm, known as Eulerian Video Magnification (EVM), provides the capability of detecting and amplifying near invisible changes in both motion and colour [1] through the usage of various image processing techniques. The main aim of this FYP is to create a toned down version of the original EVM algorithm that is capable of amplifying the minor colour changes in a facial region caused by blood circulation in a real-time contact-free manner. Unlike the original algorithm, the proposed solution is expected to obtain correct results in any realistic condition, and not just in a room with a still background and constant lighting conditions. After evaluating the results, the developed application reported heart rates nearly identical to a finger monitor with an average accuracy of 97% over all accepted tests. Scenarios in which there are major changes in lighting or large amounts of movement cause the test to fail, however it is assumed that the user is aware of these limitations.

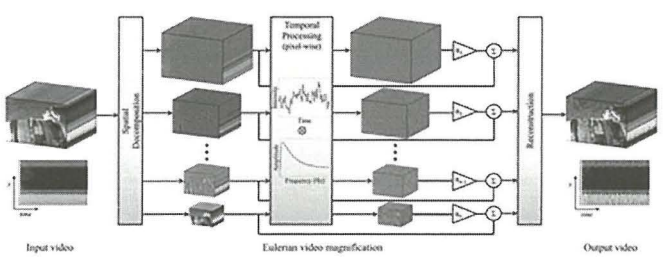


Figure 1: Eulerian Video Magnification Overview [1]

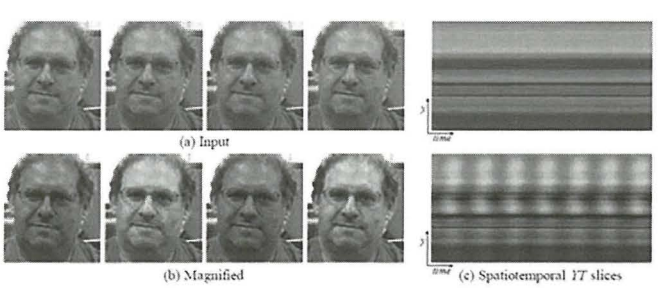


Figure 2: Effects of applying Eulerian Video Magnification algorithm

REFERENCES

[1] H.-Y. Wu, M. Rubinstein, E. Shih, J. V. Guttag, F. Durand, and W. T. Freeman, "Eulerian video magnification for revealing subtle changes in the world," 2012.

[2] M. T. Jensen, P. Suadicani, H. O. Hein, and F. Gyntelberg, "Elevated resting heart rate, physical fitness and all-cause mortality: a 16-year follow-up in the Copenhagen male study," Heart, vol. 99, no. 12, pp. 882–887, 2013.

PROJECT NO. 16

A Model to Optimise Consistency in the Production of Food for Franchise Industry

JESSICA BONNICI

SUPERVISED BY
DR CONRAD ATTARD

Technology plays an important part when designing business process, particularly for manufacturing. A major challenge that factories, especially franchises, face is to track the various stages of several cycles used to produce food, using HACCP. The aim is to study the production processes that are currently being adopted in the factory and optimise it using a business model (Figure 1). The fragmented work that is being done manually will be transferred into a dashboard using an ERP solution. This optimization will allow stakeholders to identify the hazards of the factory in less time, thus making the process more efficient.

Visits were made to a local food factory. Food processes were observed and interviews were conducted with a HACCP expert to get further data relating to HACCP. HACCP is a management system that monitors every step of the process, therefore the chances for hazards to occur is decreased [1]. Four case studies were chosen and represented in a business model using BPMN. The case studies chosen were: Figolla, Kwarezimal (Figure 2), Meat Pie and Qagħaq tal-Għasel. The aim of the BPMN is to provide a notation that all business users can understand [2]. After modelling the case studies, they were carefully analysed to identify at which points a

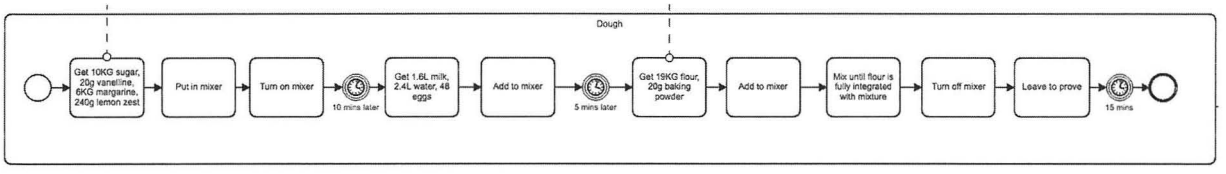


Figure 1: Figolla Production Snippet

hazard could occur. In fact, the four hazard simulations were selected based on these models.

The prototype for this FYP is a dashboard user interface using an ERP, Acumatica which allows different stakeholders to identify hazards happening in the factory related to the food production process. This was done using the data collected from observations and interviews. Two variations of the dashboard were created. A usability experiment was conducted to determine which dashboard is preferred. The data gathered from the usability study was analysed. Participants agreed that it is a very important system to have in as they said that it is helpful to control food safety hazards since that will improve consistency that franchise industry would require. It was proven that the system makes it easier to identify problems. A number of recommendations have been suggested to further improve the prototype suggested.

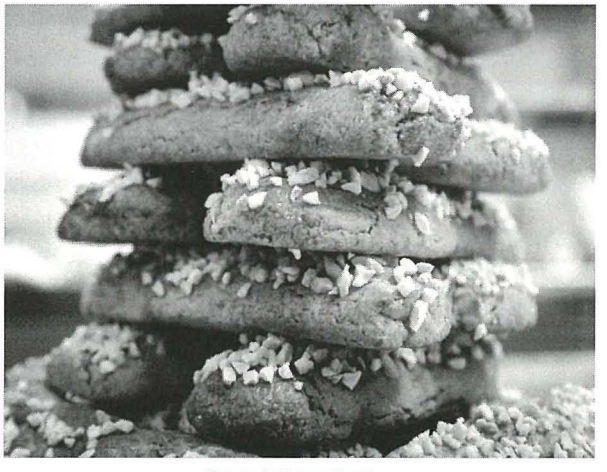


Figure 2: Kwarežimal

REFERENCES

[1] "HACCP Questions and Answers." [Online]. Available: http://haccpalliance.org/alliance/haccpqa.html. [Accessed: 16-Dec-2016].

[2] S. A. White, "Introduction to BPMN," 2006.

PROJECT NO. 18

Implementation of an Automated Event Scheduling System

JAMES BORG

SUPERVISED BY
DR JOSEPH VELLA

Automation is becoming more frequent in many of our daily tasks. Scheduling meetings in an appropriate candidate for automation because of the tedium and difficulties it presents when coordinating with other people.

The aim of this study is to design and implement a scheduler that is capable of finding a suitable time slot for all the invitees based on constraints set by the proposer and with minimal human interaction. Since scheduling is, in general, an NP-Complete problem, heuristics needed to be developed in order to ensure that the system did not spend too long, and waste many resources, trying to solve the scheduling problem.

The scheduling process involves checking if the invitees are all available in the indicated time slot, and attempts to reschedule the meeting if not. Indications of invitees on when they do not wish to meet are also considered by the scheduler when checking for availability.

All data pertaining to a meeting is stored in a database which is in turn a basis for the scheduling algorithms to make decisions during the scheduling process. The two algorithms implemented in this project are Round Robin and Priority Round Robin. The latter is used only when a meeting of importance needs to be scheduled. The user interacts with the system using a web application, which allows the user to create new meeting, view details of scheduled meetings and accept or reject invitations. Users are also able to schedule a meeting on one of two third party applications, i.e. Google Calendar and Doodle, using their respective API.

The insights gained and results from this artefact development include: negotiation between invitees is key to the success of a scheduler allocation; the use of two scheduling methods to allow the scheduler flexibility in its search for a solution is very effective; having a priority levels for meeting and invitees offers a more fair allocation of meeting slots.

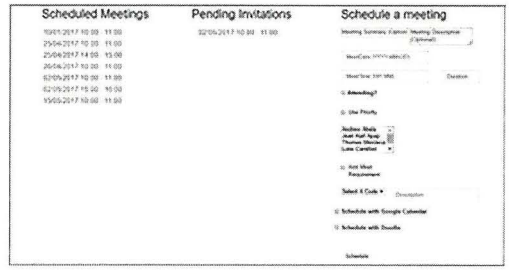


Figure 1: Main Screen of Scheduler

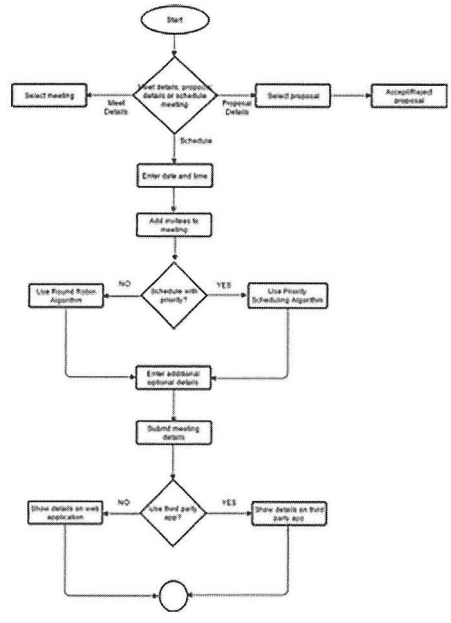


Figure 2: Process of Scheduling Application

PROJECT NO. 42

A Generic Chess-Like Game Player

JONATHAN BORG

SUPERVISED BY
PROF. GORDON PACE AND DR JEAN-PAUL EBEJER

The objective of this project is to investigate the relationship between the two features, **compactness** and **expressivity**, of a domain-specific language. From this relationship one can develop a domain-specific language which fits this relationship.

Domain-specific languages are programming languages which are targeted towards a specific problem [1, 2]. When trying to describe a class of chess-like variants, domain-specific languages provide a plausible notation, as these can be altered and modified to cater for our problem domain.

Prior to the publication of this research, numerous domain-specific languages (DSL) catering for the description of chess-like variants, were developed.

Pell [4] and Browne [5] created DSLs capable of describing a number of diverse chess variants. Pell took a more generic approach, by making numerous assumptions and reducing the number of chess variants which can be described by his language. This resulted in a compact language, with restricted expressivity. On the other hand, Browne had an expressive language which lacked compactness.

In this project, we reviewed and evaluated Pell and Browne's domain-specific languages which tackled the problem while taking different approaches. For this study, we measured the strengths of the domain-specific languages by the level of expressivity and compactness. The level of expressivity measures how many chess-like game variants can be described using a single domain-specific language. The level of compactness measures the ease of describing a variant without any unnecessary or redundant definitions.

Within "The classified encyclopaedia of chess variants", Pritchard [3] describes approximately 1500 chess variants. From this list of chess variants, we chose two random subsets. The first subset, called the language design set, will be used to model the DSL, depending on any key attributes, similarities and any information deemed important to a game description. The second subset, the evaluation set, will be used to measure how many chess variants are completely describable, partially describable or indescribable using the DSLs. We will also measure how compact a DSL is be comparing similar descriptions against each other.

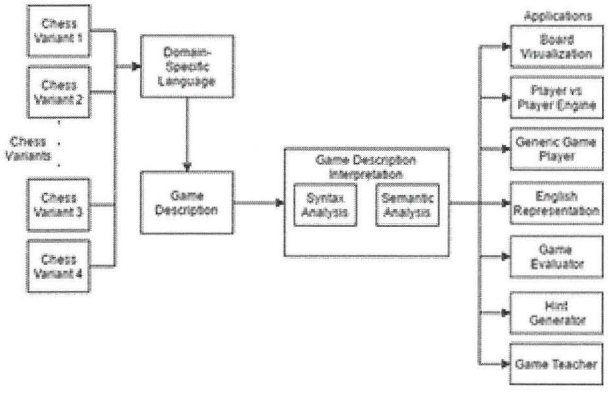


Figure 2: The flow of operations from describing a chess variant to creating an application

Once the balance between compactness and expressivity of the DSL has been found, some applications making use of the game description are developed, as shown in Figure 1.

A generic game player is a specifically designed application able to play any specified game without any human interaction, making use of game playing algorithms such as alpha-beta pruning [6]. The generic game player takes the rules and specifications of the game as an input and returns a player able of evaluating and applying the input specified. The English representation application, provides a better explanation of the rules defined through the DSL. This is done by converting the definitions into a more readable format. The player versus player application allows for players to play against each other, while the application monitors their actions, ensuring that each move is valid according to the game description.

REFERENCES

[1] D. Ghosh, DSLs in action. Manning Publications Co., 2010.

[2] S. Mauw, W. T. Wiersma, and T. A. C. Willemse, "Language-driven system design," in Proceedings of the 35th Annual Hawaii International Conference on System Sciences, pp. 3637–3646, Jan 2002.

[3] D. B. Pritchard and J. D. Beasley, The classified encyclopedia of chess variants. J.Beasley, 2007.

[4] B. Pell, Strategy generation and evaluation for meta-game playing. PhD thesis, Citeseer,1993.

[5] C. Browne, Automatic generation and evaluation of recombination games. PhD thesis, Queensland University of Technology, 2008.

[6] R. Stuart and N. Peter, "Artificial intelligence-a modern approach 3rd ed," 2016.

PROJECT NO. 19

Increasing Taxpayer Value for Money Through Increased Usage of eHealth Services

JOSEPH BORG

SUPERVISED BY
DR PETER XUEREB

eHealth allows citizens to digitally monitor their own health data online and supports health care professionals with tools to improve the care they offer. Despite *eHealth*'s benefits, a lack of uptake *eHealth* services in Malta suggests a need to promote and increase awareness of these services.

After a survey with 205 respondents carried out as part of this project, it emerged that 23% of Maltese citizens make use of *eHealth* services, while only 7% access *myHealth*, the portal where citizens can monitor their health data online. Although significant investments have been made in the development of the Ministry of Health's *eHealth* portal (http://www.health.gov.mt) [1] and its *myHealth* portal (https://myhealth.gov.mt) [2], which requires authentication in order to use, the current usage of eGovernment health services in Malta is very low.

This project aimed to discover why the usage of such a widely applicable service is so low in Malta. Domain expert interviews and usability testing on the *eHealth* and *myHealth* portals were also carried out, in addition to the survey. Maltese citizens expressed their opinion regarding Malta's public *eHealth* services, suggested how could they be improved and recommended various new services which could be introduced.

Following the process of data collection and detailed data analysis, an *eHealth* prototype portal was developed to improve the usability of the current (live) *eHealth* portal and add new services, as requested by users. Elements of the *myHealth* portal that should not require authentication were moved into the *eHealth* portal for greater accessibility.
The results showed that the introduction of completely new services to the *eHealth* website would attract more users to it (see Figure 1); also that the newly accessible *myHealth* services from the *eHealth* portal would encourage more users to start using those *myHealth* services that still required authentication.

Furthermore, the usability testing of the newly redesigned *eHealth* prototype portal showed a marked improvement in time taken to find information on the site (see Figure 2); moreover, users managed to successfully complete substantially more tasks in the *eHealth* prototype portal than in its currently live counterpart, the official *eHealth* portal of the Government of Malta.

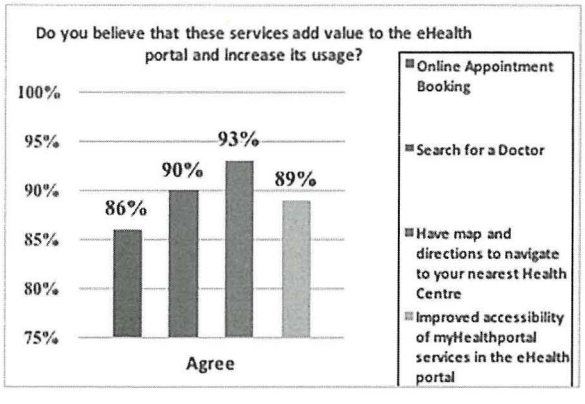


Figure 1: Woulf the introduction of these services give an added value to the eHealth portal, hence increasing its usage?

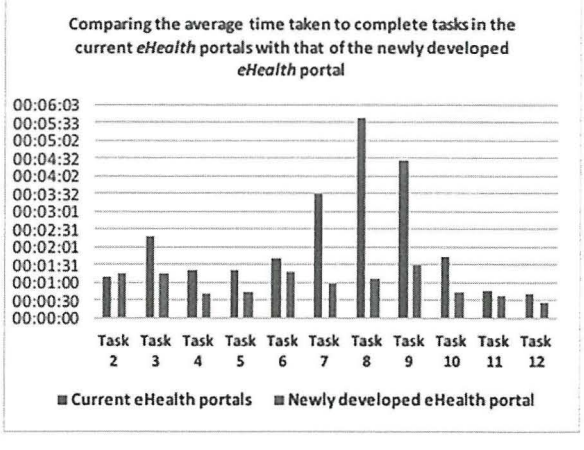


Figure 2

REFERENCES

[1] "health.gov.mt," Ministry For Health, [Online]. Available: https://www.health.gov.mt

[2] "myHealth," Ministry for Health, [Online]. Available: https://myhealth-ng.gov.mt

PROJECT NO. 20

Augmented Reality in Exploratory Testing Activities

SAMUEL BORG

SUPERVISED BY
DR CHRIS PORTER

Exploratory software testing is the process of inspecting all or part of a software application with the aim of deciding if the current state of the software satisfies its requirements or not. Thus, the process plays an essential role in software engineering as it requires executing the software manually to identify any errors, unexpected behaviour or missing functionality against the specified requirements [1]. Software testers need to design their test strategy based on the information given, served from a variety of sources including system documentation, system change logs, textbooks and results of previous tests, to name a few [2]. System testers need also maintain awareness of any information changes in order to preserve the relevance of the tests.

As software systems are rapidly evolving and becoming increasingly more complex, testers can be overwhelmed with the amount of information that is required to be understood, made aware of, and made use of. Hence, such issues commonly lead to a phenomenon known as 'information anxiety' [3] which in turn could result in avoidance of information [4] and which may negatively affect the quality of the end-product. This project proposes that state-of-the-art technology and research techniques borrowed from the field of HCI (Human-Computer Interaction) can support the human tester to be more efficient throughout the software testing processes. It is believed that information anxiety can be reduced by taking advantage of the early yet rapidly evolving augmented reality technology to present just-in-time and contextually relevant information, related to what the tester is currently doing on the system-under-test (SUT).

The augmented reality workbench enables the use of information overlays around the user's monitor with each overlay accurately positioned according to the user's preference. Physical markers on which overlays are projected enable the possibility of organizing what information is to be displayed while allowing for custom setups to improve work ergonomics [5], much like a real workbench – albeit dynamically adapting to the tester's current context within the SUT.

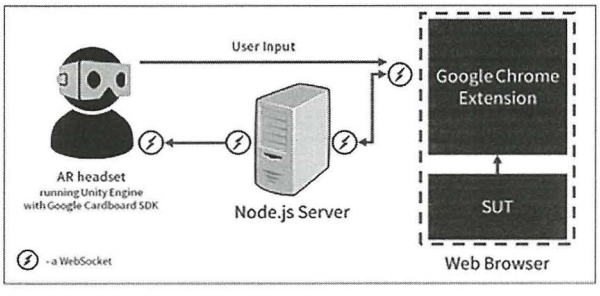


Figure 1: Overview of the system architecture

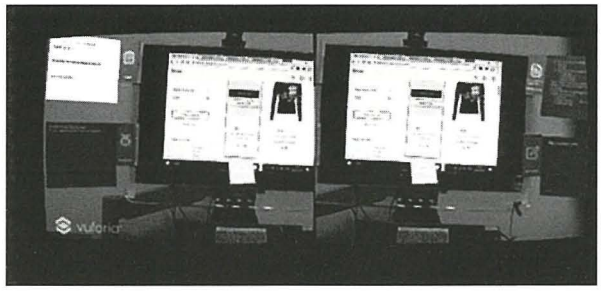


Figure 2: Screenshot from the Augmented Reality application

REFERENCES

[1] J. Bach, "Exploratory testing explained," 2003.
[2] G. Myers, C. Sandler, and T. Badgett, "*The Art of Software Testing*", 2011.
[3] R. Wurman, "Information anxiety," pp. 4–6, 1989.
[4] D. Bawden and L. Robinson, "The dark side of information: overload, anxiety and other paradoxes and pathologies," *J. Inf. Sci.*, 2009.
[5] S. Henderson and S. Feiner, "Exploring the benefits of augmented reality documentation for maintenance and repair," *IEEE Trans. Vis.*, 2011.

PROJECT NO. 66

N.I.R.V.A.N.A – iNtelligent Information RetrieVal ANd Aggregation

ANDRE' BRIFFA

SUPERVISED BY
DR JOEL AZZOPARDI

With today's huge amount of social media content, it is almost impossible to keep track of all the new content on each happening event. Social Media includes all sources where peers and friends can interact with each other such as Blogs, Facebook, Twitter etc. Hence, if a social media user is searching for a piece of news, it could take him more than a couple of minutes to retrieve it. Moreover, if the user is in a hurry, it could take him a long time to read a piece of social media content to discover whether or not an event is of interest to him.

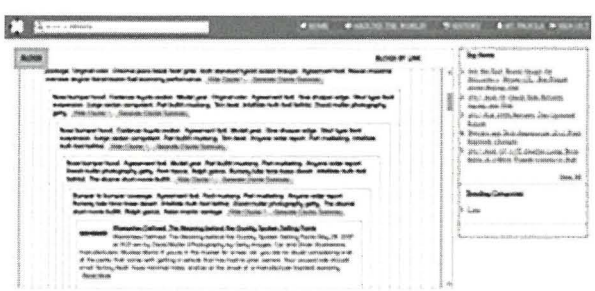
We are proposing a system whereby a user can obtain a quick overview of the latest blog and social media posts. Our proposed solution is made up of different components such as text categorization, sentence extraction summarization and keyphrase extraction summarization. Moreover, we propose a solution to solve the topic identification problem beforehand by applying text categorization on social media posts. Posts related to each particular event is placed within a seperate cluster and the system will display these event clusters to the user.

We also implement automatic summarization systems to generate short 'overviews' of the posts' content. We developed two main approaches i.e. Sentence Extraction where actual sentences are exacted from the original text and Keyphrase extraction approach where keyphrases are extracted from source text. This is done by utilising different techniques such Parts of Speech Tagging, Named Entity Recognition etc.

Finally the categories, Sentences and keyphrases are displayed in a drill-down manner as can be seen in the adjacent figure; as well as in structural manner and utilising keyword clouds as shown in the figure below.

As for the results, results were quite satisfactory especially when it comes to Text Categorization. However, the summarization algorithms underperformed at times.

Apart from quantitative evaluation, we also perform a qualitative evaluation where survey to a number of students was handed out and the results were analysed to come up with a conclusion that our system would be helpful in the real world as it would help the user optimise his time even better.



PROJECT NO. 43

Automated Checking for Deterministic Monitor Behaviour

ANDREW BUHAGIAR

SUPERVISED BY
DR ADRIAN FRANCALANZA

Monitors are computational entities typically *instrumented* with other computational entities, referred to as *systems*, with the intent of collecting system information or verify its behaviour [1]. Monitors are central to several software engineering techniques including monitor-oriented programming and runtime verification, a lightweight verification technique where monitors observe the trace produced during a system run. Monitors are often considered to form part of the Trusted Computing Base (TCB), and as a result are expected to be correct. A frequent correctness requirement is often assumed that they should exhibit *deterministic behaviour*.

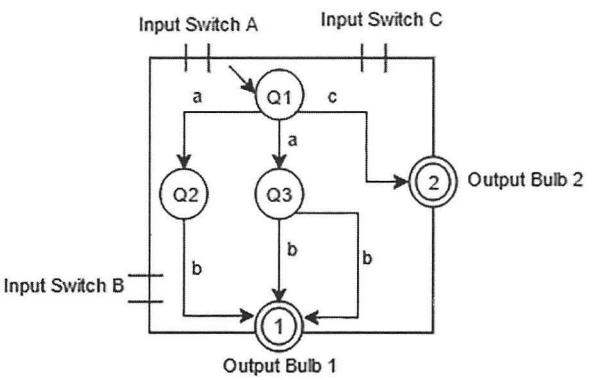
One such definition of determinism is through the notion of *controllability* [2] [3], which in discrete event systems, refers to the ability to steer a (passive) entity to designated terminal states via a series of admissible controls. In [4], Francalanza develops a *co-inductive* definition for *controllability* that can be used to automate the process of checking a monitor description for determinism.

Intuitively, a monitor is controllable if for a particular trace it consistently reaches the same verdict. Figure 1 illustrates a simple lightbulb system with three input switches (Input Switch A, B and C which accept inputs 'a', 'b' and 'c' respectively), and two lightbulbs as outputs. The internal automaton (which is typically hidden) is illustrated for the sake of this example. Certain sequences of inputs will switch on either one of the two lightbulbs. Controllability allows for internal non-determinism. For instance, there are three different paths which can be followed for the input sequence *'a.b'*. After observing an input *'a'*, the system may transition to either state **Q2** or **Q3**. From these two states, the system can only proceed to switch on light bulb 1 by observing input *'b'*. Therefore, the trace *'a.b'* produces consistent results regardless of which path is followed internally.

We say that this system is controllable. Systems are said to be controllable when they exhibit consistent results every time the same trace is analysed.

Francalanza [4], proceeds to define a second definition for determinism, termed *symbolic controllability*, which uses *symbolic events* and *conditions* instead of concrete events, thus being able to observe a potentially infinite number of events from every monitor state. He conjectures that both definitions are implementable and also claims that an implementation based on the latter can analyse monitors more effectively.

In this study, we set out to verify the implementability claims for controllability and its symbolic counterpart by producing algorithms for the formal definitions of [4]. Moreover, we present an investigation on the claim that *symbolic controllability* is more effective in detecting monitor determinism when compared to the original definition.



REFERENCES

[1] A. Francalanza, "A Theory of Monitors," in *Foundations of Software Science and Computation Structures: 19th International Conference, FOSSACS 2016*, Springer Berlin Heidelberg, 2016, pp. 145-161.

[2] J. Klamka, "Control System, Robotics and Automation," in *System Characteristics: Stability, Controllability, Observability*, vol. 7, EOLLS, 2009.

[3] J. Dorsey, Continuous and Discrete Control Systems: Modeling, Identification, Design, and Implementation, NY, USA: McGraw-Hill, Inc., 2001.

[4] A. Francalanza, "Consistently-Detecting Monitors," 2017, *Submitted for publication*.

PROJECT NO. 67

Automatic Segmentation of Indoor and Outdoor scenes from Visual Lifelogging

JUAN BUHAGIAR

SUPERVISED BY
DR GEORGE AZZOPARDI, DR NICOLA STRISCIUGLIO AND PROF. NICOLAI PETKOV

Visual lifelogging and indoor & outdoor classification have been researched on several occasions. The focus of this research is to identify and implement the most effective methods to automatically classify images, captured from a visual lifelogger, into indoor and outdoor scenes. Figure 1 show an example of a visual lifelogger in the form of a wearable camera. Additionally, we incorporated these statistics in a smart phone application where the user can monitor his/her activities.

To automatically classify images into indoor & outdoor scenes, we have used pre-trained convolutional neural networks on Imagenet [1] and Places2 [3] and applied transfer learning to adapt to the problem at hand. Figure 2 shows how we have extracted the features from the images. Once the features have been extracted, we have used classifiers to learn and predict indoor and outdoor scenes. This method proved to be useful instantly, obtaining accuracies above 90%. Therefore, we investigated ways to improve the classification further by joining different features, using intermediate layer information from the CNNs (FC7) and improving the classifiers used. We have performed tests on two datasets UBRug dataset and the SUN397 dataset [2]. The UBRug dataset contains 1000 images from a wearable camera while the SUN397 contains 108,000 images from a normal perspective.

When tested on the UBRug dataset, our best method, joining the FC7 intermediate layer's output from both CNNs with a Random Forest classifier, obtained around 98% accuracy. We repeated the tests on the SUN397 dataset and compared our results with previous tests done on the SUN397 dataset. Our methods outperformed these tests with the best performing method from these tests obtained an accuracy rate of 94.2% while our best performing method obtained an accuracy of 97.06%. Moreover, we tested re-training a pre-trained CNN on the UBRug dataset, introducing, the rst Ego-centric indoor/outdoor CNN.

Figure 1: An example of a visual lifelogger

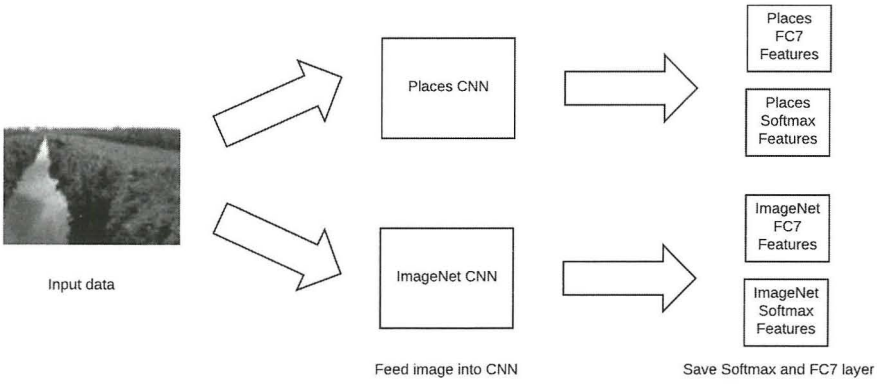


Figure 2: Extracting features using Pretrained CNN

REFERENCES

[1] Krizhevsky, A., Sutskever, I., and Hinton, G. E. Imagenet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems 25, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097-1105.

[2] Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., and Torralba, A. Sun database: Large-scale scene recognition from abbey to zoo. In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (June 2010), pp. 3485-3492.

[3] Zhou, B., Khosla, A., Lapedriza, A., Torralba, A., and Oliva, A. Places: An image database for deep scene understanding. arXiv preprint arXiv:1610.02055 (2016).

PROJECT NO. 34

Security Issues in Controller Area Networks in Automobiles

ROBERT BUTTIGIEG

SUPERVISED BY
MR CLYDE MELI AND DR ING. MARIO FARRUGIA

Exploratory software testing is the process of inspecting all or part of a software application with the aim of deciding if the current state of the software satisfies its requirements or not. Thus, the process plays an essential role in software engineering as it requires executing the software manually to identify any errors, unexpected behaviour or missing functionality against the specified requirements [1]. Software testers need to design their test strategy based on the information given, served from a variety of sources including system documentation, system change logs, textbooks and results of previous tests, to name a few [2]. System testers need also maintain awareness of any information changes in order to preserve the relevance of the tests.

As software systems are rapidly evolving and becoming increasingly more complex, testers can be overwhelmed with the amount of information that is required to be understood, made aware of, and made use of. Hence, such issues commonly lead to a phenomenon known as 'information anxiety' [3] which in turn could result in avoidance of information [4] and which may negatively affect the quality of the end-product.This project proposes that state-of-the-art technology and research techniques borrowed from the field of HCI (Human-Computer Interaction) can support the human tester to be more efficient throughout the software testing processes. It is believed that information anxiety can be reduced by taking advantage of the early yet rapidly evolving augmented reality technology to present just-in-time and contextually relevant information, related to what the tester is currently doing on the system-under-test (SUT).

The augmented reality workbench enables the use of information overlays around the user's monitor with each overlay accurately positioned according to the user's preference. Physical markers on which overlays are projected enable the possibility of organizing what information is to be displayed while allowing for custom setups to improve work ergonomics [5], much like a real workbench – albeit dynamically adapting to the tester's current context within the SUT.

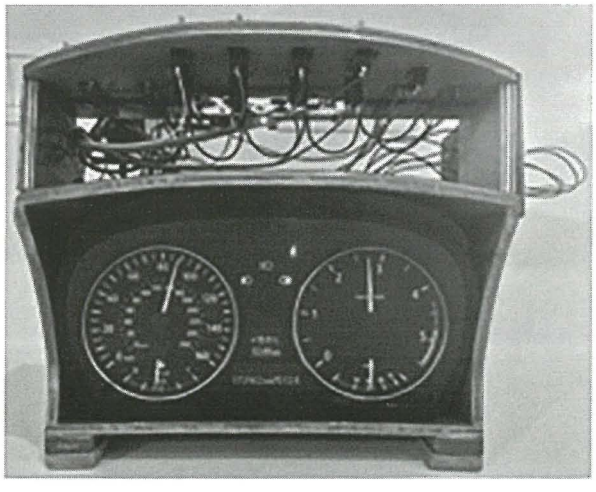


Figure 1: The manual mode allows the user to control the instrument cluster using a hardware interface consisting of slide potentiometers, push button switches and toggle switches.

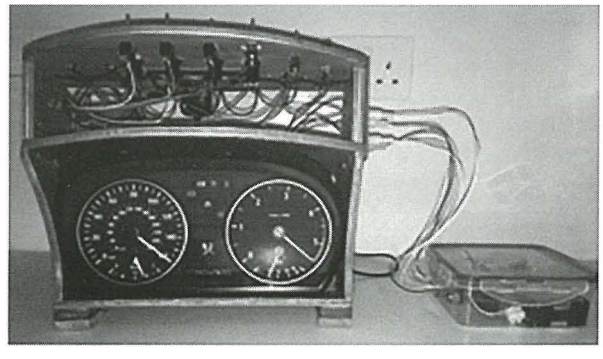


Figure 2: The connection of the rogue device as a man in the middle attack between the vehicle simulator and the instrument cluster. The instrument cluster has been hacked by displaying an RPM value of 5500, a speed of 260km/h and both ABS and airbag systems are disabled.

REFERENCES

[1] O. P. Sharma and P. A. Bhargav, 'Implementation of Controller Area Network', Oct. 2016.

[2] R. Currie, 'Developments in Car Hacking'. SANS Institute, 2016.

[3] P. Carsten, T. R. Andel, M. Yampolskiy, J. T. McDonald, and S. Russ, 'A System to Recognize Intruders in Controller Area Network (CAN)', in *Proceedings of the 3rd International Symposium for ICS & SCADA Cyber Security Research*, 2015, pp. 111–114.

[4] H. Ueda, R. Kurachi, H. Takada, T. Mizutani, M. Inoue, and S. Horihata, 'Security Authentication System for In-Vehicle Network', SEI TECHNICAL REVIEW, no. 81, pp. 5–9, Oct. 2015.

[5] W. Choi, H. J. Jo, S. Woo, J. Y. Chun, J. Park, and D. H. Lee, 'Identifying ECUs Using Inimitable Characteristics of Signals in Controller Area Networks', arXiv preprint arXiv:1607.00497, 2016.

PROJECT NO. 44

JavaScript Operating System Simulation of Processes for Instructional Use

MICHAEL CALLEJA

SUPERVISED BY
DR JOSHUA ELLUL AND DR KEVIN VELLA

Operating Systems is a core topic in computer science, but difficult to learn as a careful balance of theory and practice is needed. This is because purely theoretical models alone do not ensure complete understanding since some of the concepts require detailed explanations. Practical implementation is required as it help users understand how the underlying concepts apply to real world examples [1].

In a lecturing environment, operating systems concepts lectures tend to be challenging to give since no practical implementation is done in class and there is often a large number of students enrolled for the course [2]. The variety in the students' background is also an issue as they may not have the necessary background textbooks expect. Programming tasks are helpful in these cases as hands on experience can help users understand how the concepts apply to real operating systems. These practical tasks also help improve communication between the students and lecturers [3].

In this final year project, the focus is on developing a web based simulation of operating system concepts, focusing on the concepts related to processes. The processes are simulated down to the execution of their individual instructions, in order to simulate processes in a way that can show synchronization issues in a multiprogramming environment. The system runs solely in JavaScript which reduces time spent to set up the system since users have immediate access to web browsers and also since this increases compatibility with the users' hardware and existing software.

The simulation provides the practical aspect required by allowing users to specify and modify their own processes and the instruction set architecture which the process's individual instructions execute. The system also allows users to plug-in and modify their own process scheduling algorithm. Similarly certain hardware parameters of the system can also be created or modified such as the registers and I/O devices the instruction set operates on.

In order to hide the internal simulation components from the users so that they can better develop these implementation examples, the user instead interacts with a framework built on top of the simulator.

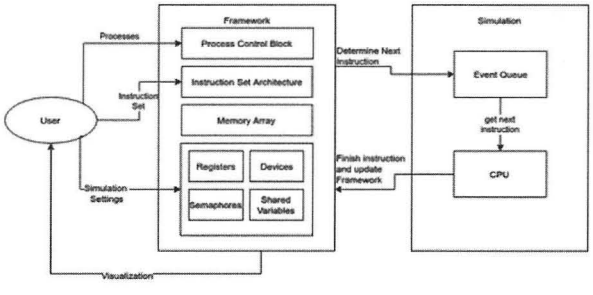


Figure1: System overview

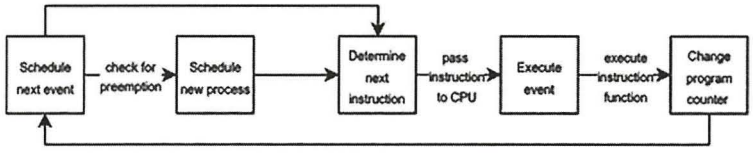


Figure2: How the instruction cycle is simulated using events

REFERENCES

[1] Wayne A Christopher, Steven J Procter, and Thomas E Anderson. The nachos instructional operating system. In USENIX Winter, pages 481–488, 1993.

[2] Felix Buenda and Juan-Carlos Cano. Webgene: A generative and web-based learning architecture to teach operating systems in undergraduate courses. IEEE Transactions on Education, 49(4):464–473, 2006.

[3] Aristogiannis Garmpis and Nicolaos Gouvatsos. Innovative teaching methods in operating systems: The linux case.

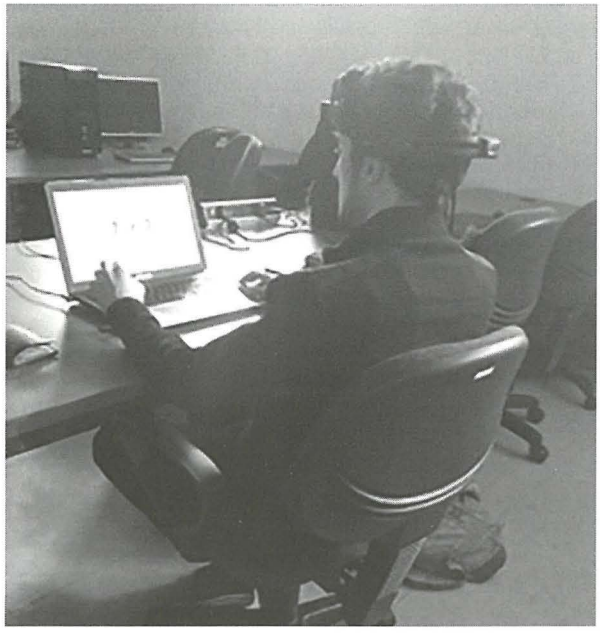
PROJECT NO.
45

Investigating Ways to Improve the Effectiveness of BCI Techniques to Mitigate the Oracle Problem

ELISE CALLUS

SUPERVISED BY
DR MARK MICALLEF AND DR CHRIS PORTER

Software Testing involves examining the behaviour of a system in order to discover faults in the system. Given an input to the system, the Oracle Problem is described as the challenge of distinguishing the correct behaviour of a system from the incorrect behaviour. This study explores the extent to which one can leverage brain computer interface (BCI) devices to mitigate the Oracle Problem. This study improves on Galea et al.'s study [GMP16] by utilising expert feedback, better stimuli presentation and using a device having more channels than the headset used their study. The Emotiv Epoch+1 was used to collect the EEG data. Simple mathematical equations were shown to the participants where the data was then analysed to obtain features for classification. The features used are the average power of the delta, theta, alpha and beta frequency bands obtained from different electroencephalograph (EEG) channels and these were fed to the Linear Discriminant Analysis (LDA) classifier. The best features were chosen using visual and statistical analysis while the classifier was chosen using expert feedback and literature. The results obtained show that using an individual model for every participant obtains an accuracy result of 60% while a global model obtains an average accuracy of 52%. Overall, the global model performs poorly when classifying simple arithmetic equations so complex test cases would probably perform worse. This indicates that the hypothesis that the Oracle Problem can be solved using BCI devices is unlikely to hold. Having said that, this work utilises a specific set of parameters, such as using the power of the frequency bands as features for the LDA classifier and using the arithmetic equations as stimuli, achieving better-than-random results. Whilst the results do not convey sufficient confidence, we would recommend further investigation before the hypothesis is abandoned.



REFERENCES

[1] https://www.emotiv.com/epoc/v[GMP16] Daniela Galea, Mark Micallef, and Chris Porter. Investigating ways to mitigate the oracle problem using hci techniques. 2016.

Investigating the Suitability of COSFIRE Filters for Pedestrian Recognition

ANGIE ANN CAMILLERI

SUPERVISED BY
DR GEORGE AZZOPARDI

Pedestrian recognition is a widely researched area due to its application in pedestrian detection systems as in the cases of surveillance in crowded places and assisted technology to the visually impaired. However, certain factors such as occlusion, light and appearance of pedestrians may hinder the performance of any pedestrian recognition system. Throughout the years, multiple state-of-the-art systems have been developed to cater for such issues including the widely used Histogram of Oriented Gradients (HOG). These systems typically employ classification techniques such as SVM to determine whether any particular scene contains any pedestrian or not. However, some aspects of these systems still fail to handle occlusion and thus misclassification occurs.

The purpose of this project was to investigate the suitability of COSFIRE Filters [1] in relation to pedestrian recognition. Given a pattern of interest, multiple COSFIRE filters are configured by selecting random regions of interest in a pedestrian's body, so that pedestrian descriptors can be configured and trained in an automatic process. The implemented COSFIRE approach is based on the work of Azzopardi et al. in [2]. Moreover, the Histogram of Oriented Gradients [3] approach is implemented to compare and evaluate the performances of these two systems. Since classification techniques vary in speed, we propose the use of two classification techniques, namely Support Vector Machines (SVM) and Random Forests to evaluate the effectiveness of these two approaches. Moreover, three well-established datasets, namely INRIA Person dataset, Daimler Mono Pedestrian Dataset and TUD- Brussels Pedestrian Dataset, which all vary in size and type of images they contain are used to test the two approaches.

The highest accuracy result achieved is that of 98.67% on the INRIA Person dataset using the SVM classifier. This result is remarkably good when compared to other state-of-the-art systems such as HOG, where the highest accuracy rate obtained is 93.55%. The research undertaken for this project resulted in some interesting conclusions on pedestrian recognition and detection problems, which could serve as a foundation for other works in this relatively ever-growing research area.

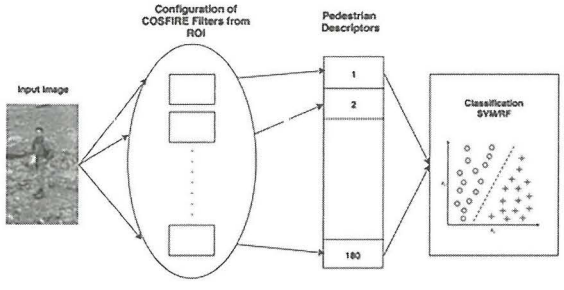


Figure 1: Illustration of system's design

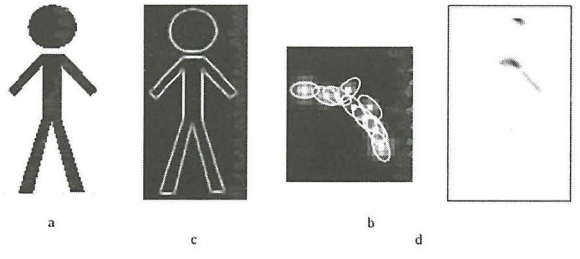


Figure 2: COSFIRE method (a) The training image, with the red-circled region indicating the prototype pattern of interest. (b) The superimposed (inverted) response maps of a bank of Gabor filters with 16 orientations and a single scale. (c) The structure of the COSFIRE filter that is configure on the prototype pattern in the first image. (d) The (inverted) response map of the concerned filter to (a). The darker the pixel, the higher the response.

REFERENCES

[1] George Azzopardi and Nicolai Petkov. Trainable cosfire filters for keypoint detection and pattern recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(2):490–503, 2013.

[2] George Azzopardi, Antonio Greco, and Mario Vento. Gender recognition from face images with trainable cosfire filters. In Advanced Video and Signal Based Surveillance (AVSS), 2016 13th IEEE International Conference on, pages 235– 241. IEEE, 2016.

[3] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1:886–893, 2005.

PROJECT NO. 21

Automating Point Cloud Acquisition Methods For Scene Understanding

JONATHAN CAMILLERI

SUPERVISED BY
DR SANDRO SPINA AND DR KEITH BUGEJA

3D cameras or scanners use depth sensors to measure the distance to the objects in view and produce a set of points in space, a point cloud. This data may hold information such as position and colour (Figure 1), acquired from real-world objects, which can then be reproduced in virtual environments. Advances in scanning hardware and computer visualization technology have made the use of this data more feasible. Applications making use of point clouds have increased, with the acquisition of 3D point information becoming customary practice in many areas including architecture, cultural heritage, manufacturing and urban planning.

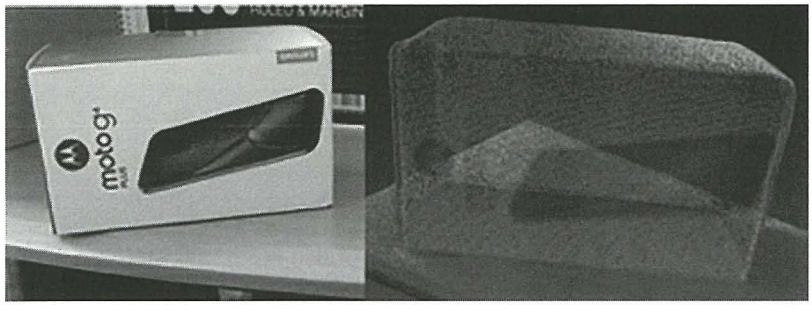


Figure 1. Visualisation of point cloud (right) sampled from box on a desk (left)

Scene understanding (SU) algorithms have traditionally been employed to classify 2D images of scenes into semantic categories; for instance, an input data set classified as an office space. Recently, SU algorithms have also been applied on point clouds to automate the process of determining which objects or structures are present in a scene. A thorough evaluation of these algorithms necessitates the production of several point cloud data-sets, which is generally a lengthy process. The aim of this project was to develop a simulation platform to help alleviate some of the drawbacks of 3D scanning (i.e. producing point cloud data sets) while facilitating other tasks such as scene setup and scanner calibration. This allows for the generation of data sets with user specified characteristics in custom environments.

To achieve this a number of objectives were set:

- Model and simulate the functionality of a real 3D scanner.
- Model specific object surface materials.
- Simulate the 3D acquisition pipeline.
- Compare the point clouds generated in a virtual scene with authentic ones.
- Use the data sets generated to evaluate point cloud processing algorithms typically used in SU.

To generate point clouds with authentic characteristics, a virtual scanner needs to simulate the behaviour of a real 3D scanners. However, replicating all the phenomena involved is a very complex process. For this reason, a heuristic approach is taken to create a mathematical model of the scanner. The scanner would then require a movement mechanism determining where to move and what to scan in a given scene. At the start of a scan a path is automatically generated which the scanner then follows. Objects are scanned depending on their size, geometric complexity and the already acquired samples.

The point clouds generated by the simulation platform were comparable to authentic point clouds in terms of the noise distribution. Moreover, the algorithms developed for the simulation can produce point clouds with high surface coverage in open areas and good surface coverage in occluded areas of the scene. The data sets were used as input to algorithms commonly used in SU and managed to provide reasonable results suggesting that this system has potential in more comprehensively evaluating the SU algorithms.

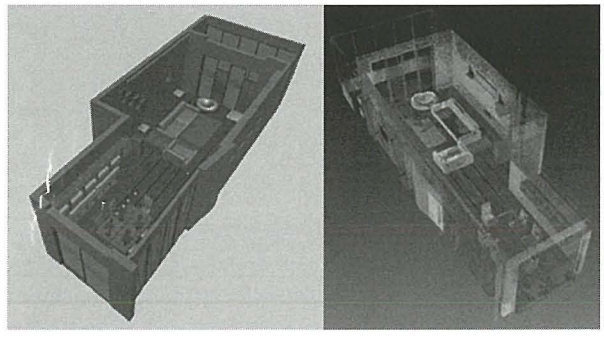


Figure 2: The virtual scene (left) is sampled using our simulator to generate a corresponding point cloud (right)

PROJECT NO. 23

Finding an Optimal Solution to the Car Traffic Congestion Problem

JULIAN PAUL CAMILLERI

SUPERVISED BY
MR TONY SPITERI STAINES

In these modern times and because of urbanization, transport systems are facing several new problems. The quality of service is expected to be high but because of population density in urban regions in Malta, there are traffic issues all around the island. As many major Maltese towns and cities were built before cars were introduced in Malta, there were little to no standards and conventions on how to build a road. Most roads were build narrow, which at the time were more than sufficient.
Currently, there are approximately 335,000 registered vehicles including commercial vehicles, motorcycles, coaches and minivans. This number is constantly rising as around 10,000 new vehicles are being introduced to the Maltese roads every year. Also, apart from new vehicles, new drivers are contributing to the problem. Between the years 2010 and 2016, 6792 learners got their driving permit, hence being able to drive on the Maltese roads [1].

There are about 1.2 billion cars on the road, and it is estimated that by 2035 around 2 billion cars will be in use [2]. Many car manufacturers believe that the autonomous vehicle industry is one that will impact our future very soon and are investing a lot into creating these modern vehicles. Autonomous vehicles will be able to drive closer to the car in front as will have faster reaction times, drive within the speed limits and will behave less randomly than humans as we sometimes tend to overreact in certain situations. As autonomous vehicles do not feature aggressive driving (over braking, over accelerating) styles like humans, less traffic congestion will take place, as the amount of jams due to the butterfly effect and shockwave propaganda will diminished. Furthermore, on a tactical level, cars will behave in a more rational way, as they can communicate with other cars and quickly choose the optimal route according to traffic density and maintenance, much faster than a human can [3].

The primary outcome of this dissertation is to build and implement a web based system which will let the user assess the flow of traffic through three different environments. As Malta is quite a small country, and there is no space to expand the road networks, we have to work with what we have got. As polls suggest, Malta's population is ever increasing. Focusing on drivers is much more essential and sufficient to tackle traffic congestion. Therefore, we will be comparing the differences between the current traffic network with human drivers against a network were autonomous vehicles will be driven. Furthermore, the user will be able to change vital parameters to the system, if need be, to make it more realistic and flexible to mimic real life scenarios.

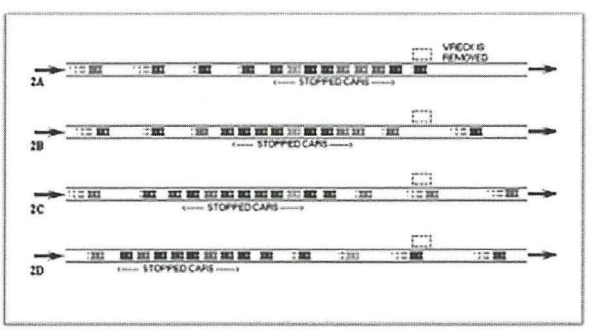


Figure 1: Effects of traffic shockwaves

REFERENCES

[1] National Statistics Office SO, Transport Statistics 2016
[2] Cheah W., "Factor of Two: Halving the Fuel Consumption of New U.S. Automobiles by 2035.", MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 2008.
[3] L. Wilson. (2016, Sept 3). *How math and driverless cars could end traffic jams.*[Online] Available: http://www.businessinsider.com/math-driverless-cars-end-traffic-jams-2016-9.

PROJECT NO. 69

Automatic Runway Detection for Unmanned Aerial Vehicles (UAV)

KEITH CAMILLERI

SUPERVISED BY
DR GEORGE AZZOPARDI AND PROF. ING. DAVID ZAMMIT MANGION

In this project, a system which uses computer vision to automatically detect runways is presented. By using and adapting CORF with push-pull inhibition [1], the runway edges are detected and used to determine the position of the runway threshold midpoint. A computer vision-based approach was chosen over a GPS-only system to counter the lack of accuracy of GPS when calculating distances.

The system uses a dataset of 631 aerial images portraying the possible conditions of a runway, such as light and weather conditions and features of the runway, to maximize effectiveness and accuracy of the runway detection process. The images are obtained both from real-life landings (107 images) and from a flight simulator (324 images), to ensure that most possible scenarios are covered.

A training set made up of 150 random images from the simulation dataset is used to determine the number of detected contours to consider and which set of y-coordinates should be used to select the edges which best describe the runway. Subsequently, these parameters are used to determine the distance error in pixels between the detected and the actual runway threshold midpoints for each image in the dataset.

Dr George Azzopardi and Prof. Ing. David Zammit Mangion The system obtains an average distance of error of approximately 101 pixels, where day and night scenarios obtained average distance errors of approximately 31 and 23 pixels respectively, while rain and live images prove to be more challenging with average distance errors of 125 and 258 pixels respectively. Additionally, factors such as the slant distance of the UAV from the runway, image noise, visibility, and light conditions are shown to have a significant effect on the runway detection algorithm.

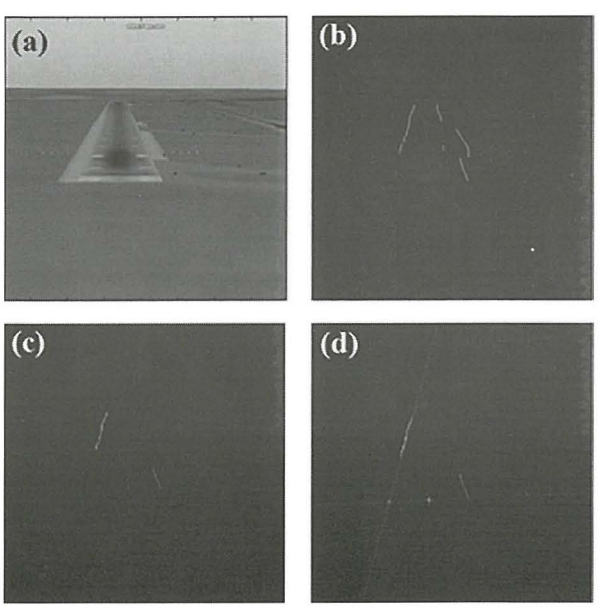


Figure 1 – Visual illustration of the system process for runway detection

a. *A frame from the UAV's point of view is obtained*

b. *Approximately vertical edges are extracted from the frame using the CORF algorithm*

c. *Two lines which best describe the runway are selected based on their length and y-coordinates*

d. *The landing point is estimated using various properties of the selected lines*

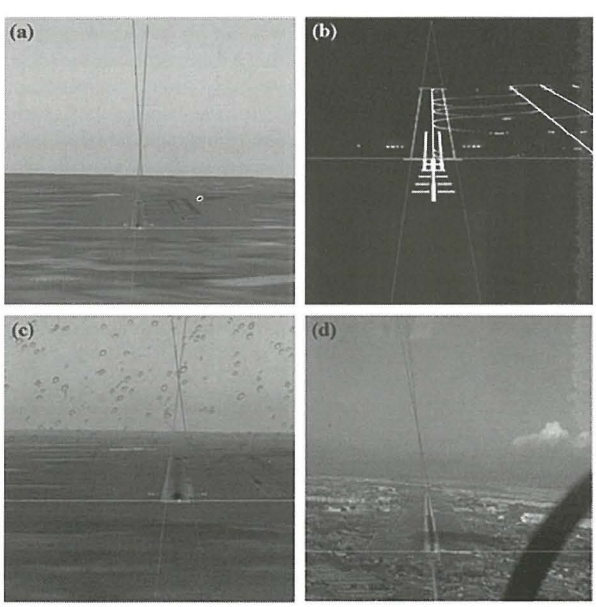


Figure 2: Contour detection in different scenarios

a. *Day scenario*

b. *Night scenario*

c. *Rain scenario*

d. *Live Images*

REFERENCES

[1] G. Azzopardi, A. Rodríguez-Sánchez, J. Piater, and N. Petkov, "A push-pull CORF model of a simple cell with antiphase inhibition improves SNR and contour detection," PLoS One, vol. 9, no. 7, 2014.

PROJECT NO. 17

The Impact of Keyboard Layouts on Dwell-Free Eye-Driven Typing Performance

BERNINE CARUANA

SUPERVISED BY
DR CHRIS PORTER

The modern technique of dwell-free eye typing enables more efficient communication than the conventional, dwell-based technique [1]. Nevertheless, speech and motor impaired individuals who are also digital immigrants, may experience challenges with this technique due to their unfamiliarity with keyboard layouts. Whereas studies evaluated the performance of different keyboard layouts with dwell-based eye typing [2] and physical keyboards [3], their findings cannot be generalised for dwell-free eye typing since this technique is characterised by fundamental interaction differences. The purpose of this study is to provide speech and occupational therapists with insights on keyboard recommendations based on the users' aptitudes for and experience with technology. This study considers three standard keyboards; QWERTY, Dvorak and the alphabetical layout, in light of deriving an optimised keyboard for users with limited typing experience. Twenty digital immigrants used all three keyboards during 45-minute experiment sessions. Results emerging from this study informed the design of a new layout (DF-AJR) which mitigates observed shortcomings in standard layouts for dwell-free eye-typing.

Performance results on dwell-free eye typing using existing keyboard layouts show that digital immigrants (having some years of experience with technology) perform best on the QWERTY layout, followed by the alphabetical and Dvorak keyboards respectively. However, participants who had less than one year of experience with technology performed equally on both QWERTY and alphabetical layouts, suggesting that the QWERTY layout may not yet be clearly imprinted in the minds of this group of users. The need for an alternative keyboard layout optimised for this group was subsequently investigated. Whereas previous studies, which used alternative modes of input, concluded that Dvorak has the potential to outperform QWERTY, this was not observed with dwell-free eye typing. Observed shortcomings for standard layouts were mitigated in the redesigned keyboard, DF-AJR, which utilised the English-letter frequency list and adopted an alphabetical bias. Initial results show that participants' performance improved with the DF-AJR keyboard and was also preferred by participants. Moreover, while introducing dwell-free eye typing over the web, the optimised and open sourced client-side web-based tool developed for this study reduces the financial burden associated with acquiring commercial eye-driven AAC devices.

Figure 1: Contour detection in different scenarios

1. *One of the generated heat maps with gaze plots which identified standard keyboards' shortcomings.*
2. *Dwell-free eye-typing user, Bjorn Formosa, communicating through the developed tool.*

Figure 1: Positioning of characters in the DF-AJR optimised keyboard layout

REFERENCES

[1] P. O. Kristensson and K. Vertanen, "The Potential of Dwell-free Eye-typing for Fast Assistive Gaze Communication," in Proceedings of the Symposium on Eye Tracking Research and Applications, 2012, pp. 241-244.

[2] S. Raupp, "Keyboard Layout in Eye Gaze Communication Access: Typical Vs. Als.", Dept. Comm Disorders, East Carolina Univ., 2013.

[3] D. A. Norman and D. Fisher, "Why alphabetic keyboards are not easy to use: Keyboard layout doesn't much matter," Hum. Factors J. Hum. Factors Ergon. Soc., vol. 24, no. 5, pp. 509-519, 1982.

PROJECT NO. 47

Towards an Alternate Characterisation of the Actor Safety Preorder

CAROLINE CARUANA

SUPERVISED BY
DR ADRIAN FRANCALANZA

Concurrency deals with the structure and composition of independently running processes, as a way to increase efficiency [1]. In fact, in a concurrent environment, tasks may be run in parallel, but can also be interleaved or executed sequentially on a single processor. This can create problems of *non-determinism* if not handled properly; that is, for a particular input, a process can emit different outputs on different runs. Hence, a number of concurrency models emerged to take advantage of the properties of concurrency [2].

In this work, we consider the Actor Model of concurrency, specifically about *actor systems*: a setup of uniquely named actors, interacting with one another via *asynchronous* message passing, *spawning* other actors or changing their own current behaviour in response to messages received [3]. Focusing on how actor systems relate to each other, we start from the widely accepted notion of *safety testing*, where we compare two actor systems by arguing that if they are able to pass all the same tests, only then they are related. Although intuitive, to show this relation we must consider all possible tests, of which there are an infinite number. This renders the process hard to automate.

We therefore shift to relating actors by their *traces*, where a trace refers to a series of input and output interactions occurring between an actor system and its external environment. Rather than relating actor systems by the tests they pass, we relate them by comparing their traces. If an actor system makes the same traces as another, we can say they are related.

We examine a particular trace based technique proven to be correct with respect to safety testing in a previous work [2]. It is already easier to reason about actor systems using traces, since most times, we only have a finite number of traces to examine, making the process easy to automate.

However, recursive systems do not fall under this category, as they generate an infinite number of traces. We therefore construct a new method of relating actors based on coinductive techniques. This allows us to reason about and compare the states reached by each system, with the intuition that if two systems can reach the same states, then they must be related. Consequently, when dealing with recursive systems, once a state has been examined, we do not have to examine its traces again if it is looped back to.

For example, suppose system A starts at state **X**, and makes a transition *a* to reach state **Y**, and loops on the state **Y** with the transition *b*. Suppose we also have an actor system B, starting at state **P** which can make the same transitions (possibly others too). Then we can say that A and B are related by the proposed coinductive technique. This is shown in Figure 1.

To evaluate this technique, ERLANG was chosen to reason about the actor model itself [4], [6]. Using a formalisation of the language, our proposed technique is finally shown to be sound with respect to the trace-based technique through formal proofs. Hence, it is also sound with respect to the original safety testing technique.

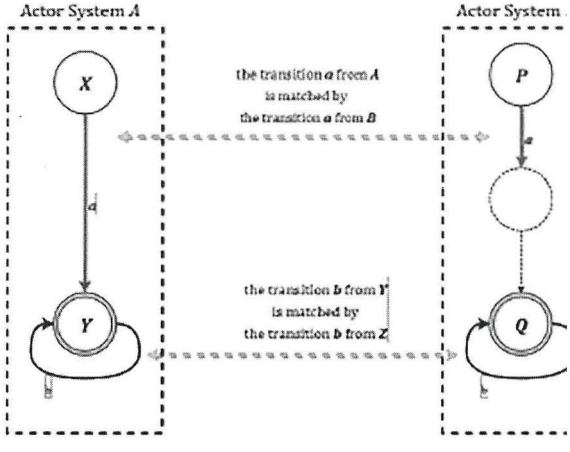


Figure 1: Demonstrating how Actor Systems A and B are related using the proposed technique

REFERENCES

[1] P. Haller and F. Sommers Actors in Scala - *Concurrent programming for the multi-core era*. Artima Press, California, 2011.

[2] E. Tanti *Towards Sound Refactoring in Erlang*. M.Sc Thesis, 2014.

[3] G.A. Agha *Actors: A Model of Concurrent Computation in Distributed Systems*. PhD Thesis, 1985-86.

[4] J. Armstrong *Programming Erlang: Software for a Concurrent World*. Pragmatic Bookshelf, 2007.

[5] C. Hewitt *Description and Theoretical Analysis (Using Schemata) of PLANNER: A Language for Proving Theorems and Manipulating Models in a Robot*. PhD Thesis, 1971.

[6] S.Vinoski, Concurrency with ERLANG, *IEEE* Internet Computing, vol. 11 issue 5, pp. 90 93, Sep-Oct, 2007.

A Hybrid Smartphone-Cloud based Rendering Algorithm for Natural Video Scenes on Mobile Device

JAMES CARUANA

SUPERVISED BY
PROF. ING. CARL JAMES DEBONO

Free-Viewpoint television is one of the most complex applications that requires a substantial amount of processing power [1]. This application requires a set of depth and texture camera views that will be used to create a virtual viewpoint for user consumption. The Free-Viewpoint algorithm can be broken down into a number of discrete tasks. These tasks can then be distributed between the cloud and mobile device in an optimal manner to benefit from the processing capabilities of both the cloud and the mobile device. While the cloud can easily scale in terms of processing capacity, the mobile device is limited to its on-board processing capability and battery capacity [2]. Although mobile device technology is constantly evolving with ever increasing processing capabilities being introduced with each new generation of devices, consideration needs to be taken to a wide variety of Mobile devices already in the market whose capabilities vary significantly. These constraints all need to be taken into account in the mobile side implementation. One further important consideration is dependant on the quality of the network that provides the connectivity between the cloud and the mobile devices. Mobile and WIFI networks are also evolving at a fast pace offering ever increasing speeds and low latency, however wireless networks are shared media which vary significantly in performance depending on the radio conditions and loading.

After analysing the Free-Viewpoint algorithm and its stages it was decided to implement a system where the cloud executes the warping of the left and right frames, and the generation of a blended image. This image is then compressed using a lossy compression algorithm and is sent to the mobile device over a TCP socket on request. Once the mobile device receives the view it performs in-painting to fill the holes and displays the image.

The implemented hybrid system generated the blended view, using OpenMP parallelisation libraries, in 0.7891 seconds on the cloud side, but further improvement could be done. Two in-painting techniques were implemented and tested on the mobile device in order to assess the effectiveness of the algorithms and their performance in terms of speed of execution. Both algorithms performed in-painting effectively however there was a significant difference in the time taken to process a frame. The natively implemented in-painting technique took on average 1.6 seconds on the latest mobile device while the in-built in-painting function provided with the OpenCV libraries took on average 0.089 second, on the same device.

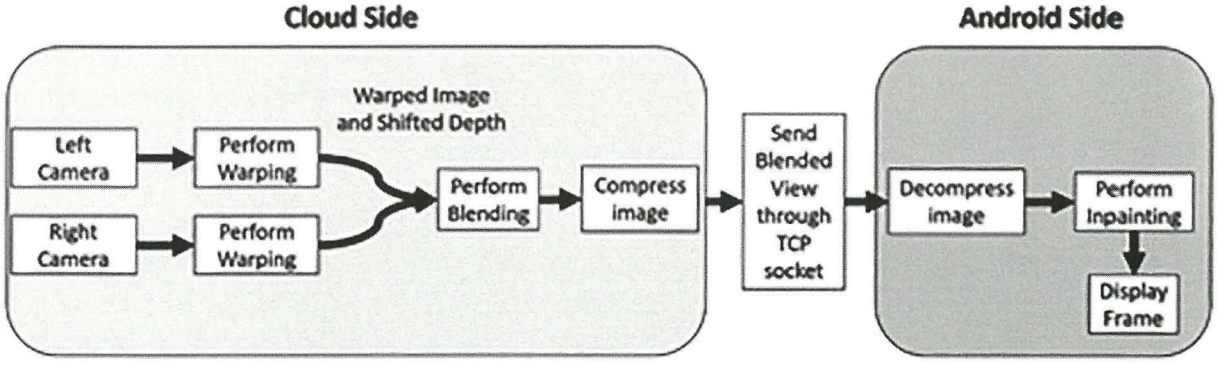


Figure 1: Free-Viewpoint Algorithm System Design

REFERENCES

[1] Masayuki Tanimoto. Overview of free viewpoint television. *Signal Processing: Image Communication*, 21(6):454-461, 2006.

[2] Dan Miao, Wenwu Zhu, Chong Luo, and ChangWen Chen. Resource allocation for cloud-based free viewpoint video rendering for mobile phones. In *Proceedings of the 19th ACM international conference on Multimedia*, pages 1237-1240. ACM, 2011.

PROJECT NO.
59

Implementing an e-Health Monitoring and Supporting System

JULIA CARUANA

SUPERVISED BY
MR ANTHONY SPITERI STAINES

In today's fast paced world, where time is precious, people, the working class in particu-lar, spends the majority of the day shuttling between several tasks and therefore tend to neglect their health and fitness. Even a simple appointment with our doctor in a clinic can take a lot of time, since it can consists of various tests for diagnosis, prescription and finally treatment. Therefore, many patients only seek medical attention and go to a clinic when they are in great pain or suffering from a serious illness. Hence, many individuals are now seeking for an alternative, such as an obtrusive device that can be worn on the body, which constantly monitor the users' health in real time while also providing timely insights on several health parameters to the users and their physicians.

Moreover, the overcrowded hospitals are resulting in a push towards an outpatients or home based treatment of some forms of situations, such as elderly patients. This is where health wearables can help monitoring patients from the comfort of their home and therefore, there is no more need of regular check up appointments.

Nowadays, everyone wants to be a touch away from information access, especially per-sonal information. Our personal health is an important aspect in our lives, thus having a system which will support and monitor our health, and can be accessed via web is of ma-jor importance. People can especially benefit of such system if the information can be accessed on their mobile devices appropriately.

Figure 1: Photo showing the wearable testing prototype

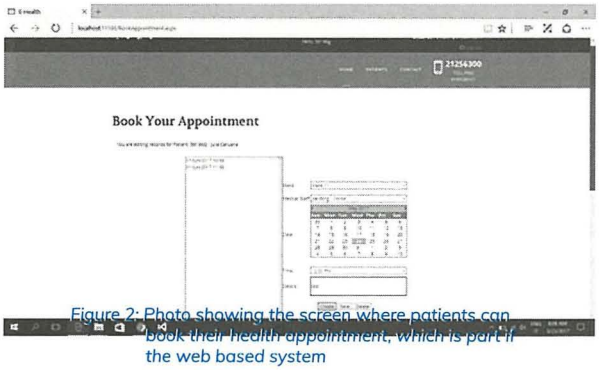


Figure 2: Photo showing the screen where patients can book their health appointment, which is part if the web based system

REFERENCES

[1] K. E. Pasierb, T. Kajdanowicz, and P. Kazienko. "Current Trends and Difficulties in Knowledge-Based e-health Systems," Communications in Computer and Information Science, vol. 278, pp.135-140, Sept 2011.

[2] D. L. Larkai and W. Ruiheng, "Wireless Heart Rate Monitor in Personal Emergency Response System," in Design and Diagnostics of Electronic Circuits &Systems (DDECS), 2015 IEEE 18th International Symposium on, 2015, pp.299-300.

[3] A. Gund. "On the Design and Evaluation of an eHealth System for Management of Pa-tients in Out-of-Hospital Care". Göteborg, Sweden, 2011.

PROJECT NO. 49

Characterizing the Effect of Covariates on the Rate of Admission and Duration of Stay in the Hospital

ROBERTA CARUANA

SUPERVISED BY
DR LALIT GARG

Forecasting the daily resource requirements is important to provide a way for healthcare planners to ensure optimal allocation of scarce resources which in turn could help the problem of long waiting lists [1]. Hospital admissions and patient Length of Stay (LOS) have an influence on the availability of healthcare resources. Several factors may affect a patients' LOS and admission patterns, where the factors considered in this study include: patient characteristics, weather, air pollution and the day of the week. The study was based on hospital admissions data obtained from Mater Dei Hospital, Malta for the period between January 2011 to December 2015, while weather and pollution covariates were obtained from Free Meteo and Environment & Resource Authority (ERA) for the respective period.

Within this study, Survival Trees were used to cluster patients' based on LOS and admissions data. Phase-Type Survival Trees and Gaussian Mixture Survival Trees were modelled using Coxian Phase-Type Distribution (C-PHD) and Gaussian Mixture Distribution (GMD) where the Weighted-average Information Criterion (WIC) statistic was computed as a likelihood parameter. In the case of C-PHD, patient flow was modelled as a Markov Process where patients move progressively between hospital phases allowing them to leave the system at any point, due to discharge or death. The GMD distribution modelled patients in the care system based on the mean and variance. The models created aimed to improve within node homogeneity by selecting covariates providing the most prognostic significance (highest gain in WIC) [2]. The nodes yielding most improvement were located in the higher levels of the Survival Tree. Survival Trees were generated by recursively partitioning the covariates based on the value of WIC until no further improvement is observed; this is the stopping criteria defined.

Furthermore, the models' Goodness-Of-Fit (GOF) was evaluated based on the WIC statistic [3], where C-PHD yielded best improvement for both LOS and admission models, thus it was expected to provide more accurate results than the GMD model. Additionally, the variability in data was analysed which illustrated that C-PHD in the case of the LOS model and GMD in the case of the admissions model, were the distributions that explained variability best. Further to this, Cumulative Distribution Function and Empirical Cumulative Distribution Function plots [4] were used to illustrate how well the distribution values represent empirical data. These plots concluded that C-PHD and GMD could be used effectively to characterize the effect of different factors on the patient LOS and admission patterns. Within this project, the relationship between patient characteristics, weather and pollution data was modelled to patient LOS and admissions patterns to identify the factors effecting the increase in admission and bed occupancy rates. Developing such forecasting models could allow healthcare planners to have a better indication on the expected number of available beds and resources required, which leads to more adequate planning of hospital resources.

REFERENCES

[1] L. Garg, S. McClean, M. Barton, B. Meenan and K. Fullerton, "Intelligent patient management and resource planning for complex, heterogenous and stochastic healthcare systems," *IEEE Trans. Syst. Man Cybern A. Syst. Humans*, vol. 42, no. 6, pp. 1332-1345, 2012.

[2] L. Garg, S. McClean, B. Meenan and P. Millart, "Phase-type survival trees and mixed distribution survival tree for clustering patients' hospital length of stay," *Informatica*, vol. 13, no. 2, pp. 57-72, 2011.

[3] T. Wu and A. Supulveda, "The weighted average information criterion for order selection in time series and regression models," *Stat. and Probability Letters*, vol. 39, no. 1, pp. 1-10, 1998.

[4] R. P, L. Bodrog and A. Danikina, "Phase-type distributions," *Resilience Assesment and Evalutaion of Computing Systems* , pp. 84-114, 2012.

PROJECT NO. 50

Building Machine Learning Models to Explore Protein-Ligand Interactions for Drug Discovery

DANIEL CATANIA

SUPERVISED BY
DR JEAN-PAUL EBEJER

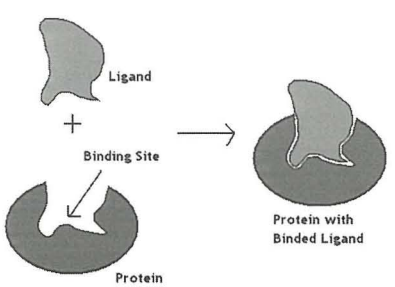
Computational methods have become increasingly popular in drug discovery. In recent years, such methods have been used to aid physical experiments during the initial stages of the drug discovery process, in order to reduce expense and the time taken to run these experiments. Despite the usage of these methods, over recent years we have witnessed only a slight increase in new medicinal drugs approved by the FDA (U.S. Food and Drug Administration). Approximately 29 new molecular entities are approved annually [1, 2].

The aim of this final year project was to gain a better understanding of protein-ligand interactions, and to build machine learning models based on this understanding. These computational models help us identify small-molecules (also known as ligands) which interact with proteins, a necessary requirement for most medicinal drugs, as shown in Figure 1a. These small-molecules interact, or bind, with proteins at some strength (binding affinity). The binding affinity is measured experimentally in wet-laboratories using three alternate measures (IC_{50}, K_i and K_d). These measures are not easily convertible from one to another (without knowing the experiment's parameters). An example of a protein-ligand interaction is aspirin binding to a human protein (shown in Figure 1b).

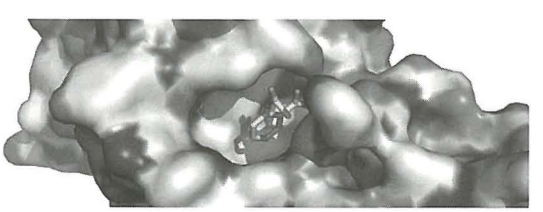
Presently, the only existing database of protein-ligand interactions, CREDO, is not publicly available. Our database was built by first acquiring 3D protein structures from the Protein Data Bank and filtering them to keep those which contained small molecules. The small-molecules and their nearby atoms on the protein were then extracted and standardised to correct inconsistencies. The small-molecules were then checked for their drug-likeness by applying Lipinski's rule of five (these are filters which distinguish drug-like from non-drug-like molecules). Thereafter, the pharmacophores (features) of these small-molecules and their nearby atoms were extracted. Finally, the interacting feature counts were calculated for every structure and stored. This database was then augmented with the experimental binding affinity data obtained from the PDBbind database. The binding affinity measures were converted to nanomolar (nM / 10^{-9}M), as these ranged from molar (M) to femtomolar (fM / 10^{-15}M). Machine learning models were then built to predict the binding affinities based on interactions between a protein and a ligand. Several machine learning approaches were explored, including Nearest Neighbours, Support Vector Machines, Random Forest and Artificial Neural Networks to find the best model which describes this binding relationship.

While acknowledging that predicting binding affinity based on only feature interactions is hard, we also concluded that the model selection depended on the type of binding affinity measure. The K_i experimental measure was the best measure from the three experimental measures of the binding affinity, as a Support Vector Machine with a Linear Kernel produced an average accuracy of 56% over 20 prediction classes.

The new database of protein-ligand interactions together with their respective binding affinity data have been made publicly available on Github [3] and may allow for further research in this area.



(a) A schematic diagram showing how a Ligand binds to a Protein at the Binding Site.



(b) Aspirin (shown in Green) binding to a Human Protein (which surface is shown in White/Red/Blue).

Figure 1: Protein-Ligand Interactions

REFERENCES

[1] Z. Brennan. New fda drug approvals: Breaking down the numbers. http://raps.org/Regulatory-Focus/News/2017/01/04/26501/New-FDA-Drug-Approvals-Breaking-Down-the-Numbers. Last Accessed: 2017-05-24.

[2] FDA. New molecular entity and new therapeutic biological product approvals. http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugInnovation/default.htm. Last Accessed: 2017-05-24.

[3] Daniel Catania. Github repository - protliginteractiondb. https://github.com/dcatania318/ProtLigInteractionDB, 2017.

PROJECT NO. 71

Using Automatic Greyscale Image Colourisation As a Compression Technique

KRISTINA CATANIA

SUPERVISED BY
MR KRISTIAN GUILLAUMIER

This project is about image compression. We make use of a compression algorithm applied to a greyscale image, and an automatic colourisation algorithm to regenerate the colour. Our aim is to test whether we are able to produce a comparable image to JPEG's. Specifically we expect our image to have a smaller file size at the expense of colour fidelity. In this project, our modified version of JPEG (Joint Photographic Experts Group) image compression is applied to the image to reduce data by removing redundant information. This allows data to be stored or transmitted in a more efficient form due to its smaller size. For our purpose, we will be storing the luminance values only as opposed to all the image data. Image colourisation is then applied. This is a process by which colour is added to greyscale images. Colour data points are extracted from the original image before compression is applied. This is done by segmenting the image where data points are then selected from each segment. Random data points are also added to further enhance colourisation. By storing only a few colour data points rather than colour data for each pixels we aim to reduce the file size even further. Our hypothesis states that neighbouring pixels with similar luminance are likely to have similar chrominance values. This is the underlying rationale supporting our automatic colourisation hypothesis. Another hypothesis is that human eyes are less sensitive to colour variance then they are to luminosity. For this reason, slight variations in colour may be unnoticed, especially if the original raw image remains unseen. Our results show that using our algorithms we were able to produce comparable images to JPEG's. However, our file size did not reach the required size reduction due to the amount of colour data points needed in order to recolour the image. Although slight imperfections can be found in the resultant image, such defects may remain unnoticed on small devices, such as mobile phones. Finally, given that our projections demonstrate comparable results to JPEG's, this suggests that further research would result in better results.

Figure 1: Left image shows the result produced by JPEG at quality factor 50. Right image shows our results after compression at quality factor 50 and automatic colourisation.

REFERENCES

[1] A. Sreedharan, "Jpeg 101 - how does it work?." Available at: http://arjunsreedharan.org/post/146070390717/jpeg-101-how-does-jpeg-work.
[2] A. Levin, D. Lischinski, and Y. Weiss, "Colorization using optimization," ACM Transactions on Graphics, 2004.
[3] C. Ebejer, "Chroma lab colourisation of grayscale images and videos," June 2005.
[4] "What is color space?." Available at: http://www.arcsoft.com/topics/photostudio-darkroom/what-is-color- space.html.
[5] "Definition of ycbcr." Available at: http://www.pcmag.com/encyclopedia/term/55147/ycbcr.
[6] "Ycbcr color space-an intro and its applications." Available at: http://www.roman10.net/2011/08/18/ycbcr-color-spacean-intro-and-its- applications/, August 2011.
[7] J. Vargas, "Hsl, hsb and hsv color: differences and conversion." Available at: http://codeitdown.com/hsl-hsb-hsv-color/, October 2013.
[8] J. W. Shipman, "The hue-saturation-value (hsv) color model." Available at: http://infohost.nmt.edu/tcc/help/pubs/colortheory/web/hsv.html, October 2012.
[9] "Grayscale images." Available at: http://homepages.inf.ed.ac.uk/rbf/HIPR2/gryimage.htm.

[10] "Whatis.com : grayscale." Available at: http://whatis.techtarget.com/definition/grayscale.
[11] "Lossy compression." Available at: https://cs.stanford.edu/people/eroberts/courses/soco/projects/data- compression/lossy/index.htm.
[12] "Lossless compression: An overview." Available at: https://cs.stanford.edu/people/eroberts/courses/soco/projects/data- compression/lossless/index.htm.
[13] "Lossy vs lossless compression." Available at: https://optimus.keycdn.com/support/lossy-vs-lossless/, July 2016.
[14] S. Bradley, "Why luminance is the key component of color." Available at: http://vanseodesign.com/web-design/color-luminance/, December 2014.
[15] M. Archive, "How does jpeg actually work?." Available at: https://blogs.msdn.microsoft.com/devdev/2006/04/12/how-does-jpeg-actually- work/, April 2006.
[16] D. Austin, "Image compression: Seeing what's not there." Available at: http://www.ams.org/samplings/feature-column/fcarc-image-compression.
[17] V. Sanchez and A. Basu, "The jpeg2000 image compression standard." Available at: https://webdocs.cs.ualberta.ca/anup/Courses/604/NOTES/slide jpeg2000.pdf.
[18] M. Archambault, "Jpeg 2000: The better alternative to jpeg that never made it big." Available at: https://petapixel.com/2015/09/12/jpeg-2000-the-better-alternative-to-jpeg-that-never-made-it-big/, September 2015.
[19] A. Jakulin, "Baseline jpeg and jpeg2000 artifacts illustrated." Available at: http://www.stat.columbia.edu/jakulin/jpeg/artifacts.htm.
[20] S. Chastain, "Jpeg artifacts." Available at: https://www.thoughtco.com/jpeg-artifacts-1701754, November 2015.
[21] "Colorization." Available at: https://www.cs.unc.edu/lazebnik/research/fall08/lec06 colorization.pdf.
[22] E. Reinhard, M. Ashikhmin, B. Gooch, and P. Shirley, "Color transfer between images," IEEE Computer Graphics and Applications, 2001.
[23] T. Welsh, M. Ashikhmin, and K. Mueller, "Transferring color to greyscale images," ACM Transactions on Graphics, 2002.
[24] R. Irony, D. Cohen-Or, and D. Lischinski, "Colorization by example," 2005.
[25] M. S. AL-Ani and F. H. Awad, "The jpeg image compression algorithm," International Journal of Advances in Engineering and Technology, May 2013.
[26] M. B. Martin and A. E. Bell, "New image compression techniques using multiwavelets and multiwavelet packets," IEEE Transactions on Image Processing, vol. 10, April 2001.
[27] D. Taubman, "High performance scalable image compression with ebcot," IEEE Transactions on Image Processing, vol. 9, July 2000.
[28] L. Cheng and S. V. N. Vishwanathan, "Learning to compress images and videos," 2007.
[29] S. Ono, T. Miyata, and Y. Sakai, "Colorization-based coding by focusing on characteristics of colorization bases," 2010.
[30] H. Noda, N. Takao, and M. Niimi, "Colorization in ycbcr color space and its application to jpeg images," Pattern Recognition, 2007.
[31] T. Horiuchi and S. Tominaga, "Color image coding by colorization approach," EURASIP Journal on Image and Video Processing, 2008.
[32] K. Pullicino, "Image colourisation as a compression techniques," May 2013.
[33] R. Obryk and J. Alakuijala, "Announcing guetzli: A new open source jpeg encoder." Available at: https://research.googleblog.com/2017/03/announcing-guetzli-new-open-source-jpeg.html, Marc 2017.
[34] "Perceptual jpeg encoder." Available at: https://github.com/google/guetzli.
[35] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Transactions on Image Processing, 2004.
[36] A. Hor´e and D. Ziou, "Image quality metrics: Psnr vs. ssim," 2010 International Conference on Pattern Recognition, 2010.

PROJECT NO. 24

Investigating HCI Techniques to Support Mentors Working with Autistic Children

DYLAN CORTIS

SUPERVISED BY
DR CHRIS PORTER AND DR PAUL A. BARTOLO

In this exploratory work, we investigate different techniques used by speech & language pathologists to help children on the autism spectrum improve their communication skills [1]. With ASD-diagnosed children, challenges emerge when using technology since they have less opportunities to gain the communication partner's attention [2]. Key behaviors necessary for conversation, such as maintaining eye contact, persistence and undistracted observation of surroundings are typically not required when interacting with a device [3].

This study hypothesized that there may exist an interactive technology which can be streamlined for speech and language pathologists, to reap the benefits of using technological devices (i.e. good motivators, attention grabbing), while avoiding drawbacks typically associated with them (e.g. undistracted observation of surroundings) – to encourage students to communicate by interacting with the real world as part of their learning process.

Initially, semi-structured interviews with domain experts were carried out to understand the main issues encountered in practice. Braun and Clarke's Thematic Analysis technique [4] was used to structure insights arising from these interviews. Following this, direct observations of communication therapy sessions further helped to understand what this interactive technology should do. A scoping exercise highlighted a frequent challenge encountered by pathologists, which is the teaching of preverbal skills (e.g. matching words/visuals to real life objects). This, together with issues such as the children's proneness to frustration (hence requiring damage-resistant technology) and clutter reduction (hence reducing the number of peripherals and cables) informed the development of an artifact which was subsequently evaluated and further developed through iterative field trials.

The interactive system created makes use of NFC tags and an Android application - Interactiverb. Verbs are prompted on screen by the application, and the child is encouraged to choose the most appropriately related object. Interaction with the real world is required since the application will only confirm the child's decision once the device is physically close to one of the related objects. (e.g. verb: eat > object: apple).

Through feedback collected from the field trials and co-design interactions with speech & language pathologists', this study confirmed the system's success in satisfying the hypothesis. Children with different levels of severity responded favorably to it, with higher-functioning children wanting more challenges and lower-functioning children displaying advanced levels of questioning because of it. A System Usability Scale survey passed to expert participants got an overall score of 90, which deems the system highly recommendable.

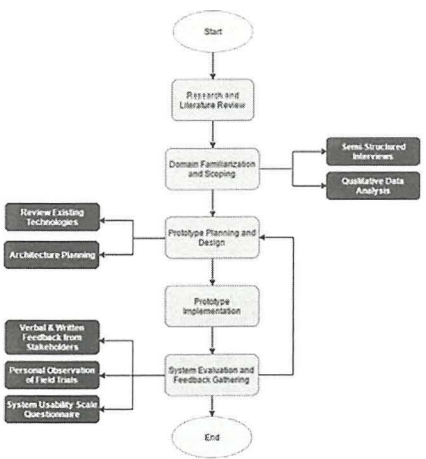


Figure 1: Setup of final artifact

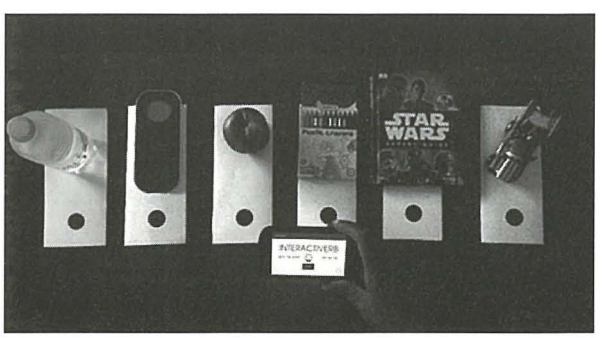


Figure 2: Flowchart of the process followed and techniques used

REFERENCES

[1] A. P. Association, Diagnostic and statistical manual of mental disorders, 5th edition (DSM-5), 5th ed. Washington, DC, United States: American Psychiatric Publishing, 2013.

[2] S. Son, J. Sigafoos, M. O'Reilly and G. Lancioni, "Comparing two types of augmentative and alternative communication systems for children with autism", Pediatric Rehabilitation, vol. 9, no. 4, pp. 389-395, 2006.

[3] M. Flores, K. Musgrove, S. Renner, V. Hinton, S. Strozier, S. Franklin and D. Hil, "A Comparison of Communication Using the Apple iPad and a Picture-based System", Augmentative and Alternative Communication, vol. 28, no. 2, pp. 74-84, 2012.

[4] V. Clarke and V. Braun, "Using thematic analysis in psychology," Qualitative Research in Psychology, vol. 3, no. 2, pp. 77-101, 2006

PROJECT NO.
51

Non-Deterministic Contract Automata

ANDRÉ CROUCHER

SUPERVISED BY
PROF. GORDON J. PACE

Contracts provide a legal binding between two or more parties and used when they are en- gaging in operations in order to obtain some common agreement. The specification of deontic norms [10] such as obligations and permissions allow for this interaction between communicating parties. At a specific point in a contract, a party may be given a number of options which it should do and this is when choices in contracts arise.

Despite the existence of several logics used to represent contracts we focus our attention on contract automata [5], which are a formalism used to represent and reason about contracts. They provide an action-based view of contracts between parties with deontic clauses tagged in each state and allow for the representation of a very important feature, that of interaction between participating parties. However, they lack expressivity when it comes to the notion of choice — there currently exists no explicit way of knowing which party is deciding to make a choice in a contract.

We provide two extensions to the definitions of contract automata in order to handle choices in contracts. Our first extension consists of mapping every action set to the participating party deciding to make the choice whilst the second extension consists of distinguishing between deterministic (where no choice is available) and choice transitions, represented using forking actions. The representation of choice of these two extensions may be seen in Figure 1 and Figure 2 respectively, below. In the figures, 'P' denotes the party making the choice, upper case letters (A, B) denote the set of actions being made whilst lower case letters (a, b, c, d) denote the deontic modality (obligation, permission, prohibition) that has to be obeyed by a party, in a particular state.

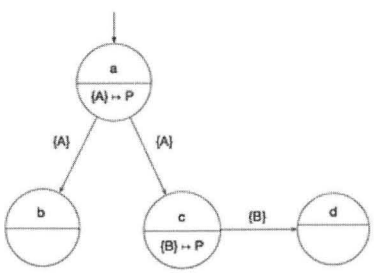


Figure 1: First extension to represent explicit choices in contracts

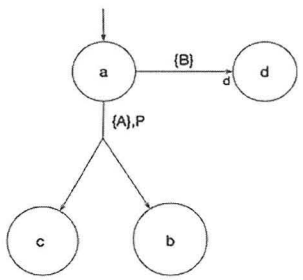


Figure 2: Second extension to represent explicit choices in contracts

Both extensions, despite having some differences, provide similar expressivity when rep- resenting choices explicitly, removing any element of ambiguity. By proving about several propositions we check the sanity of our formalisms and by also providing a use case between a student and the university, based on the assessment regulations of the University of Malta [1], we demonstrate the effectiveness of our extensions and the way they are implemented.

Our work adds to the expressivity of contract automata by being able to handle external choices, explicitly stating which party is making the choice. This adds a new dimension to contract representation when using contract automata and provides a basis for future research.

REFERENCES

[1] U. of Malta, "University assessment regulations," https://www.um.edu.mt/__data/assets/ pdf_file/0010/24868/Assessment_Regulations.pdf, 2009.

[2] C. Prisacariu and G. Schneider, "A formal language for electronic contracts," in Formal Methods for Open Object-Based Distributed Systems, 9th IFIP WG 6.1 International Conference, FMOODS 2007, Paphos, Cyprus, June 6-8, 2007, Proceedings, ser. Lecture Notes in Computer Science, M. M. Bonsangue and E. B. Johnsen, Eds., vol. 4468. Springer, 2007, pp. 174–189. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-72952-5_11

[3] G. J. Pace and A. P. Ravn, Eds., Proceedings Sixth Workshop on Formal Languages and Analysis of Contract-Oriented Software, FLACOS 2012, Bertinoro, Italy, 19 September 2012, ser. EPTCS, vol. 94, 2012. [Online]. Available: http://dx.doi.org/10.4204/EPTCS.94

[4] G. J. Pace and F. Schapachnik, "Contracts for interacting two-party systems," in Proceedings Sixth Workshop on Formal Languages and Analysis of Contract-Oriented Software, FLACOS 2012, Bertinoro, Italy, 19 September 2012., ser. EPTCS, G. J. Pace and A. P. Ravn, Eds., vol. 94, 2012, pp. 21–30. [Online]. Available: http://dx.doi.org/10.4204/EPTCS.94.3

[5] S. Azzopardi, G. J. Pace, F. Schapachnik, and G. Schneider, "Contract automata — an operational view of contracts between interactive parties," Artif. Intell. Law, vol. 24, no. 3, pp. 203–243, 2016. [Online]. Available: http://dx.doi.org/10.1007/s10506-016-9185-2

[6] E. Martínez, G. Díaz, M. Cambronero, and G. Schneider, "A model for visual specification of e-contracts," in 2010 IEEE International Conference on Services Computing, SCC 2010, Miami, Florida, USA, July 5-10, 2010. IEEE Computer Society, 2010, pp. 1–8. [Online]. Available: http://dx.doi.org/10.1109/SCC.2010.32

[7] S. Fenech, G. J. Pace, and G. Schneider, "Automatic conflict detection on contracts," in Theoretical Aspects of Computing - ICTAC 2009, 6th International Colloquium, Kuala Lumpur, Malaysia, August 16-20, 2009. Proceedings, ser. Lecture Notes in Computer Science, M. Leucker and C. Morgan, Eds., vol. 5684. Springer, 2009, pp. 200–214. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-03466-4_13

[8] S. L. P. Jones, J. Eber, and J. Seward, "Composing contracts: an adventure in financial engineering, functional pearl," in Proceedings of the Fifth ACM SIGPLAN International Conference on Functional Programming (ICFP '00), Montreal, Canada, September 18-21, 2000., M. Odersky and P. Wadler, Eds. ACM, 2000, pp. 280–292. [Online]. Available: http://doi.acm.org/10.1145/351240.351267

[9] S. Fenech, G. J. Pace, J. C. Okika, A. P. Ravn, and G. Schneider, "On the specification of full contracts," Electr. Notes Theor. Comput. Sci., vol. 253, no. 1, pp. 39–55, 2009. [Online]. Available: http://dx.doi.org/10.1016/j.entcs.2009.09.027

[10] G. H. von Wright, Norm and Action: A Logical Enquiry. Routledge and Kegan Paul, 1963.

[11] "Deontic logic," Mind, vol. 60, no. 237, pp. 1–15, 1951.

Dual-Axis Light Tracker

PROJECT NO. 2

MARK DALLI

SUPERVISED BY
DR ING. OWEN CASHA

With an ever growing dependence on technology and intelligent systems, automatic control and microcontroller based systems play an increasingly important role in our daily lives. Feedback control systems have been extremely popular over the past few decades, providing elegant solutions to complicated control problems in applications ranging from climate control systems in automobiles to high precision systems in space vehicle control.

This project involved the design and implementation of a dual-axis light tracker which is able to locate a source of light in both the panning and tilting axis of rotation. A block diagram of the implementation is presented in **Figure 1.** Four light sensors were strategically placed around a cross-shaped shade, in such a way that all the sensors receive an equal light level when the system directly faces the light source (refer to **Figure 2)**. A digital control algorithm based on proportional, integral and derivative control was implemented on an Arduino microcontroller driving two servo motors, one for each axis of rotation.

The system also contains a number of of additional key features to improve its performance, amongst which are gain scheduling and integral anti-windup. Gain scheduling is a technique used in non linear systems which require different control settings depending on the current conditions of the external environment. In the case of the dual axis light tracker, the tuning parameters of the controller need to be selected based on the ambient light intensity. A set of parameters was determined experimentally for a wide range of different light intensity and set automatically on the fly. Integral anti-windup techniques were also explored. The introduction of integral control can often result in a phenomenon called integral windup, whereby the system shows unstable response or lag. This is often due to the physical limitations of the system, preventing it from reaching its desired final point. A number of conditions need to be imposed on the integral action depending on the specific physical limitation of the system. The results were compared to a simple yet widely used algorithm for light tracking. This system is based on the idea of measuring the intensity of light in two separate regions and orienting the system towards the higher light intensity region. The system stops when the two readings are equal, indicating that the light source has been located.

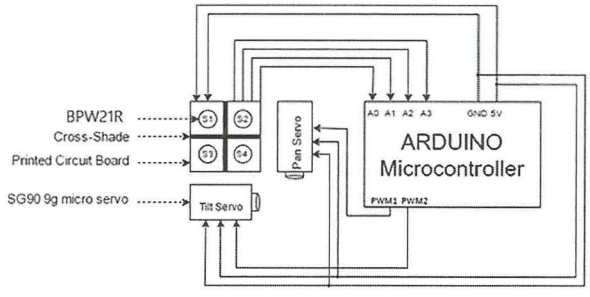


Figure 1: Block diagram of the dual-axis light tracker

Figure 2: Implementation of the dual-axis light tracker

REFERENCES

[1] K. AAstrom and T. Hagglund, *Advanced PID Control*. ISA-The Instrumentation, Systems, and Automation Society, 2006.

[2] K. Ogata, *Modern Control Engineering*, ser. Instrumentation and controls series. Prentice Hall, 2010.

[3] F. Golnaraghi and B. Kuo, *Automatic Control Systems*, Tenth Edition. McGraw-Hill Education, 2017.

[4] J.-M. Wang and C.-L. Lu, *Design and implementation of a sun tracker with a dual axis single motor for an optical sensor-based photovoltaic system*, Sensors, vol. 13, no. 3, pp. 3157–3168, 2013.

PROJECT NO. 26

Business Analysis Tool for Small/Medium Enterprises using Spatial Databases

KURT DEBONO

SUPERVISED BY
MR MICHEL CAMILLERI

All businesses strive to compete with one another and try to find methods to improve on their sales with each year. This can be done using various different methods, starting from diverse marketing techniques to strategic business decisions. One important factor that can significantly affect their sales is their geographic location (Ambarwati & Hanum, n.d.). Small and Medium Enterprises (SMEs), however, due to the way that they are built and operated, often do not possess sufficient knowledge, data and even financial resources, to effectively exploit this component in their business model (Gourova, 2015).

This study sought to investigate the usefulness to and usability by local SMEs of a specifically developed Business Analysis tool that uses location-based data. This included a study of its ability to maximize the efficiency of part of their day-to-day operations and to support their strategic business decision making processes. Different visualizations were implemented to present various types of information relevant to the user (Figures 1 and 2). These visualizations were designed to be easily understood and interpreted by these SMEs whilst requiring the least use of resources by the same SMEs.

A questionnaire was distributed to a set of local SMEs, asking them about their views and opinions about the role such a tool could play in an SME environment. The responses indicated the type of functionality that is relevant to such SMEs and influenced the design and development of the prototype tool. During the testing of the prototype, the accuracy of the predictive function of the model system was estimated through specific experiments and data analysis.

The eventual prototype system was evaluated by participants, similar to the ones who participated in the first questionnaire, but this time through an interview. The results of this evaluation gave an indication of whether the functionality implemented in the prototype tool would be of use in a real world scenario and highlighted the general usability and usefulness of the prototype.

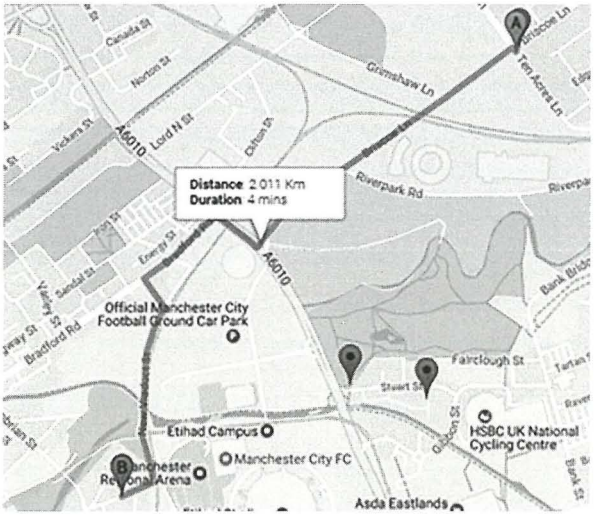


Figure 1: Visualization Example - Routing

Figure 2: Visualization Example - Prediction

REFERENCES

[1] Ambarwati, L., & Hanum, Y. (n.d.). Web Based Geographic Information System for the Spread of Small and Medium Enterprises across Jakarta.

[2] Gourova, E. (2015). Knowledge management strategy for Small and Medium Enterprises. *International Conference on Applied Computer Science Knowledge*, (October), 639–648.

PROJECT NO. 73

Towards Bookmark Folder Recommendation Using Simulated User-Guided Classification, and Clustering Techniques

NIKOLAI DEBONO

SUPERVISED BY
DR CHRIS STAFF AND DR COLIN LAYFIELD

Webpage bookmarking involves users storing their favourite webpages into an organised structure called a bookmark hierarchy or repository. We developed a user-guided bookmark folder recommendation system to aid users in maintaining their repository's organisation.

Our system builds a flat-cluster semantic model of the hierarchy using Latent Semantic Analysis [1]. This evolves alongside users through their Indirect guidance whenever a bookmark is created. Clusters model folders containing pages using a mean centroid representation. If the repository is small or empty, LSA is performed on the user's history instead. Using the model's centroids, it recommends up to three folders most similar to the page. If no folder's similarity is above a threshold, the system suggests creating a new one. We also developed a folder restructuring user-tool using DBSCAN [2] for folders containing random webpages or multiple pages from different topics, thus aiding recommendation.

The recommendation system was evaluated using the SemEval 2013 Task 11 corpus [3]. We simulated the creation and evolution of multiple hierarchies, performing recommendation before adding each bookmark. We modified a K-NN [4] algorithm to give up to three recommendations, for system comparisons. We also used a new metric, User Cost Index, that measures recommendation quality with respect to the 'cost' for a user to fix it. The restructuring tool was evaluated using the same corpus.

We reached our aims through a system requiring no external data that creates and maintains an evolving, user-guided model for recommendation. Using LSA we achieved up to 67.6% overall accuracy, which rises to 72.4% when the repository's size increases. We also discovered that using a static LSA parameter on an evolving hierarchy deteriorates accuracy by up to 4.4%. Meanwhile, the folder restructuring evaluation showed that its positive results are not always consistent for word sense based clustering.

These tools were incorporated into a browser addon. This shows that a fully functional web browser addon for intelligent bookmark management is possible.

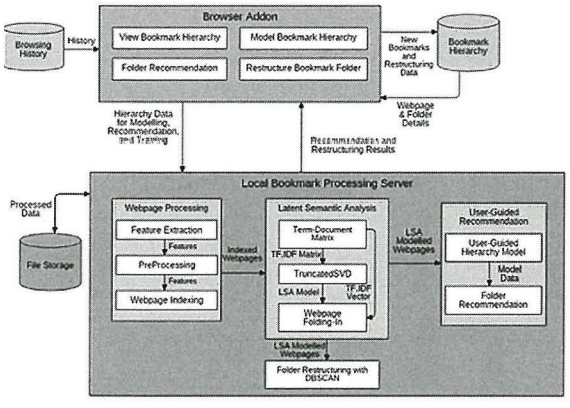


Figure 2: System Diagram

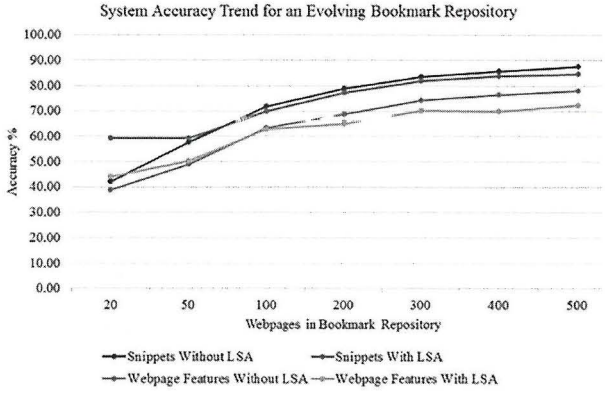


Figure 2: System Accuracy Trend

REFERENCES

[1] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman, "Indexing by latent semantic analysis," *J. Am. Soc. Inf. Sci.*, vol. 41, no. 6, pp. 391–407, Sep. 1990.

[2] M. Ester, H. P. Kriegel, J. Sander, and X. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," *Proc. 2nd Int. Conf. Knowl. Discov. Data Min.*, pp. 226–231, 1996.

[3] "Evaluating Word Sense Induction & Disambiguation within An End-User Application." [Online]. Available: https://www.cs.york.ac.uk/semeval-2013/task11/index.php%3Fid=task-description.html.

[4] S.-H. Myoung and J.-M. Choi, "BClassifier: A Personal Agent for Bookmark Classification."

PROJECT NO. 27

Intelligent Bookmark Organiser and Recommender

MAXIMILIAN DEPASQUALE

SUPERVISED BY
DR COLIN LAYFIELD AND DR JOEL AZZOPARDI

With the increasing use of the World Wide Web in day to day life, it is of growing importance for the average web user to have a convenient web browser setup. As such the continued development of the WWW spurs the development of feature rich tools which have the objective of improving the web browsing experience. A sizeable subset of such tools aim to do so by offering a bookmark organisation service or a bookmark recommendation service. However the great majority of these tools require supervision to operate, for instance most bookmark organisers provide users with an interface to manually go through bookmarks and sort them one by one, but they do not automate this process for them. Thus, this dissertation aims to further exploration into the automation of such tools.

To explore this, an artefact which can organise and recommend bookmarks in an unsupervised manner was developed. Two distinct organisation techniques were explored, one clustering via user defined labels (mimicking folder names) and one clustering without any user input at all, using the Vector Space Model (as shown in Figure 1). Latent Semantic Analysis, a technique for analysing semantic relationships, was also tested with non user labelled clustering. Two distinct recommendation techniques were also developed. The first of these techniques made use of the Vector Space Model to compare the similarity of webpages visited by the user to those in his bookmark list. The second technique used a search engine query approach, where the most semantically significant words in a bookmark folder were extracted and submitted to a search engine as a query, with the top 4 results returned being recommended to the user.

To evaluate the artefact's performance three tests were devised. The first test compared the artefact's clustering algorithms with the top performing algorithm in SemEval-2013 task 11 workshop, with the algorithms used in the artefact producing marginally better results. For the second test, the artefact's bookmark organisation performance was tested on real bookmark lists. The artefact was installed on 10 users' computers. Users were asked to produce a list of how they would manually cluster their bookmarks (without seeing any results produced by the artefact). Then the clusters produced by the artefact were compared to the user produced lists. A simple accuracy score of 90.43% was achieved when using LSA and 88.72% when using labelled clustering, which is significantly higher than accuracy scores of similar bookmark organisational systems. For the third test, users' were asked to mark recommendations made by the artefact as either useful or not useful. In total, 91.78% of recommendations were marked as useful, which was on par with the best performing similar fully automated recommendation system.

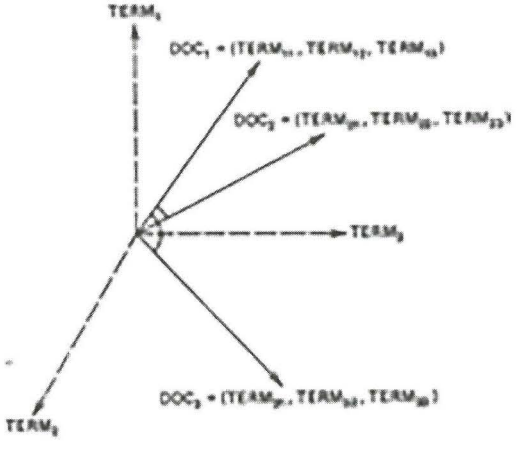


Figure 1: Vector Space Model Representation

Modelling Techniques for Cloud-Based Solutions

GLEB ELISEEV

SUPERVISED BY
PROF. ERNEST CACHIA

Cloud computing is nowadays on everyone's mouth and in every developer's mind [1]. Various solutions were developed or rewritten and ported to the cloud. Even though adoption of Cloud technology and services offers specific benefits to its users, in certain instances it might present unexpected development issues. [2, 3]

Most of the software that has been developed uses modeling techniques which allow one to foresee possible bottlenecks and challenges as well as the selection of tools that would be used to best leverage development [4, 5]. In addition, modelling techniques allow the breakdown of a software solution into its constituent and more detailed parts. Many modelling techniques have been around for quite some time and have a well-known track record [6]. However, in order to create software solutions that make use of modern technologies and approaches, researchers build different extensions to them [7, 8, 9]

At present, Cloud computing is a technological aspect which is not comprehensively supported by software modeling techniques [10, 11]. This spurred the idea of proposing a solution that would allow developers and other relevant solution stakeholders to model software solutions with Cloud technology and services in mind, allowing, amongst other things, the modelling and provisioning of Cloud technology as part of the presented solution.

The purpose of this project is to help developers who are planning to either develop for, or port to, the Cloud, to reduce development costs that may be arise from improper cloud integration and deployment [12, 13]. It is proposed that this be done by introducing appropriate extensions to one of the most popular modelling language nowadays - the Unified Modelling Language (UML) [14, 15].

This project will be tackling the three aspects of the cloud deployment service models, which are Software as a Service, Platform as a Service, and Infrastructure as a Service [2, 12]. The reason for this is that deployment of the application should be as closely mapped to its model so as to safeguard consistency in the later stages of development and to facilitate maintenance of the solution in the future resulting in reduced maintenance costs. Each service model was tackled differently, but each was implemented within the same solution. The proposed solution presents a set of extensions as classes for each of the previously mentioned models as well as providing a new type of relationship to connect the existing classes to the newly introduced ones. The main aim of the project is to reduce both the overall cost of software solutions using Cloud technologies as well as to reduce the risks of additional, and sometimes unjustified, costs linked to incorrect usage of Cloud resources [16, 17].

The validity and outcomes of the newly proposed extension and relationship type are then tested on the case-study by comparing the estimated costs incurred for two scenarios: one using the extension for that given scenario, the other scenario involves using the default features and modelling capabilities offered by the providers, namely Microsoft's Azure and Rackspace [18, 19].

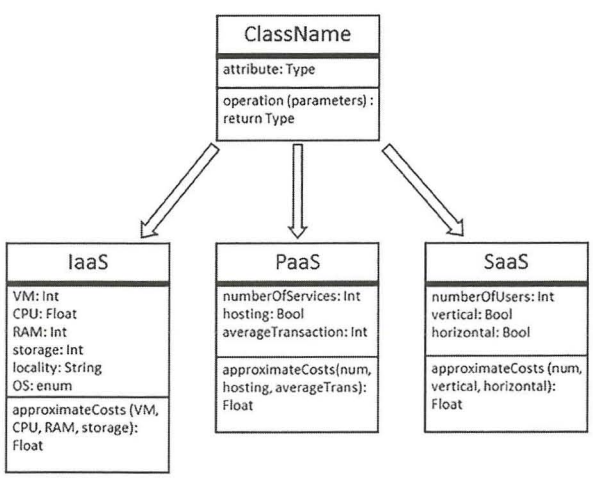


Figure 1

Figure 2

REFERENCES

[1] G. Pallis, "Cloud Computing: The New Frontier of Internet Computing," IEEE Internet Computing, vol. 14, no. 5, pp. 70-73, 2010.

[2] Q. Zhang, L. Cheng and R. Boutaba, "Cloud computing: state-of-the-art and research challenges," Journal of internet services and applications, vol. 1, no. 1, pp. 7-18, 2010.

[3] H. Singh, "Cloud Computing: Issues," 2010.

[4] J. den Haan, "15 reasons why you should start using Model Driven Development," The Enterprise Architect, 25 November 2009. [Online]. Available: http://www.theenterprisearchitect.eu/blog/2009/11/25/15-reasons-why-you-should-start-using-model-driven-development/. [Accessed 04 06 2017].

[5] E. S. K. Yu and J. Mylopoulos, "Understanding "Why" in Software Process Modelling, Analysis, and Design," in 16th International Conference on Software Engineering, Sorrento, 1994.

[6] R. N. Taylor, N. Medvidovic and E. M. Dashofy, "The Big Idea," in Software Architecture: Foundations, Theory, and Practice, Hoboken, John Wiley & Sons, Inc., 2009, pp. 1-23.

[7] N. Koch, H. Baumeister, R. Hennicker and L. Mandel, "Extending UML to Model Navigation and Presentation in Web Application," in Proceedings of Modelling Web Applications in the UML Workshop, York, 2000.

[8] J. Farhad, "The UML Extension Mechanisms," Department of Computer Science, University College London, London, 2002.

[9] B. Bauer, "UML Class Diagrams Revisited in the Context of Agent-Based Systems," in AOSE '01 Revised Papers and Invited Contributions from the Second International Workshop on Agent-Oriented Software Engineering II, Berlin, 2002.

[10] A. Bergmayr, J. Troya, P. Neubauer, M. Wimmer and G. Kappel, "UML-based Cloud Application Modeling with Libraries, Profiles, and Templates," in 2nd International Workshop on Model-Driven Engineering on and for the Cloud, Valencia, 2014.

[11] S. Alajrami, A. Romanovsky, P. Watson and A. Roth, "Towards Cloud-Based Software Process Modelling and Enactment," in 2nd International Workshop on Model-Driven Engineering on and for the Cloud, Valencia, 2014.

[12] T. Dillon, C. Wu and E. Chang, "Cloud Computing: Issues and Challenges," in 24th IEEE International Conference on Advanced Information Networking and Applications, Perth, 2010.

[13] W. Kim, S. D. Kim, E. Lee and S. Lee, "Adoption Issues for Cloud Computing," in 7th International Conference on Advances in Mobile Computing and Multimedia, Kuala Lumpur, 2009.

[14] J. Cheesman and J. Daniels, "Preface," in UML Components, EUA: Addison-Wesley, 2001, p. xiii.

[15] C. F. Lange and M. R. Chaudron, "In Practice: UML Software Architecture and Design Description," IEEE Software, vol. 23, no. 3, pp. 40-46, 2006.

[16] A. Gohad, N. C. Narendra and a. P. Ramachandran, "Cloud Pricing Models: A Survey and Position Paper," in IEEE International Conference on Cloud Computing in Emerging Markets (CCEM), Bangalore, 2013.

[17] D. Bellenger, J. Bertram, A. Budina, A. Koschel, B. Pfänder, C. Serowy, I. Astrova, S. Gatziu Grivas and M. Schaaf, "Scaling in Cloud Environments," in 15th WSEAS International Conference on Computers, Corfu, 2011.

[18] Rackspace, "RACKSPACE CLOUD CALCULATOR," RACKSPACE US, INC., [Online]. Available: https://www.rackspace.com/en-gb/calculator. [Accessed 11 05 2017].

[19] Microsoft, "Pricing calculator," [Online]. Available: https://azure.microsoft.com/en-us/pricing/calculator/. [Accessed 7 05 2017].

What's up? Research

PROJECT NO. 75

ALAN FALZON

SUPERVISED BY
DR JOEL AZZOPARDI

Nowadays, ample information is available on the internet, causing users to be overwhelmed, leading to information overload. This makes it difficult to find conferences on the internet, which interest them enough to publish papers and/or attend to. COnference REcommender (CORE) utilises Call For Papers (CFPs) to recommend conferences to researchers to solve this issue.

Recommender systems attempt to filter out information which would have been disregarded by users, and only present what is predicted to intrigue them. Amongst recommenders, content-based and collaborative-filtering are the most common. Content-based recommenders, analyse a user's previous high ratings of items and find similar items using the item content's features [1, 2, 3]. Using Collaborative filtering, positively rated items of users are suggested to those with similar interests [4]. Initially, in either method, the system will not contain prior knowledge, thus will not be able to accurately recommend items [5]. This issue is known as the cold start problem [5]. Knowledge-based recommenders give suggestions based on knowledge gathered beforehand about users and items, overcoming the cold start problem [4, 6].

Knowledge-based, content-based and collaborative-filtering techniques, together with a hybrid approach combining the last two, were implemented in CORE. The first step of CORE is collecting information about conferences by obtaining CFP data. This information is retrieved and stored inside a graph database, creating the conference layer. The second step is collecting information about each user, and constructing the user model based on his/her past actions. This information is collected from online publication sources, creating the second layer of the graph; the user layer. This information retrieval of CFPs and users shows the knowledge-based nature of CORE. The two layers are connected using content-based (through conference series and research topics) and collaborative-filtering (through user co-authors) links. Finally, using CFPs, conferences are recommended to users based on those links.

After evaluating the different implemented techniques, the best result was obtained using the hybrid approach, giving a recall of 0.311, a precision of 0.248 and an f1 of 0.257.

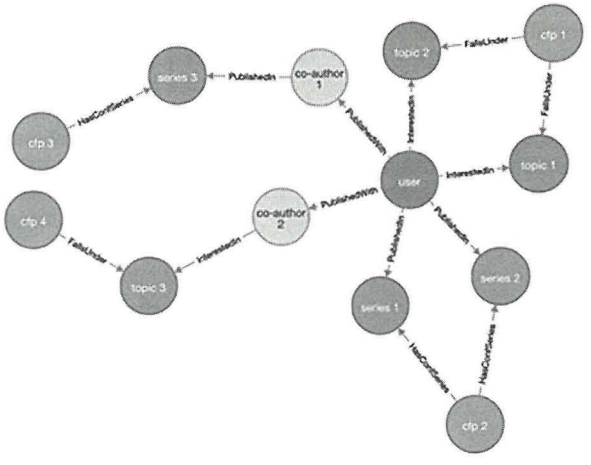


Figure 1: Figure showing the user connected to CFPs, using content-based techniques through topic and series links, and using collaborative filtering through co-author links.

REFERENCES

[1] N. Agarwal, E. Haque, H. Liu, and L. Parsons, "Research paper recommender systems: A subspace clustering approach," Advances in Web-Age Information Management, pp. 475–491, 2005.

[2] [2] P. Lops, M. De Gemmis, and G. Semeraro, "Content-based recommender systems: State of the art and trends," in Recommender systems handbook, pp. 73–105, Springer, 2011.

[3] M. Balabanovi´c and Y. Shoham, "Fab: content-based, collaborative recommendation," Communications of the ACM, vol. 40, no. 3, pp. 66–72, 1997.

[4] F. Ricci, L. Rokach, and B. Shapira, Introduction to recommender systems handbook. Springer, 2011.

[5] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender sys- tems: A survey of the state-of-the-art and possible extensions," IEEE transactions on knowledge and data engineering, vol. 17, no. 6, pp. 734–749, 2005.

[6] S. Trewin, "Knowledge-based recommender systems," Encyclopedia of library and in- formation science, vol. 69, no. Supplement 32, p. 180, 2000.

PROJECT NO. 76

Convolutional Neural Networks and the Bag-Of-Visual-Words Approach: A Comparative Analysis Based on Variation of Training Set Size

MICHELLE FALZON

SUPERVISED BY
DR GEORGE AZZOPARDI AND DR REUBEN FARRUGIA

Image classification has a variety of practical applications in computing, most notably medicinal (e.g. detecting diseases), industrial (e.g. locating faulty parts) and security (e.g. facial recognition). Despite their popularity, little research has been carried out on the comparison between two common image classification algorithms: the Bag-of-Visual-Words approach (BoVW) and Convolutional Neural Networks (CNN), in terms of the size of the training set required.

These methods have proven to be highly reliable. However, they each have different memory and time requirements as well as different optimal training set characteristics. The tendency towards using the current top performing algorithm, without regard for its best uses, has led many to use CNNs despite other methods being possibly more suitable for the task.

This work aimed at studying the BoVW approach and CNNs to determine their performance when presented with varied sized training sets. Subsets of the CIFAR-100 data set, containing tiny images and multiple classes, and the FERET data set, consisting of larger images and only two categories, were chosen for this project.

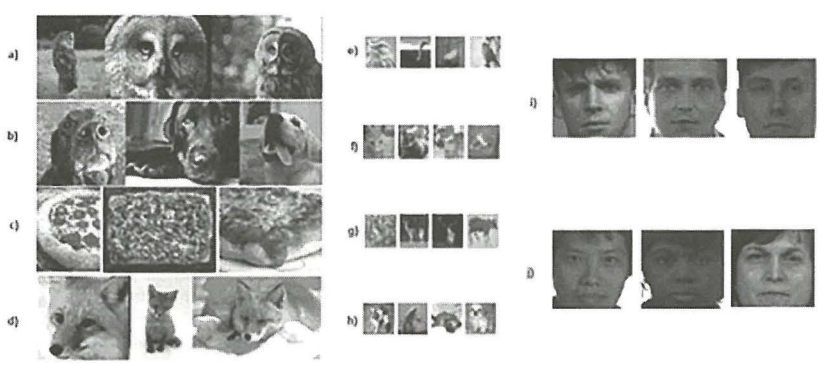


Figure 1:
Left - Images from ImageNet data set used to pre-train the CNN.
a) Great Grey Owl b) Labrador Retriever
c) Pizza d) Red Fox

Middle - Images from the CIFAR-100 data set
e) Bird f) Cat
g) Deer h) Dog

Right - Images from the FERET data set
i) Male faces j) Female faces

Upon training the entire CNN on the CIFAR-100 data set, the error rate did not decrease below 0.9, suggesting that more time and training images were required. In contrast, training using transfer learning was successful (Fig. 2), obtaining an accuracy of 77.55% on 20 classes using 200 images/class. The BoVW approach was less successful reaching a top accuracy of 54.05%. Re-initialising more layers of the CNN led to a lower performance - probably due to the fact that more training images were required. Adding more levels to the BoVW spatial pyramid resulted in a higher accuracy.

The opposite was seen on the FERET data set. The BoVW drastically outperformed the CNN, peaking at 92.82% with 200 images/category. The CNN required more than 50 images/category in order to provide decent results and, even then, its accuracy was slightly less than that of the BoVW.

From the results, one can conclude that using transfer learning on pre-trained CNNs is more suitable for large and complex data sets, while the BoVW approach performs best on simpler and smaller data sets.

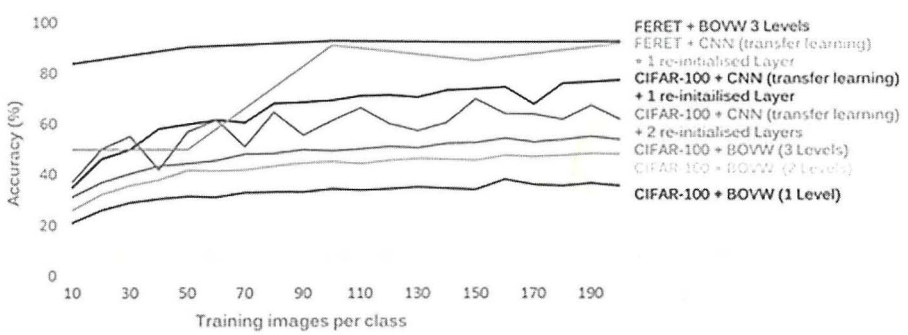


Figure 2: A graph showing the accuracies obtained using different configurations of the CNN and BoVW algorithms on the FERET and CIFAR-100 data sets

PROJECT NO. 29

Bringing Business Intelligence to Bear on Long-Term HR Planning

CLAIRE FARRUGIA

SUPERVISED BY
DR PETER A. XUEREB

Business Intelligence (BI) is "a broad field that combines people skills, technologies, applications, and business processes to make better strategic and tactical business decisions" [1]. It supports an organisation's decision making by combining "data gathering, data storage, and knowledge management with analytical tools to present complex internal and competitive information to planners and decision makers" [2]. Businesses are realising the importance and effectiveness of making decisions based on solid data rather than intuition, and Business Intelligence would seem the way forward for any business which wants to thrive.

While many businesses are adopting BI in many areas and reaping the associated benefits, the use of BI in Human Resource (HR) management and planning remains scant [3]. Human Resources, which may be defined as the human capital of an enterprise, and their management, are increasingly being identified as potential sources of competitive advantage for businesses [4]. Applying Business Intelligence to this important area can help realise this potential. One of the reasons for the scant use of BI in this field could be the lack of awareness of the capabilities of BI to provide insightful information to guide decisions and actions regarding Human Resources, their management and planning.

This study aimed to showcase the relevance of Business Intelligence to HR personnel, and subsequently assess the applicability, usefulness and benefits of this technology in Human Resource management and long-term planning. Taking employee absence as the main area of focus, and using Microsoft Power BI as a visualization tool, a series of dashboards and tailor-made visualizations displaying associated data were produced, based on the input of various HR professionals. The visualizations produced were then submitted for evaluation by a different group of HR professionals.

The results and positive feedback show that while different users have different requirements, Business Intelligence and data visualization can support Human Resource personnel very effectively in their roles as managers and planners. Professionals from both fields of BI and HR management can work together to provide the solution most appropriate to the particular business. Moreover, the degree of interest that participants in this study showed in adopting this technology would suggest it opportune to continue this investigation into areas of Human Resource management and planning other than employee absence.

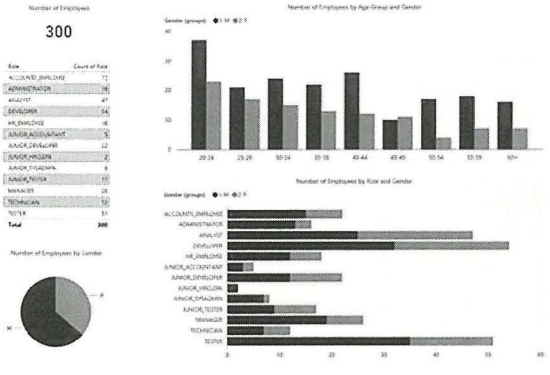


Figure 1: A General Dashboard

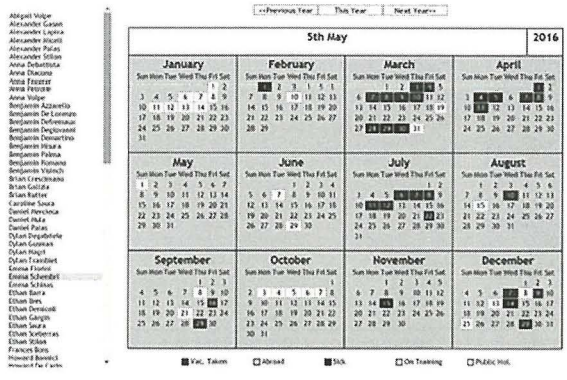


Figure 2: A Tailor-made Visualization

REFERENCES

[1] Bhushan Kapoor, "Business Intelligence and Its Use for Human Resource Management," *The Journal of Human Resource and Adult Learning*, vol. 6, no. 2, pp. 21-30, December 2010.

[2] Solomon Negash, "Business Intelligence," *Communications of the Association for Information Systems*, vol. 13, no. 15, pp. 177-195, February 2004.

[3] Bhushan Kapoor and Joseph Sherif, "Human resources in an enriched environment of business intelligence," *Kybernetes*, vol. 41, no. 10, pp. 1625-1637, 2012.

[4] Rūta Kazlauskaitė and Ilona Bučiūnienė, "The Role of Human Resources and Their Management in the Establishment of Sustainable Competitive Advantage," *Engineering Economics*, vol. 5, no. 60, pp. 78-84, 2008.

PROJECT NO. 32

Enhancing a Simulator and Visualization Platform For Human Activity Datasets – PacSim

GABRIEL FARRUGIA

SUPERVISED BY
DR CONRAD ATTARD

Business Intelligence (BI) is "a broad field that combines people skills, technologies, applications, and business processes to make better strategic and tactical business decisions" [1]. It supports an organisation's decision making by combining "data gathering, data storage, and knowledge management with analytical tools to present complex internal and competitive information to planners and decision makers" [2]. Businesses are realising the importance and effectiveness of making decisions based on solid data rather than intuition, and Business Intelligence would seem the way forward for any business which wants to thrive.

While many businesses are adopting BI in many areas and reaping the associated benefits, the use of BI in Human Resource (HR) management and planning remains scant [3]. Human Resources, which may be defined as the human capital of an enterprise, and their management, are increasingly being identified as potential sources of competitive advantage for businesses [4]. Applying Business Intelligence to this important area can help realise this potential. One of the reasons for the scant use of BI in this field could be the lack of awareness of the capabilities of BI to provide insightful information to guide decisions and actions regarding Human Resources, their management and planning.

This study aimed to showcase the relevance of Business Intelligence to HR personnel, and subsequently assess the applicability, usefulness and benefits of this technology in Human Resource management and long-term planning. Taking employee absence as the main area of focus, and using Microsoft Power BI as a visualization tool, a series of dashboards and tailor-made visualizations displaying associated data were produced, based on the input of various HR professionals. The visualizations produced were then submitted for evaluation by a different group of HR professionals.

The results and positive feedback show that while different users have different requirements, Business Intelligence and data visualization can support Human Resource personnel very effectively in their roles as managers and planners. Professionals from both fields of BI and HR management can work together to provide the solution most appropriate to the particular business. Moreover, the degree of interest that participants in this study showed in adopting this technology would suggest it opportune to continue this investigation into areas of Human Resource management and planning other than employee absence.

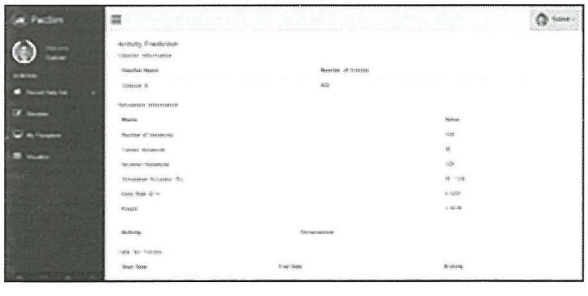


Figure 1: Image showing simulation of an uploaded dataset

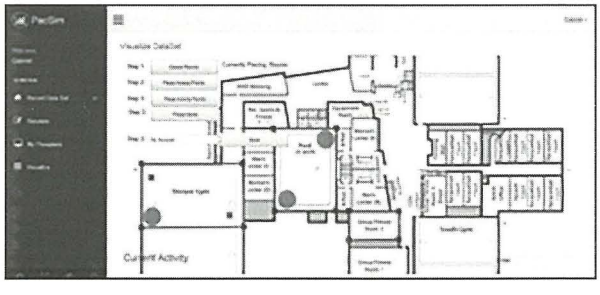


Figure 2: Image showing visualization of an uploaded dataset

REFERENCES

[1] Grech and Attard 2016, Generating datasets of human activities in pervasive spaces when using wearable devices, https://www.um.edu.mt/library/oar//handle/123456789/14751, University of Malta

[2] Eibe Frank, Mark A. Hall, and Ian H. Witten (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016

Optimising Path Tracing with Genetic Algorithms

JAMES FARRUGIA (161296M)

SUPERVISED BY
DR KEITH BUGEJA AND DR KEVIN VELLA

Presented with a scene description — a vector of object models in a 3-dimensional Cartesian coordinate space — the path-tracing algorithm propagates volumes of light rays into the scene and simulates the ray-object interactions that ensue. It is consequently possible to elegantly recreate a physically-correct image that is representative of a realistic environment, while supporting phenomena such as caustics and global illumination amongst others.

The class of ray-tracing algorithms, to which path-tracing belongs, presents approaches aimed at solving the rendering equation proposed by Kajiya [1]. Kajiya's work shows that producing perceptually noiseless physically-correct renders involves calculating a scattering integral over an inordinate number of samples. Path-tracing's largest contribution to its complexity is its attempt at approximating this integral via importance sampling over a finite set of values, in a process known as Monte Carlo integration.

This work presents a genetic programming framework to locate solutions in a search-space of program syntax-trees with the intent of finding code-segments in the path-tracing algorithm that outperform their standard implementation, while maintaining a reasonable degree of accuracy (see Figure 1). While similar approaches have been undertaken — such as the work by Dutrè et al. regarding Monte Carlo path generation using genetic algorithms [2] — this approach is unique as it uses genetic programming acting at the level of path-tracing source-code as opposed to higher-level constructs such as ray paths.

This work set out to determine whether it was possible to use a strongly-typed genetic programming framework to improve path-tracing source-code performance, and to develop alternative implementations that could guide further research (see Figure 2). It was in fact possible to approach the path-tracing problem from a computer-generated perspective; providing insight into parsimonious code-management using constrained function-sets. Techniques were developed to improve convergence rates for general problem solving, and the shortcomings of the system along with potential improvements (such as the need for meta-heuristics) were discussed.

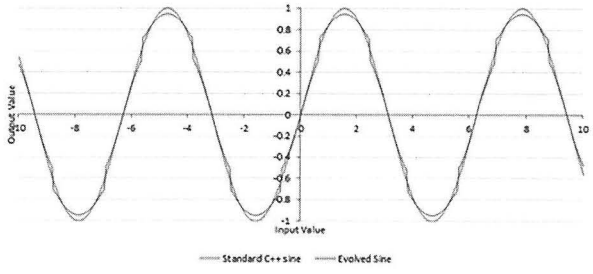


Figure 1: An example of an evolved sine function superimposed over its ground truth.

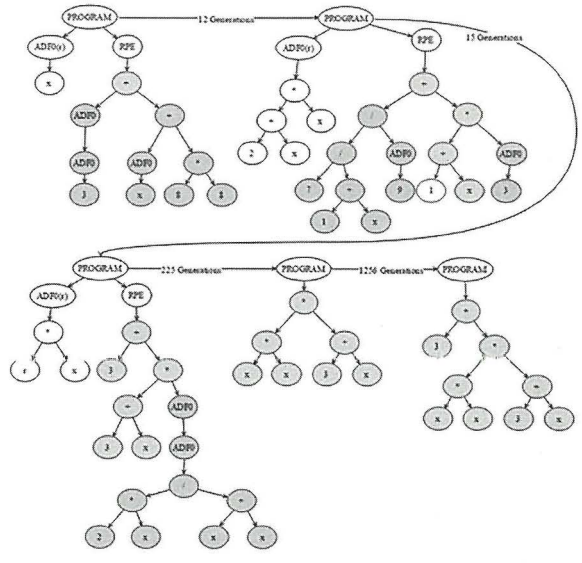


Figure 2: Snapshots of a simple example GP run at various generations, with a cubic ground-truth of x3 + 3x2 + 3, using a fitness = error10 + variance10 + bloat2. Good genes are coloured in green, good yet incomplete genes in yellow, and deleterious genes in red.

REFERENCES

[1] J. T. Kajiya. The rendering equation. In ACM Siggraph Computer Graphics. ACM, 1986.
[2] P. Dutrè, F. Suykens, and Y. D. Willems. Optimized Monte carol path generation using genetic algorithms. 1998.

PROJECT NO. 74

EXCITE – Extracting Geographic Information from Text

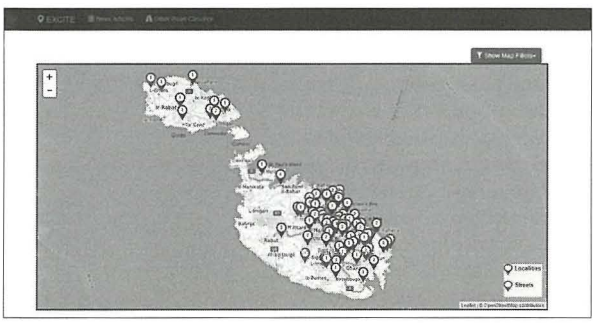
JAMES FARRUGIA

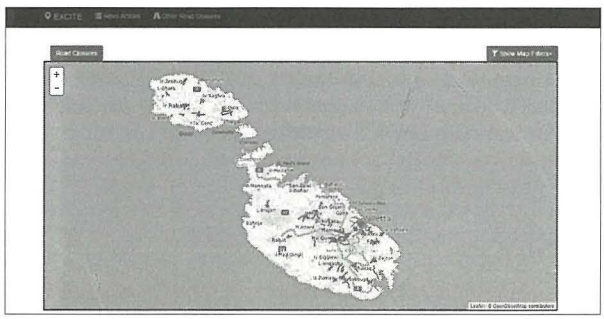
SUPERVISED BY
DR JOEL AZZOPARDI AND DR CHARLIE ABELA

Geotagging is an automatic process that detects geographic locations in unstructured text. In addition it maps the detected locations to latitude-longitude pairs [1]. EXCITE is a geotagging system that is applied on a number of online local news articles and other official notices such as police road closures and scheduled power cuts. It detects references to localities and streets of the Maltese Islands within collections of unstructured text (e.g. news reports). In addition, EXCITE visualises the detected locations on a map of the Maltese Islands accessed via a simple and interactive interface. The system accepts news articles written both in English and in Maltese. EXCITE is available online on *http://ec2-54-244-103-196.us-west-2.compute.amazonaws.com/*.

Geotagging is performed using a gazetteer based lookup method [2]. Our gazetteer contains references to localities and streets mapped to latitude-longitude pairs, in the form of a hierarchy. Our gazetteer for the Maltese Islands has been constructed by utilising freely available OpenStreetMap data. The geotagging system then applies Named Entity Recognition on English articles and, given the absence of NLP tools in Maltese, applies a number of rules to extract locations in Maltese articles. The system must disambiguate any ambiguities such as when a detection can refer to either a location or to another entity, or when a detection can refer to two distinct locations. EXCITE utilises the gazetteer in order to solve ambiguity and to ground each detection to latitude-longitude pairs. EXCITE also introduces the geotagging process to the Maltese language and establishes groundwork in the detection of location entities when considering the lack of language tools for Maltese.

EXCITE achieved satisfactory results when evaluated. The results show that the gazetteer contains references to all localities in Malta and Gozo and also refers to approximately 75% of streets. EXCITE also detects 91% and 86% of the referenced locations in English and Maltese news articles respectively. These results are very encouraging. The achieved results for the Maltese language are particularly satisfying since the implemented methods were constructed from scratch and make no use of language tools such as NERs.





REFERENCES

[1] C. Wang, X. Xie, L. Wang, Y. Lu, and W.-Y. Ma. Detecting geographic locations from web resources. In Proceedings of the 2005 workshop on Geographic information retrieval, pages 17-24. ACM, 2005.

[2] D. Buscaldi and B. Magnini. Grounding toponyms in an Italian local news corpus. In Proceedings of the 6th Workshop on Geographic Information Retrieval, page 15. ACM, 2010.

PROJECT NO. 53

Preliminary Investigations on Runtime Enforcement Implementations

KARL FARRUGIA

SUPERVISED BY
DR ADRIAN FRANCALANZA AND MR IAN CASSAR

Runtime Enforcement is a practical analysis technique that ensures the system behaves correctly, according to some behavioural specifications, via runtime checks. This technique adds an extra authentication layer that ensures that only correct behaviour is exhibited by the system to the outside environment.

Currently, there is little work on how to best conduct runtime enforcement in actor systems. Therefore, it is not yet clear which technique one should employ when creating a runtime enforcement monitor in actor systems.

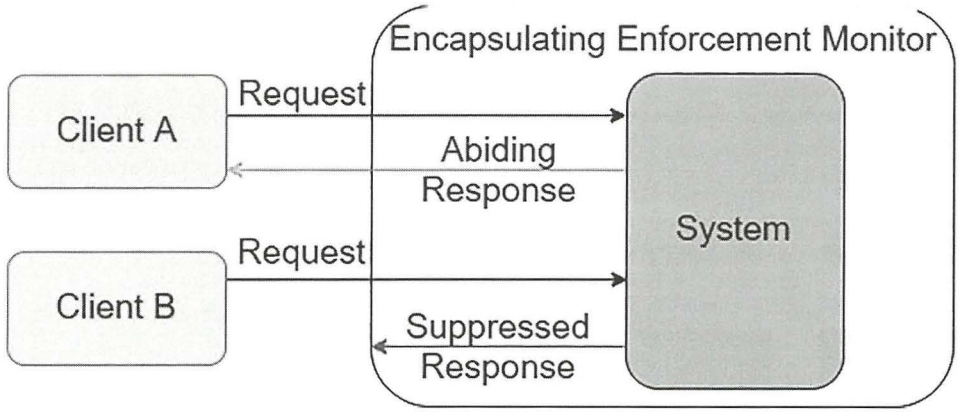
Actors are concurrent entities, uniquely identified by their process identifier, that execute asynchronously to one another. Actors can spawn other actors or communicate through asynchronous message passing.

An enforcement monitor prevents incorrect behaviour from being observed by the external environment. This is done by suppressing or altering the execution when it starts to deviate from the expected properties [1, 2, 3].

This work investigates three novel instrumentation strategies of how to apply a runtime enforcement monitor with actor systems. The three strategies we put forth intend to achieve the instrumentation of the system in a distinct manner. The first implementation is that of a send intercept strategy which is able to intercept outgoing messages from the system. A second implementation concerns a serialised strategy that only allows one client to execute at any moment in time. This strategy switches the client's address with its own as to deceive the system into communicating with the monitoring entity. The third and final strategy is similar to the serialised strategy in structure. This strategy assigns each client their own monitoring stub, and proceeds to swap the client's address with the address belonging to the monitoring stub. This work continues by explaining the reasoning behind each instrumentation strategy and accompanies it with a detailed description of how their implementation is carried out.

Finally, these strategies are applied on a third party software where an analysis in terms of their efficiency, in terms of overheads, and their genericity, in terms of their ability to encapsulate different types of systems, is performed. The chosen software is the Ranch [4] socket acceptor pool for TCP protocols. This software is chosen with the intent of increasing objectivity of results

Through our evaluation we conclude that the send intercept strategy yields the best results. However, this strategy induces the highest instrumentation burden and as a consequence it is not always viable to be used. In cases where the burden is too great to be handled it would be ideal to use the monitoring stubs as an alternate strategy. This strategy gives similar time efficiency but at a higher overhead cost. Ideally the serialised strategy should be avoided as it gives the worst performance with little to no overhead gains.



REFERENCES

[1] Y. Falcone. You should better enforce than verify. In International Conference on Runtime Verification, pages 89{105. Springer, 2010.

[2] Y. Falcone, J.-C. Fernandez, and L. Mounier. What can you verify and enforce at runtime? International Journal on Software Tools for Technology Transfer, 14(3):349{382, 2012.

[3] J. Ligatti and S. Reddy. A theory of runtime enforcement, with results. In European Symposium on Research in Computer Security, pages 87{100. Springer, 2010.

[4] Ranch socket acceptor pool. url: https://github.com/ninenines/ranch/.

PROJECT NO.
65

Aggression Detection in Urban Environments based on Audio Analysis

EDWARD FLERI SOLER

SUPERVISED BY
DR GEORGE AZZOPARDI

Amid escalating security concerns, improvements in technology and reductions in costs, the Electronic Security Systems market is projected to exceed US $80 billion by 2020 [1]. Surveillance systems have become common-place within urban settings such as public transport systems, parks and motorways. Current equipment is restricted by the man power required to manually monitor and report activity, and is susceptible to human error [2]. In this project, computer vision techniques are applied for the development of an automated aggression detection system for the detection of gunshots, breaking glass and screams in audio surveillance files.

Audio surveillance sensors are a cheaper, less-intrusive alternative to conventional Closed-Circuit Television (CCTV) cameras and are able to capture events with no visual counterpart [3]. A novel approach for the off-line detection of aggressive events in noisy, suboptimal conditions typical of urban environments is proposed. As opposed to traditional pipelines founded on the basis of low-level audio feature analysis [2,4,5,6,7], this proposal has been adapted to employ computer vision techniques upon time-frequency audio representations, in order to attain a high robustness to noise.

Training of the system commences with the translation of audio event snippets into their visual counterparts, through the application of logarithmic-frequency spectrograms. Image features are then extracted and, through the Bag-of-Words model [8], utilized for the generation of high level event descriptors. These descriptors are then employed for the training of classifiers in order to distinguish between salient events of different classes, as well as background noise.

A standalone localisation stage is implemented prior to classification, in order to analyse test input and detect temporal regions of suspected importance. The application of such a filtering stage prior to classification reduces the computational expense of standard windowing classifiers adopted within literature. A set of Support Vector Machines (SVM) are trained for each class at varying noise conditions,

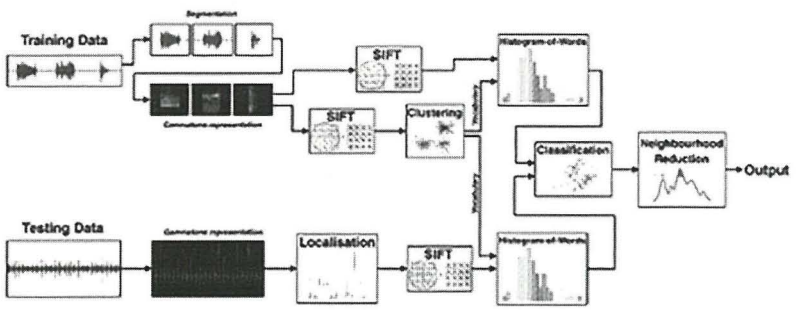


Figure 1: Pipeline of the proposed system

to produce the final classification for each detected region of importance.

The comprehensive MIVIA Audio Events Dataset [8], composed of event recordings, superimposed with complex, realistic background noise at varying Signal-to-Noise Ratios (SNR), has been acquired for the thorough analysis of the system. The robustness of the system to high levels of varying background noise has been confirmed, surpassing accuracy ratings achieved by similar systems within literature, and thus cementing the applicability of such an architecture within a forensic audio analysis context.

REFERENCES

[1] Electronic Security Systems (ESS) Market Trends." *Electronic Security Systems Market Trends*. N.p., Sept. 2015. Web. 24 May 2017.

[2] J-L Rouas, Jerome Louradour, and Sebastien Ambellouis. Audio events detection in public transport vehicle. In Intelligent Transportation Systems Conference, 2006. ITSC'06. IEEE, pages 733{738. IEEE, 2006.

[3] Pasquale Foggia, Nicolai Petkov, Alessia Saggese, Nicola Strisciuglio, and Mario Vento. Reliable detection of audio events in highly noisy environments. Pattern Recognition Letters, 65:22{28, 2015.

[4] Vincenzo Carletti, Pasquale Foggia, Gennaro Percannella, Alessia Saggese, Nicola Strisciuglio, and Mario Vento. Audio surveillance using a bag of aural words classifier. In Advanced Video and Signal Based Surveillance (AVSS), 2013 10th IEEE International Conference on, pages 81{86. IEEE, 2013.

[5] Chloe Clavel, Thibaut Ehrette, and Gael Richard. Events detection for an audio-based surveillance system. In Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on, pages 1306{1309. IEEE, 2005.

[6] Giuseppe Valenzise, Luigi Gerosa, Marco Tagliasacchi, Fabio Antonacci, and Augusto Sarti. Scream and gunshot detection and localization for audio-surveillance systems. In Advanced Video and Signal Based Surveillance, 2007. AVSS 2007. IEEE Conference on, pages 21{26. IEEE, 2007.

[7] Pasquale Foggia, Nicolai Petkov, Alessia Saggese, Nicola Strisciuglio, and Mario Vento. Audio surveillance of roads: a system for detecting anomalous sounds. IEEE Trans-actions on Intelligent Transportation Systems, 17(1):279{288, 2016.

[8] Zellig S Harris. Distributional structure. Word, 10(2-3):146{162, 1954.

A Comparison of Touch-Screen Interfaces and Voice-User-Interfaces: An Explorative Study for Multi-Modal Design

DANIEL EDWARD GALEA

SUPERVISED BY
DR PETER A. XUEREB

Modern mobile devices allow people to perform their day-to-day tasks anywhere in the world. Therefore, one would assume that these devices have interfaces which can accommodate the different contextual scenarios that users might find themselves in. However, generally mobile devices make use of a touch-screen, which uses a Graphical User Interface (GUI) [1] to display information to the users. While this type of interface has been widely accepted, it does not cater for scenarios in which a user's eyes and hands might be busy [2], and this scenario can be very common in mobile contexts. A user interface which can cater to these scenarios is a Voice User Interface (VUI) [1].

The aim of this dissertation is to build a GUI and a VUI for a small-screen mobile device, compare and evaluate them, and then combine them into Multi-Modal User Interface (MMUI) [1-3] so that this interface can also be compared and evaluated against the previous two. The GUI and VUI are compared using, as a measure, the time it takes for a user to complete various tasks, and the user's preference of input/output modalities. The results obtained are then used to evaluate the weaknesses and strengths of each UI in each test. Furthermore, the evaluation is used to determine whether the MMUI's added flexibility in terms of I/O has made it simpler for users to perform tasks, or more complex, as well as to see whether having multiple modes of I/O promotes greater 'safety' of use in specific environmental contexts, while still allowing users to accomplish certain specific goals with the application.

The results obtained show that users preferred having a MMUI rather than a unimodal interface, and that having different types of input and output possibilities made it easier to perform tasks and not more complex. Also, the MMUI did promote increased safety, as users could perform their tasks in an 'eyes busy' situation without risking injury to themselves or to others. The results also showed that, while users like having a VUI as part of the MMUI, they still have a 'natural' bias towards using GUIs – possibly due to their constant exposure to them.

Figure 1: Application main-menu

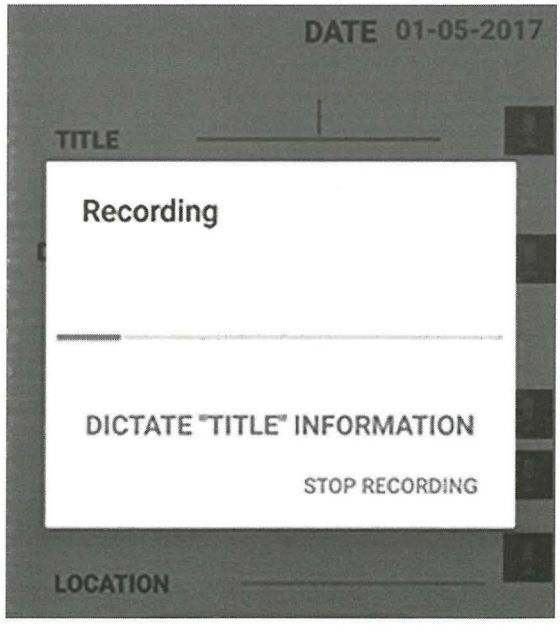


Figure 1: Voice recording dialog-box for filling a field

REFERENCES

[1] P. Kortum, HCI Beyond the GUI: *Design for Haptic, Speech, Olfactory, and Other Nontraditional Interfaces*. Morgan Kaufmann, 2008

[2] M. H. Cohen, J. P. Giangola and J. Balogh, *Voice User Interface Design*. Addison-Wesley Professional, 2004.

[3] M. Masterson. Can VUI and GUI SURVIVE an interface marriage? *Speech Technology* 18(4), pp. 20-23. 2013. Available: https://search-proquest-com.ejournals.um.edu.mt/docview/1461357210?accountid=27934

PROJECT NO. 25

Distributed Computation Frameworks for Solving Optimisation Problems

ETIENNE GALEA

SUPERVISED BY
MR MICHEL CAMILLERI

Evolutionary algorithms are extremely efficacious at finding near-optimal solutions for NP-Complete problems. However, complex combinatorial problems often demand high processing power which most of the time is unavailable to the ordinary public. A Genetic Algorithm (GA) is a stochastic-based search and optimisation technique that attempts to find a solution through an automated routine of trial-and-improvement (Goldberg, 1989). Simulating the process of natural selection, individuals compete against each other in a continuously evolving population to achieve a near-optimal solution. A Parallel Genetic Algorithm (PGA) can be employed to significantly reduce the time necessary for obtaining a near-optimal solution by executing multiple Genetic Algorithms simultaneously on several loosely-coupled machines.

After extensively researching GAs, PGAs, various communication topologies and different evolutionary computational frameworks, a tailor-made GA was developed in C++ and implemented into the EASEA framework; an open-source, artificial evolution platform that allows users to implement evolutionary algorithms and easily achieve parallelism (Maitre, Krüger, Querry, Lachiche, & Collet, 2012). PGA parameter settings such as migration rates and topology infrastructure were tactfully selected to promote optimal solution quality. A network of inter-connected, homogenous machines, much like the set-up commonly found in a school or office, was used for the PGA set-up. The PGA utilised a UDP communication protocol for the transfer of individuals, eliminating much of the communication overhead found by using a TCP/IP, as well as allowing for a higher level of robustness and scalability.

Results generated from this study indicated that time was dramatically decreased when using a PGA due to the increased amount of evaluations per second on different machines. However, given the same amount of evaluations, the GA and PGA arrived to a solution of similar quality. The PGA performed marginally better as the complexity of the knapsack problem arose. Thus, results indicate that a PGA may be a preferable option for problems with a very high complexity.

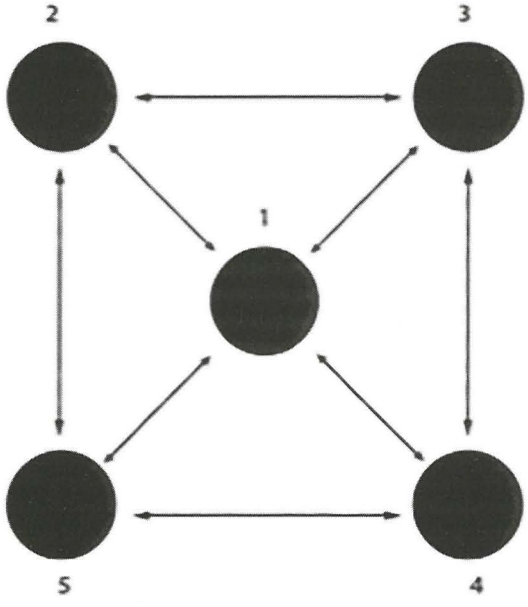


Figure 1: Designed topology for the PGA

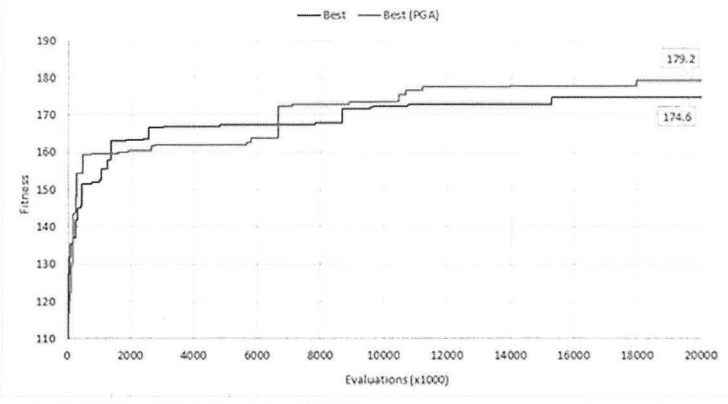


Figure 2 - GA and PGA runs for 10,000 items against total number of evaluations.

REFERENCES

[1] Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization & Machine Learning* (1st ed.). Boston, MA, USA: Addison-Wesley Longman Publishing Co.

[2] Maitre, O., Krüger, F., Querry, S., Lachiche, N., & Collet, P. (2012). EASEA: Specification and execution of evolutionary algorithms on GPGPU. *Soft Computing*, 16(2), 261–279. https://doi.org/10.1007/s00500-011-0718-z

PROJECT NO. 77

Smart Sensor for EEG Acquisition and Epileptic Seizure Detection

NOELLA GALEA

SUPERVISED BY
DR LALIT GARG AND MR KRISTIAN GUILLAUMIER

Epilepsy is a chronic neurological disorder which affects various people, independent of age. According to the World Health Organisation (WHO), around 50 million people worldwide have epilepsy, making it one of the most common neurological diseases globally [1]. Considering that seizures are unpredictable and unforeseen, epilepsy patients have no control on the seizure and this is the most threatening aspect of epilepsy. Symptoms include changes in movement, behaviour, sensation and awareness [2]. Early diagnosis of epileptic seizures is essential to the development of effective prevention treatments so the patient can undergo treatments which can delay or prevent disease progression.

This project investigates a novel and hybrid system for seizure detection and prediction by attempting to build a real time EEG analyser that strives to detect and predict seizures based on the EEG waveforms. By predicting an oncoming seizure within a given time frame, epileptic patients can prepare for the seizure and therefore remove themselves from any harm. With this goal in sight, in order to increase accuracy, irrelevant or distorted channels are removed from the EEG signals and the most important features and channels of the signals that distinguish between seizures and non-seizures states are extracted and utilised. Amongst these features, the frequency of these EEG signals together with the variation in amplitude are considered for better discrimination between abnormal and normal EEG signals. By calculating seizure likelihood in the time domain using the envelope detector and in the frequency domain by using the bandpass Butterworth filter, seizure detection and prediction using signal processing techniques is achieved. Another approach taken into consideration is classification, where using an innovative machine learning technique, other linear features of seizure and non-seizure states of the EEG signals such as skewness and variance are considered.

The system is tested using the publicly available CHB-MIT scalp EEG database to simulate real-time data which contains over 1000 hours of recorded scalp EEG signals

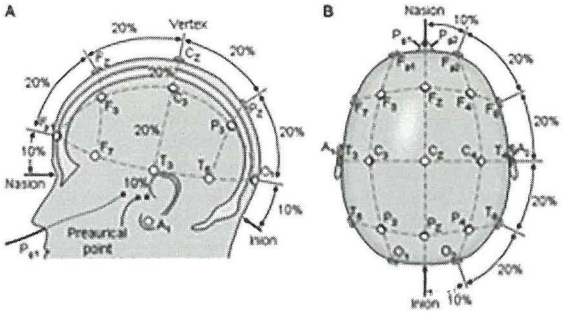


Figure 1: The international 10-20 system used to record EEG signals [3]

with over 190 seizure events of 23 distinct patients suffering from epilepsy. This proposed system offers an environment where users can test this model with minimal effort while allowing them to alter any parameters deemed fit.

Evaluation of the proposed system shows that analysis of EEG signals using digital signal processing and classification techniques can be useful in developing an efficient and accurate seizure prediction and detection system that can be employed in real time. Using the signal processing technique an average sensitivity of 100% and an average specificity of 73.65% were attained. The signal processing technique combined with the novel machine learning technique achieved an average sensitivity of 72.4% and an average specificity of 52.3%. These techniques obtained an average warning window of 74s prior to the seizure onset.

Even though this study is still in its infancy, the results obtained are promising and show considerable potential which justifies further research in order to identify the best way forward in the direction of a real life practically useful seizure detection and prediction machine. This would result in a novel and hybrid system that would be propitious especially to epileptic patients.

REFERENCES

[1] W. H. O. (WHO), 2012. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs999/en/.

[2] D. M. Ficker, "Sudden unexplained death and injury in epilepsy," *Epilepsia*, vol. 41, no. 2, pp. S7-S12, 2000.

[3] F. Sharbrough, G. E. Chatrian, R. P. Lesser, H. Luders, M. Nuwer and T. W. Picton, "American Electroencephalographic Society guidelines for standard electrode position nomenclature," *J. clin. Neurophysiol*, vol. 8, no. 2, pp. 200-202, 1991.

PROJECT NO. 33

Facilitating Parking Management Using Aerial Image Recognition

DANIEL GAUCI

SUPERVISED BY
DR CONRAD ATTARD

Aerial Image Recognition is a technological area which has enabled several solutions to problems in different sectors around the world. This study examines the use of a combination of automated multi-rotors (commonly referred to as drones), computer vision and machine learning as a solution to problems faced every day, taking parking management as a case study, and is part of SmartSkyAutomator [1].

Several different technological approaches have been taken to facilitate parking management, such as the use of an infrastructure of sensors around the parking space[2], applying computer vision techniques to footage captured by CCTV cameras[3] and using similar computer vision techniques on dash-mounted mobile phones to identify empty parking spaces[4]. This study recognizes the benefits of computer vision in this area, and the importance of a mobile platform which allows mobility of the system.

As part of the development of the proof of concept, a custom multi-rotor, shown in Figure 1, was built, around which a flight automation system was created. This multi-rotor was used to collect a dataset of the parking availability over a period of one week at one of the University of Malta parking facilities. Computer vision techniques were used to automate the process of determining the number of occupied and unoccupied parking spaces, with a sample result shown in Figure 2. An online survey was carried out with University students, lecturers and administration to collect information on the how they use all of the parking facilities at the University of Malta. A Machine Learning model was trained to predict the percentage of free parking spaces using the data captured from the multi-rotor and the survey.

Statistical analysis of the data collected showed that the computer vision technique used achieved satisfactory results, albeit with some limitations. Several patterns where observed for the parking patterns exhibited in the case study, identifying the ideal time to find a parking space in Car Park 6B at the University of Malta, which the end-user can visualize through a mobile application. The results of this study show that aerial image recognition is an effective tool to facilitate parking management.

Figure 1: Custom Built Multi-Rotor

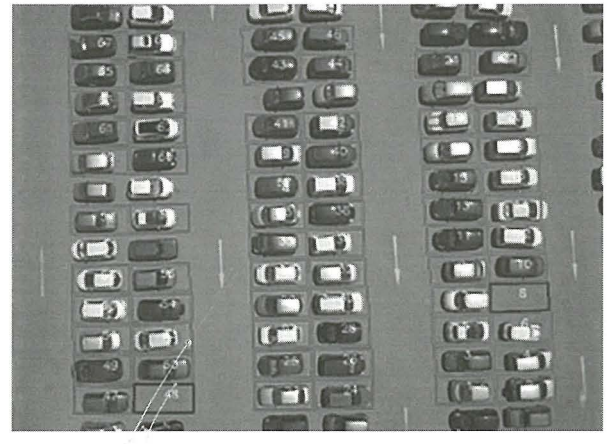


Figure 2: Processed Parking Space Image, with Red showing occupied and Blue showing unoccupied spaces

REFERENCES

[1] C. Attard, Principle Investigator of SmartSkyAutomator, Funded Project (Personal Communication)

[2] A. Bagula, L. Castelli, and M. Zennaro, "On the Design of Smart Parking Networks in the Smart Cities: An Optimal Sensor Placement Model.," *Sensors (Basel).*, vol. 15, no. 7, pp. 15443–67, Jun. 2015.

[3] M. Y. I. Idris, E. M. Tamil, Z. Razak, N. M. Noor, and L. W. Kin, "Smart Parking System using Image Processing Techniques in Wireless Sensor Network Environment," *Inf. Technol. J.*, vol. 8, no. 2, pp. 114–127, Feb. 2009.

[4] S. Patricio Figueroa Sanz, J. Tran, and C. Yu Wang, "Street Parking Detection."

PROJECT NO. 5

Collision Avoidance System for the RP Survey and Visual Inspection Train in the CERN Large Hadron Collider

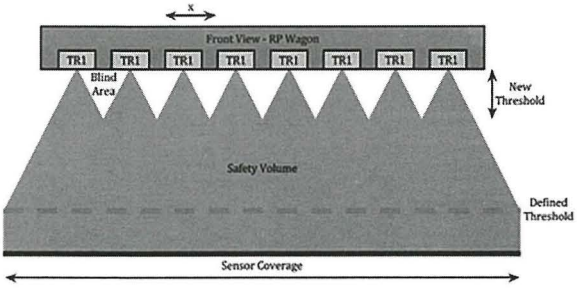
LEANDER GRECH

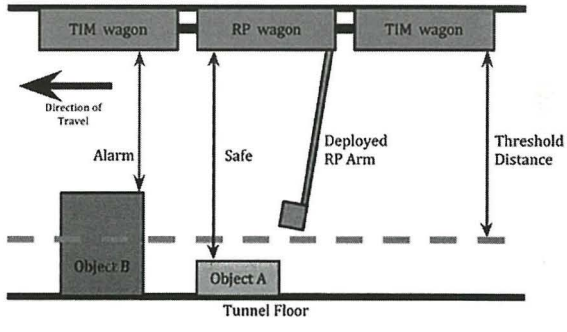
SUPERVISED BY
DR GIANLUCA VALENTINO

This Final Year Project focused on producing a working prototype of a collision avoidance system for the retractable Radio-Protection (RP) arm, in the Train Inspection Monorail (TIM), located in European Organization for Nuclear Research (CERN)'s Large Hadron Collider (LHC) Tunnel. This retractable arm is deployed in specific regions in the tunnel, after the beams have been dumped, to take radiation and oxygen level measurements around LHC equipment. In turn, these measurements allow the CERN Radiation Protection Group to assess any health hazards inside the LHC tunnel, before access to it is permitted to the maintenance personnel.

This prototype uses a series of eight distance measurement sensors based on the Infrared (IR) Time-of-Flight (ToF) technology, which scans a safety volume underneath the TIM and raises an alarm once this volume is breached. The TeraRanger One (TR1) IR ToF distance measurement sensor is used, which is interfaced via the TeraRanger Hub (TRH), by TeraBee. A Cortex-M microcontroller, is mounted on the TRH, which in turn is connected to the distance measurement sensors via a multiplexed USART connection among all eight sensors. The microcontroller also provides an interface for the user maintaining the collision avoidance system, via a second, dedicated USART connection.

Sensor characterisation tests were also devised and performed, from which calibration data for the TR1 sensors was obtained and the effects of prolonged use of the TR1 as well as that of partial Field-of-View coverage on the distance measurements, were observed. For the specific role of aiding in the data acquisition during the sensor characterisation tests, a testing rig was designed and constructed. This was designed to ensure that the distance from the sensor array to the surface being scanned, was always known during testing.





REFERENCES

[1] "numpy.gradient — NumPy v1.12 Manual," *Scipy*. [Online]. Available: https://docs.scipy.org/doc/numpy/reference/generated/numpy.gradient.html. [Accessed: 26-May-2017]

[2] "numpy.fft.ifft — NumPy v1.12 Manual," *Scipy*. [Online]. Available: https://docs.scipy.org/doc/numpy/reference/generated/numpy.fft.ifft.html. [Accessed: 26-May-2017]

[3] "Discrete Fourier Transform (numpy.fft) — NumPy v1.12 Manual," *Scipy*. [Online]. Available: https://docs.scipy.org/doc/numpy/reference/routines.fft.html. [Accessed: 26-May-2017]

[4] D. Jacquet, G. Kruk, and R. Gorbonosov, "LSA-the High Level Application Software of the LHC-and Its Performance During the First Three Years of Operation," in *Control System Upgrades*, San Francisco, CA, USA , 2014, pp. 1201–1204.

[5] T. Martin, "Chapter 6 – Developing with CMSIS RTOS," in *Designers Guide to the Cortex-M Processor family*, Elsevier Ltd., 2013, p. Pages 165–216.

[6] Ltd, Arm, "µVision User's Guide: Debugging," *Keil*. [Online]. Available: http://www.keil.com/support/man/docs/uv4/uv4_debugging.htm. [Accessed: 25-May-2017]

[7] Arm Ltd., "uVision User's Guide: System and Thread Viewer," *Keil*. [Online]. Available: http://www.keil.com/support/man/docs/uv4/uv4_db_dbg_systhread_viewer.htm. [Accessed: 25-May-2017]

[8] Ltd, Arm, "C51: CALCULATING STACK SIZE," *Keil*, 30-Jan-2015. [Online]. Available: http://www.keil.com/support/docs/192.htm. [Accessed: 25-May-2017]

[9] J. Hunter, D. Dale, E. Firing, and M. Droettboom, "Matplotlib: Python plotting — Matplotlib 2.0.2 documentation," *Matplotlib*, 10-May-2017. [Online]. Available: https://matplotlib.org/. [Accessed: 24-May-2017]

[10] "numpy.poly1d — NumPy v1.12 Manual," *SciPy*, 16-Jan-2017. [Online]. Available: https://docs.scipy.org/doc/numpy-1.12.0/reference/generated/numpy.poly1d.html. [Accessed: 24-May-2017]

[11] "numpy.polyfit — NumPy Manual," *SciPy*, 16-Jan-2017. [Online]. Available: https://docs.scipy.org/doc/numpy/reference/generated/numpy.polyfit.html. [Accessed: 24-May-2017]

[12] "NumPy," *NumPy*. [Online]. Available: http://www.numpy.org/. [Accessed: 24-May-2017]

[13] B. Earl, "Calibrating Sensors." Adafruit Industries, 10-Sep-2016 [Online]. Available: https://cdn-learn.adafruit.com/downloads/pdf/calibrating-sensors.pdf

[14] M. Lindner, I. Schiller, A. Kolb, and R. Koch, "Time-of-Flight sensor calibration for accurate range sensing," *Comput. Vis. Image Underst.*, vol. 114, no. 12, pp. 1318–1328, Dec. 2010.

[15] "Terabee - Home," *TeraBee*. [Online]. Available: http://www.terabee.com/. [Accessed: 04-May-2017]

[16] "About CERN," *CERN*. [Online]. Available: https://home.cern/about. [Accessed: 08-May-2017]

[17] "Arduino - ArduinoBoardUno," *Arduino*. [Online]. Available: https://www.arduino.cc/en/main/arduinoBoardUno. [Accessed: 10-May-2017]

[18] A. Behrens, P. Bestmann, C. Charrondiere, T. Feniet, J. L. Grenard, and D. Mergelkuhl, "The LHC Collimator Survey Train," CERN [Online]. Available: http://iwaa2010.desy.de/e107506/e107507/e113279/e119275/IWAA2010_PB_train.pdf

[19] B. Feral, "Train for RP survey and visual inspection in LHC," 8 May, 2013 [Online]. Available: https://indico.cern.ch/event/248708/contributions/1572079/attachments/429323/595780/RadWG_08_May_2013.pdf

[20] "ULN2801A, ULN2802A, ULN2803A, ULN2804A - Eight Darlington array," ST, Nov. 2012 [Online]. Available: http://www.st.com/content/ccc/resource/technical/document/datasheet/e4/fa/1c/16/4e/39/49/04/CD00000179.pdf/files/CD00000179.pdf/jcr:content/translations/en.CD00000179.pdf

[21] "Fritzing Home," *Fritzing*. [Online]. Available: http://fritzing.org/home/. [Accessed: 19-May-2017]

[22] "PM55L-048," Minebea Co., Ltd. [Online]. Available: http://www.nmbtc.com/content/pdfs/pm55l-048.pdf

[23] T. Hammer, "The-Hammer: HTerm-A Terminal Program for Windows and Linux," *Der-hammer*. [Online]. Available: http://www.der-hammer.info/terminal/. [Accessed: 17-May-2017]

[24] S. Richter, "Minimum Requirements of a Microcontroller to Run Embedded Linux?," *Stack Exchange - Electrical Engineering*, 10-Jun-2016. [Online]. Available: https://electronics.stackexchange.com/questions/240227/minimum-requirements-of-a-microcontroller-to-run-embedded-linux. [Accessed: 15-May-2017]

[25] A. Industries, "Adafruit VL53L0X Time of Flight Distance Sensor - ~30 to 1000mm," *Adafruit* [Online]. Available: https://www.adafruit.com/product/3317. [Accessed: 14-May-2017]

[26] Sharp, "GP2Y0A Series / GP2Y0D Series Application Note" [Online]. Available: http://www.sharp-world.com/products/device/lineup/data/pdf/datasheet/gp2y0a_gp2y0d_series_appl_e.pdf

[27] A. Carullo and M. Parvis, "An Ultrasonic Sensor for Distance Measurement in Automotive Applications," *IEEE Sens. J.*, vol. 1, no. 2, p. 5, 2001.

[28] Sharp, "GP2Y0A21YK0F" [Online]. Available: http://www.sharp-world.com/products/device/lineup/data/pdf/datasheet/gp2y0a21yk_e.pdf

[29] E. Freaks, "Ultrasonic Ranging Module HC - SR04" [Online]. Available: http://www.micropik.com/PDF/HCSR04.pdf

[30] S. Adarsh, M. Kaleemuddin, D. Bose, and K. I. Ramachandran, "Performance Comparison of Infrared and Ultrasonic Sensors for Obstacles of Different Materials in Vehicle/Robot Navigation Applications," presented at the Advances in Materials and Manufacturing Applications (IConAMMA 2016), Bengaluru, Karnataka, India, 2016, p. 8.

[31] R. V. Aroca and G. Caurin, "A Real Time Operating Systems (RTOS) Comparison," Escola de Engenharia de Sao Carlos (EESC) – Universidade de Sao Paulo (USP) , Apr. 2014 [Online]. Available: https://www.researchgate.net/publication/228940960_A_Real_Time_Operating_Systems_RTOS_Comparison. [Accessed: 11-May-2017]

[32] G. C. Buttazzo, *Hard Real-Time Computing Systems - Predictable Scheduling Algorithms and Applications*, vol. Real-Time Systems. Kluwer Academic Publishers, 1997.

[33] P. Raghavan, A. Lad, and S. Neelakandan, *Embedded Linux System Design and Development*. Auerbach Publications, 2006.

[34] "Arduino - ArduinoBoardUno," *Arduino*. [Online]. Available: https://www.arduino.cc/en/main/arduinoBoardUno. [Accessed: 10-May-2017]

[35] V. Parma, "Cryostat Design," *arXiv [physics.acc-ph]*, CERN, 28-Jan-2015 [Online]. Available: http://arxiv.org/abs/1501.07154

[36] CERN, "The Large Hadron Collider," *CERN*. [Online]. Available: https://home.cern/topics/large-hadron-collider. [Accessed: 08-May-2017]

[37] "About CERN," *CERN*. [Online]. Available: https://home.cern/about. [Accessed: 08-May-2017]

[38] M. Benedikt and F. Zimmermann, "Status of the Future Circular Collider Study," in *Colliders*, St. Petersburg, Russia, 2017, pp. 34–38.

[39] K. Kershaw, C. Bertone, D. Forkel-Wirth, N. Rousset, T. Feniet, J. L. Grenard, and P. Bestmann, "Remotely Operated Train for Inspection and Measurement in CERN's LHC Tunnel," in *Accelerator Technology - Subsystems*, Vancouver, BC, Canada, 2009, pp. 2902–2904.

[40] "Terabee - Home," *TeraBee*. [Online]. Available: http://www.terabee.com/. [Accessed: 04-May-2017]

[41] TeraBee, "TeraRanger One - Best distance sensor for robots!," *TeraRanger*. [Online]. Available: http://www.teraranger.com/products/teraranger-one/. [Accessed: 04-May-2017]

[42] Haryono, A. Eko Putra, J. E. Istiyanto, and A. Harjoko, "Error Detection and Correction System (EDAC) of On Board Data Handling (OBDH) in Real TimeOperating System Behaviour," *Scientific Research Journal*, vol. 10, no. 2, pp. 17–28, 2013.

[43] A. A. C. Keil™, *Keil Embedded Development Tools for ARM, Cortex-M, Cortex-R4, 8051, C166, and 251 processor families*. [Online]. Available: http://www.keil.com/. [Accessed: 04-May-2017]

[44] A. Ltd and A. G. GmbH, "Advantages of using an RTOS." [Online]. Available: http://www.keil.com/rl-arm/rtx_rtosadv.asp. [Accessed: 04-May-2017]

[45] ARM, "RTX Real-Time Operating System," *ARM*. [Online]. Available: https://www.arm.com/products/tools/software-tools/mdk-arm/middleware-libraries/rtx-real-time-operating-system.php. [Accessed: 04-May-2017]

[46] N. Lethaby, "Why Use a Real-Time Operating System in MCU Applications," 2013 [Online]. Available: http://www.ti.com.cn/cn/lit/wp/spry238/spry238.pdf

[47] R. David, N. Merriam, and N. Tracey, "How Embedded Applications using an RTOS can stay within On-chip Memory Limits" [Online]. Available: https://www-users.cs.york.ac.uk/~robdavis/papers/Euro2.pdf

[48] J. Lee, V. J. Mooney III, A. Daleby, K. Ingström, T. Klevin, and L. Lindh, "A Comparison of the RTU Hardware RTOS with a Hardware/Software RTOS," in *Proceedings of the 2003 Asia and South Pacific Design Automation Conference*, Kitakyushu, Japan, 2003, pp. 683–688.

[49] S. A. Daud, N. H. Mahmood, P. L. Leow, R. Sudirman, and M. R. A. Kadir, "Automated Sensor Rig in Detecting Shape of an Object," *Procedia Comput. Sci.*, vol. 42, pp. 153–159, Jan. 2014.

[50] Y. Liu, J. Chang, J. Lian, Z. Liu, Q. Wang, and C. Zhu, "A Time Difference Method for Measurement of Phase Shift between Distributed Feedback Laser Diode (DFB-LD) Output Wavelength and Intensity," *Sensors* , vol. 15, no. 7, pp. 16153–16161, Jul. 2015.

[51] D. Um and W. N. p. Hung, "A Novel Infrared Proximity Array Sensor for 3D Visual Sensing: Modeling and Applications," in *2006 IEEE Conference on Robotics, Automation and Mechatronics*, 2006, pp. 1–6.

[52] Y. Do and J. Kim, "Infrared Range Sensor Array for 3D Sensing in Robotic Applications," *Int. J. Adv. Rob. Syst.*, vol. 10, no. 4, p. 193, Jan. 2013.

[53] H. Park, S. Baek, and S. Lee, "IR sensor array for a mobile robot," in *Proceedings, 2005 IFFF/ASME International Conference on Advanced Intelligent Mechatronics.*, 2005, pp. 928–933.

[54] M.-C. Amann, T. Bosch, M. Lescure, R. Myllyla¨, and M. Rioux, "Laser ranging: a critical review of usual techniques for distance measurement," *Opt. Eng.*, vol. 40, no. 1, pp. 10–19, Jan. 2001.

[55] TeraBee, "TeraRanger One - Technical Specifications" [Online]. Available: http://www.teraranger.com/wp-content/uploads/2015/04/TeraRangerOne-Brochure.pdf

[56] G. Brown, *Discovering the STM32 Microcontroller*. Indiana University, 2016 [Online]. Available: https://www.cs.indiana.edu/~geobrown/book.pdf

[57] NetBurner, "Comparing RTOS to Infinite Loop Designs." 14 October, 2016 [Online]. Available: https://www.google.com.mt/url?sa=t&rct=-j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwjb1rHO-s7TAhUlV-RoKHSjwCroQFggoMAA&url=http%3A%2F%2Fwww.netburner.com%2Fmod54415%2Fdevelopment-kit%2F51-850%2Ffile&usg=AFQjC-NE7cNuXZr0jx0VRadNT6Sk3zJWslA&sig2=NFC7ZA9_KnioTy_DJFk-ow&-cad=rja

[58] ST, "STM32F103x8, STM32F103xB," Aug. 2015 [Online]. Available: http://www.st.com/content/ccc/resource/technical/document/datasheet/33/d4/6f/1d/df/0b/4c/6d/CD00161566.pdf/files/CD00161566.pdf/jcr:content/translations/en.CD00161566.pdf

[59] "Why RTOS and What is RTOS?," *Freertos*. [Online]. Available: http://www.freertos.org/about-RTOS.html. [Accessed: 11-Dec-2016]

[60] ARM, "RTX Real-Time Operating System," *ARM*, 2016. [Online]. Available: https://www.arm.com/products/tools/software-tools/mdk-arm/middleware-libraries/rtx-real-time-operating-system.php. [Accessed: 11-Dec-2016]

[61] STMicroelectronics, "STM32Cube Embedded Software," *ST*, 2016. [Online]. Available: http://www.st.com/en/embedded-software/stm32cube-embedded-software.html?querycriteria=productId=LN1897. [Accessed: 11-Dec-2016]

[62] K. Kershaw, B. Feral, J.-L. Grenard, T. Feniet, S. De Man, C. Hazelaar-Bal, C. Bertone, and R. Ingo, "Remote Inspection, Measurement and Handling for Maintenance and Operation at CERN," *Int. J. Adv. Rob. Syst.*, vol. 10, no. 11, Nov. 2013 [Online]. Available: http://journals.sagepub.com/doi/pdf/10.5772/56849

[63] European Organization for Nuclear Research, O. Bruning, and European Council for Nuclear Research, *LHC Design Report, Volume I: The LHC Main Ring*. 2004.

[64] TeraBee, "TeraRanger One - Installation and Operation Manual," Jun. 2016 [Online]. Available: http://www.teraranger.com/wp-content/uploads/2016/06/TR_One_Manual_FW5.0.Version1_0_0.pdf

[65] TeraBee, "TeraRanger Hub - Manual," 2016 [Online]. Available: http://www.teraranger.com/wp-content/uploads/2016/05/TeraRangerHubManualVersion1_1.pdf

[66] "TeraRanger Hub - Small, light, fast plug n' play multi-sensor solution," *TeraRanger*. [Online]. Available: http://www.teraranger.com/products/teraranger-hub/. [Accessed: 11-Dec-2016]

[67] M. Ruffo, M. Di Castro, L. Molinari, R. Losito, A. Masi, J. Kovermann, and L. Rodrigues, "New InfraRed Time of-Flight Measurement Sensor for Robotic Platforms," in *Research on Electric and Electronic Measurement for the Economic Upturn*, Benevento, Italy, p. 6.

[68] TeraRanger, "Time-of-Flight Principle," *TeraRanger*, 2016. [Online]. Available: http://www.teraranger.com/technology/time-of-flight-principle/. [Accessed: 11-Dec-2016]

PROJECT NO. 54

Efficient Rendering of Soft Shadows for Multiple Dynamic Area Light Sources

NICHOLAS GRISCTI SOLER

SUPERVISED BY
DR KEITH BUGEJA AND DR SANDRO SPINA

Realistic soft shadows, cast by area and volumetric light sources, are highly sought after in computer graphics. These consist of an umbra and a penumbra [3, 5] – the portion of the shadow that is partially occluded from the light source [2], resulting in a blurry silhouette. Nevertheless, soft shadows are more computationally expensive to render than traditional hard-edged shadows, due to complexities in penumbra calculations.

In this work, we are primarily concerned with real-time rendering, which puts the additional constraint of interactivity on the generation of soft shadows: dynamic illumination models where light sources and scene objects can change at runtime preclude the use of traditional approaches that precompute shadowing information and encode (or bake) it into visual elements a priori.

We propose a novel approach based on area light point-sampling [6] which allows user-parameterised trade-offs between visual quality and rendering performance. The approach is suitable for area lights which emit light uniformly along their surface (majority of light sources). The sampling process is independent of the complexity of area light geometry; it makes use of depth and normal maps generated from the area light. These maps (or buffers) are then sparsely sampled to reconstruct surface positions on the area light geometry at runtime and generate a volumetric approximation of the lighting via individual point-based light sources. The cost of rendering these individual point sources is further reduced through an adaptation of the Barnes-Hut algorithm [1] (approximation method for the n-body problem in Astrophysics), which clusters reconstructed sample points using a heuristic that takes into consideration camera proximity. The method is illustrated in Figure 2, where geometrically correct soft-shadows are generated; these can be contrasted to the hard-edged shadows in Figure 1.

Visual quality and performance metrics [4] were used to evaluate the proposed method across a variety of scenes and light source geometries. As expected, results show an improvement in performance as the number of samples decreases, albeit at the cost of image quality. Notwithstanding, perceptually-weighted image difference metrics show that even when the sample reduction is substantial (within limits), the decrease in image quality is virtually imperceptible to a human observer. This trade-off allows for the use of dynamic area light sources at interactive rates.

Figure 1: Conference Room scene rendered with hard shadows cast from a point light source

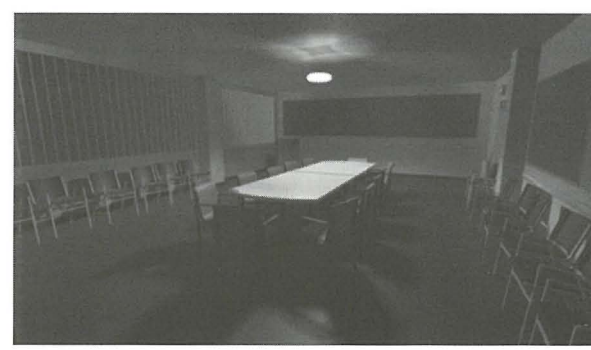


Figure 2: Soft shadows from a toroidal light source rendered using our technique

REFERENCES

[1] Josh Barnes and Piet Hut. A hierarchical O(n log n) force-calculation algorithm. nature, 324(6096):446{449, 1986.
[2] Robert L Cook, Thomas Porter, and Loren Carpenter. Distributed ray tracing. In ACM SIGGRAPH Computer Graphics. ACM, 1984.
[3] John F Hughes, Andries Van Dam, James D Foley, and Steven K Feiner. Computer graphics: principles and practice. Pearson Education, 2014.
[4] Image quality assessment online. http://driiqm.mpi-inf.mpg.de/index.php. (Accessed on 05/23/2017).
[5] Peter Shirley, Michael Ashikhmin, and Steve Marschner. Fundamentals of computer graphics. CRC Press, 2015.
[6] László Szirmay-Kalos. Monte Carlo methods in global illumination. 2000.

PROJECT NO. 55

Extending Android's Binder as a Basis for Application Monitoring

GIANCARLO LAFERLA

SUPERVISED BY
DR MARK VELLA AND DR JOSHUA ELLUL

Android is the leading operating system (OS) for mobile devices. While equipped with device diagnostic functionality it is still missing an OS-centric monitoring capability for apps, without requiring their modification. Such monitoring benefits a range of scenarios, from the verification of correct app functionality at runtime to device security e.g. to track complex permission leakage operated by malware that turns mobile phones into spying devices.

Despite being Linux based, Android does away with System V Inter- Process Communication (IPC) and uses a Remote Procedure Call (RPC) mechanism called Binder. Binder connect apps to each other as well as to device services, e.g. telephony and camera, making it an ideal location from where to provide the required OS-centric monitoring. This mechanism is tucked underneath the Java APIs exposed to developers [1]. Client apps use a proxy object to send marshalled/serialized remote call data to the Binder driver. The Binder driver in turn forwards this data to the service through its stub object as shown in Figure 1.

The first part of this work identifies suitable patch points to provide the required monitoring. An AIDL (Android Interface Definition Language)-driven app, that interacts directly with Binder, is used as a medium for fully comprehending the nature of its IPC flows [2]. The resulting Binder patches are applied at the Java level, with remote calls being intercepted immediately on the client-side. A string representation of the call data is printed using Logcat, as well as sent to a purposely-built monitoring app using Java-level inter-app communication objects: Intents.

The second part of this project focuses on the monitoring application per se. On Intent receipt the monitored remote call data is decomposed and parsed into simpler parts, enabling correct display. This is achieved using an intermediate representation of the AIDL files related to the services being monitored. Applications accessing different services were used as case studies in order to demonstrate the effectiveness of the monitoring functionality achieved. These were showcased on an Android emulator running the modified Android image, as shown in Figure 2.

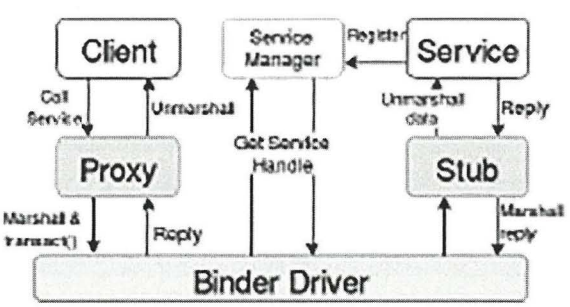


Figure 1: Binder Inter-Process Communication Flow

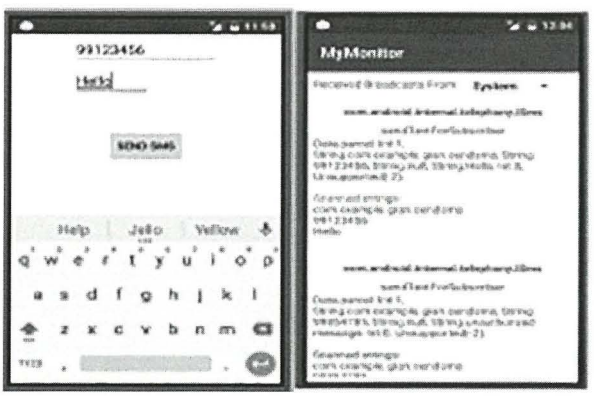


Figure 2: Left: Sample messaging application
Right: Monitoring application

REFERENCES

[1] T. Schreiber. Android binder. a shorter, more general work, but good for an overview of binder, 2011

[2] Google and O. H. Alliance. Android interface definition language. https://developer.android.com/guide/components/aidl.html [Online;accessed:2017-02-14]

FIRE: Finding Important News REports

PROJECT NO. 72

NICHOLAS MAMO

SUPERVISED BY
DR JOEL AZZOPARDI

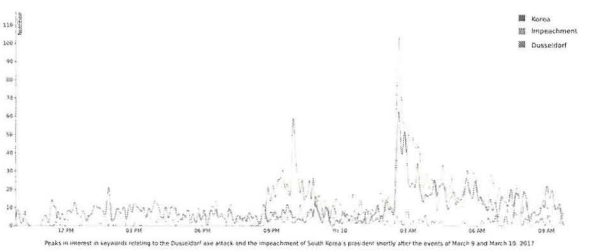
Every day, an immeasurable number of news items are published, with social media greatly contributing to the dissemination of information. As one of the social networks with the largest user bases, Twitter is not only a tool that facilitates communication between people; the succinct messages can be written in seconds, but contain a wealth of information. Thus, it comes as no surprise that Twitter has become an invaluable tool in breaking stories, serving as a news outlet for many users[1].

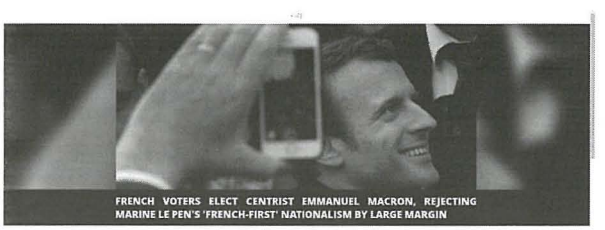
Finding Important News REports (FIRE) tackles the problem of information overload by identifying topics of interest to the general Twitter population in real time. This is achieved by exploiting the metrics available on Twitter to detect and track breaking news items, and articles that discuss them, based solely on tweet content[2].

The developed solution is split into different components, including clustering, emerging topic detection and tracking, and spam and noise detection and removal. Given the context-specific problems of Twitter, FIRE modifies the No-K-Means clustering algorithm detailed in [3] to collect as much information about topics as possible. Subsequently, an adaptation of the emerging topic detection algorithm presented in [4] is used to collect multiple breaking events.

FIRE also tackles spam and noise - a prevalent problem on Twitter[1] - by using existing and new metrics to filter out unremarkable tweets[5, 6]. Each component has a clearly-defined role in the bigger picture, and has been the subject of research in this project. These different methods, albeit widely-researched in information systems, present context-specific problems when dealing with the micro-messages of Twitter.

Ultimately, FIRE demonstrates how even a small sample of tweets can be used to extract newsworthy stories, although results indicate a dependency on a reliable spam detection and removal tool. On the other hand, Twitter's conversation habits do nothing to shackle the clustering process, with the system performing well across different news categories. Finally, it is shown that the adopted workflow can pick up emerging topics earlier than Twitter itself.





REFERENCES

[1] A. Java, X. Song, T. Finin, and B. Tseng, "Why we twitter: Understanding microblogging usage and communities," in Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis, WebKDD/SNA-KDD '07, (New York, NY, USA), pp. 56–65, ACM, 2007.

[2] M.-C. Yang and H.-C. Rim, "Identifying interesting twitter contents using topical analysis," Expert Syst. Appl., vol. 41, pp. 4330–4336, July 2014.

[3] J. Azzopardi and C. Staff, "Incremental clustering of news reports," Algorithms, vol. 5, no. 3, p. 364, 2012.

[4] M. Cataldi, L. D. Caro, and C. Schifanella, "Personalized emerging topic detection based on a term aging model," ACM Trans. Intell. Syst. Technol., vol. 5, pp. 7:1–7:27, Jan. 2014.

[5] B. Wang, A. Zubiaga, M. Liakata, and R. Procter, "Making the most of tweet-inherent features for social spam detection on twitter," arXiv preprint arXiv:1503.07405, 2015.

[6] F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida, "Detecting spammers on twitter," in In Collaboration, Electronic messaging, Anti-Abuse and Spam Conference (CEAS, 2010.

PROJECT NO. 3

Video Transmission using 10G Ethernet

SHAUN MARKHAM

SUPERVISED BY
DR TREVOR SPITERI AND MR ANDRE MICALLEF

The Memento Project is a project funded by MCST. It aims to commercialise home-grown electronics technology in the high performance vision market. This project aims to address the problem of precision synchronisation when dealing with groups of high speed cameras, enabling the capturing of subject media from a multitude of different angles. This technology would be extremely important for scientific experimentation, among other applications, where the gathering of different facets of information from contrasting angles yields a full and complete picture.

The aim of this project was to create a system whereby the data generated from the high performance imaging hardware may be transported to a centralised hub. It was designed within the Vivado Design Suite and had the final aim of reaching a benchmarked speed of 10 Gigabits per second (10G).

The system was implemented using pre-defined cores that each provided functionality suitable for the aims of this project. It encompassed an efficient and dedicated means for the transferring of high performance imaging. The project area dealt with communications and networking on a hardware implementable level. This, in turn, is a project detailed and carried out to the specification provided by the project actuators, keeping in mind the software, licenses and hardware available.

A modular approach was taken to the problem, looking to break it down to its core elements and work it back up to a final complete design. Numerous tests were performed to verify the final design's functionality, together with simulations carried out at each stage of the design process. Particular difficulty was seen with respect to the synchronisation of received video data together with paired video timing signalling and this was solved through the introduction of a FIFO buffer together with other components.

This project deals with, and conforms to, the specifications set out by the IEEE as well as the standards set by ARM. The main problem was creating a complete point to point system, in this case tested singularly in software, where this system could provide an efficient implementation of video signal generation, the transmission of this video signals, reception of this same video signal and finally the synchronisation to an industry standard output that can be interfaced to a display or storage medium of the user's choice.

In conclusion, the main aim of the project was met successfully with the 10G speed ultimately being reached. The process involved creating various designs where each design was an incremental step towards the final design. This involved firstly understanding the protocols that govern this field, and then moving onto a design phase. The design phase was taken in stages, starting from a 1G Media Access Controller (MAC) test design, then moving onto the main core, namely the 10G MAC. After that, the data generation was tackled and paired with the 10G MAC, with the receive interface being ultimately tackled. This receive interface helps with synchronising data to control signalling, and was implemented using additional IP cores and some custom written code modules.

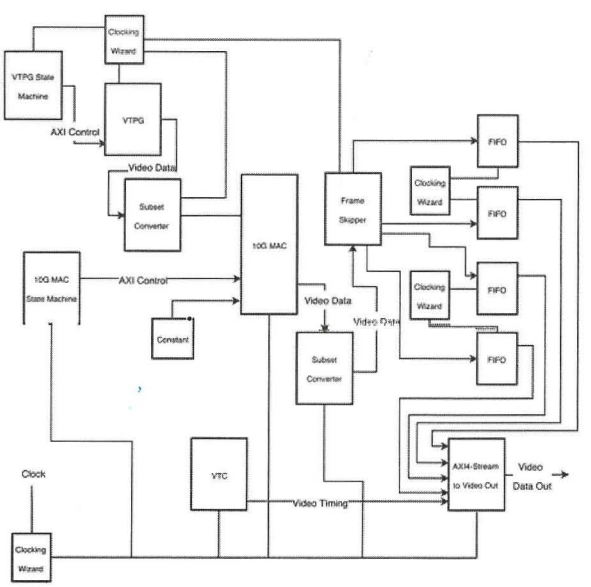


Figure 1: Binder Inter-Process Communication Flow

PROJECT NO. 35

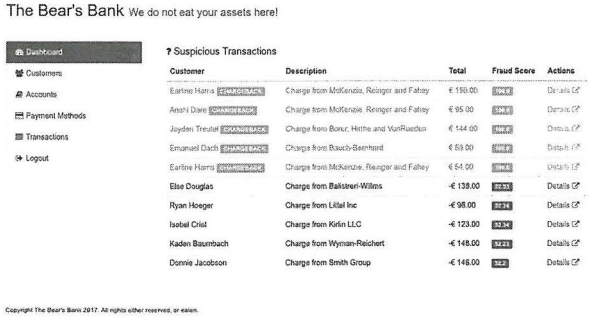
Constructing a Web Application for Identification, Detection and Analysis of Pattern Based Fraud

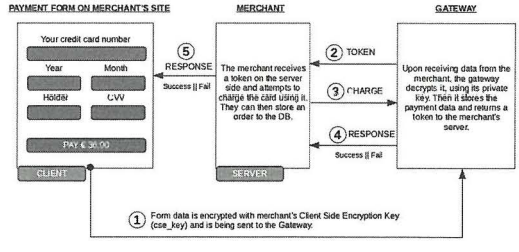
ANELIYA DICHEVA MAVROVA

SUPERVISED BY
MR ANTHONY SPITERI STAINES

The rapid development and popularization of the e-commerce business in the past few decades have made online purchasing of products preferable and more comfortable shopping method for buyers. As credit / debit cards became most common payment method for both online and regular purchases, the number of registered fraudulent transactions also increased significantly. Fraudsters seem to find back-doors in the security systems, leading to accumulation of losses for banks, merchants and customers. Therefore implementation of efficient fraud detection system has been made imperative for bank issuing systems in order to be able to predict suspicious behavior and stop fraud before it occurs. Modern techniques for fraud prevention include Machine Learning, Data Mining, use of Artificial Immune Systems etc. Based on the dataset and the present information various techniques could be used to develop security systems. This project has a goal to examine the transaction process for both genuine and fraudulent purchases of online products and using a Naive Bayes algorithm to be able to detect an flag the suspicious such. The algorithms selection is based on the dataset size (small dataset, since credit card details include sensitive data) and due to the high probability performance of Naive Bayes when data is not clustered or is partially missing.

For the scope of the project a gateway, a bank and a processor are simulated to observe behavior of credit / debit card charges. Data are being fed to the system to analyze it in order to detect suspicious transactions, trigger a fraud score an eventually stop fraud before it reaches the bank and affects merchant's of customer's account [7] [8] . This project develops a gateway and a bank along with customers, bank accounts, merchants and a processor. The standard gateway checks are divided into three groups: consistency, velocity and referral risk checks. [1] Each check is implemented through an algorithm triggering score. The goal of the scoring system is to decline a card transaction by reaching a limit of score 100, when the transaction is rejected by the gateway. On the bank side a machine learning is used to identify and flag fraudulent transactions using Naive Bayes algorithm. The algorithm is selected based on the size of the dataset used for the project and the type of data. Since credit card credentials are highly sensitive data, the database with customers and transactions is randomly generated and the size of the dataset is small (less than 2000 entries). Therefore the selected algorithm is Naive Bayes, because it is a suitable solution for classified and scattered data and in cases when some of them is missing or not full. [9]





*Transaction Workflow Schema

TOKEN – payment ID

REFERENCES

[1] (2017). *Payment Technology*. Available: https://ca-test.adyen.com/ca/ca/fraud/scoring.shtml?setActiveAccountKey =MerchantAccount.DNTradeLTDCOM.

[2] Kellie Bryan, Kristen Dunnesen, Jayson Jean, Elli Jellenc, Josh Lincoln, Michael Ligh, Mike La Pila, Ryan Olson, Andrew Scholnick, Greg Sinclair, Tom Wils and Kimbery Zenz, "Chapter 2: Cyber fraud: Principles, trends and mitigation techniques," in *Cyber Fraud: Tactics, Techniques and Procedures*, James Graham, Rick Howard, Ralph Thomas and Steve Winterfeld, Ed. Taylor and Francis Group, LLC, 2009, pp. 21.

[3] M. Unibul, "Credit Card Transaction Processing Basics," 2017. Available: http://blog.unibulmerchantservices.com/credit-card-transaction-processing-basics/

[4] (13 August 2013). The USA has more e-commerece fraud than Nigeris [E-commerce fraud]. Available: https://blog.siftscience.com/2013/global-e-commerce-fraud/.

[5] P. Braintree, "Fraud Tools: Overview," 2017. Available: https://articles.braintreepayments.com/guides/fraud-tools/overview

[6] P. DSS, "PCI DSS quick reference guide: Understanding the payment card industry," PCI Security Standard Council, LLC, October 2010.

[7] M. Meissner, "Making Payments Safer Top Priority for Global Payment Card Industry," 31 March 2017. Available: https://www.pcisecuritystandards.org/pdfs/Making-Payments-Safer-Top-Priority-For-Global-Payment-Card_Industry.pdf

[8] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, USA: Springer, 2006.

[9] (November 2016). Scikit-learn: *Machine Learning in Python* [Machine Learning]. Available: http://scikit-learn.org/stable/modules/naive_bayes.html.

PROJECT NO. 36

Multi-Dimensional Indexes in DBMSs

THOMAS MERCIECA

SUPERVISED BY
DR JOSEPH VELLA

Multi-dimensional data and operations are commonplace across multimedia and data-mining environments. A well-known and suitable indexing method is the R-Tree, which DBMSs are implementing into their core for efficient retrieval of multi-dimensional data. This structure groups data into subspaces using Minimum Bounding Rectangles but it does underuse space and overlap each other. Searching the R-Tree is thus a multipath problem, and the gap between the best and worst-case performance is very wide, making carefully crafted R-trees especially important.

The aim of this study is to investigate the R-Tree in a popular and advanced DBMS, PostgreSQL. Focus is directed on an index-building parameter as an optimisation parameter, referred to as *LIMIT_RATIO* in the DBMS, that controls the number of items per group of space. The challenge behind studying this parameter is that officially, it is only controllable through direct manipulation of the DBMS source code. This project also provides a suitable means of dealing with this overhead.

To undo the hard-coding of this parameter, the DBMS is unravelled and modified, creating a solution which controls the parameter from the SQL level. Then, the impact of *LIMIT_RATIO* is studied with respect to a set of metrics over an extensive and well-known spatial and multi-dimensional dataset and query workload. Metrics include time and storage costs of the index and query performance.

The SQL addition successfully underwent an extensive set of regression tests, ensuring correctness of its implementation. Through performance analysis, a set of values offering improved performance in terms of a specific metric are identified. Further, results suggest alternative configurations which are superior to PostgreSQL's hardcoded default. As a result of this project's implementation, data designers now have a more flexible and configurable set-up of R-tree based multi-dimensional data indexes.

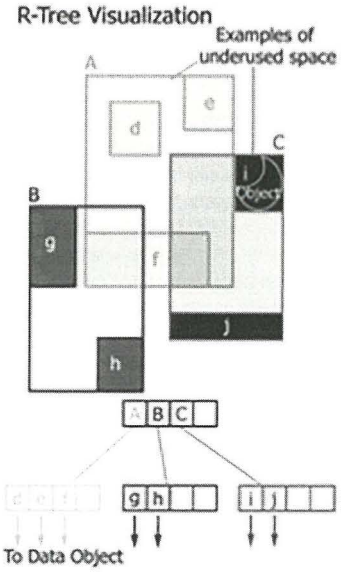


Figure 1: Two groupings of space considered by the R-Tree building process. Configuration of LIMIT_RATIO contributes to the rejection of one over the other.

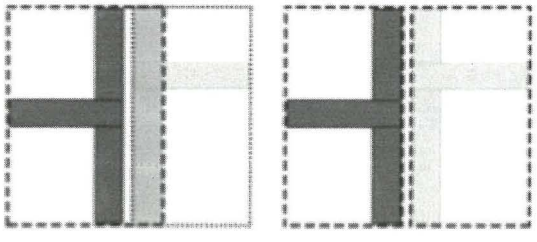


Figure 2: A visualization of a simple R-Tree built over a set of data objects.

Gamified Beacons

PROJECT NO. 80

DANIEL MIZZI

SUPERVISED BY
PROF. ALEXIEI DINGLI

The main aim of this project is to provide a platform which embeds knowledge into a gaming environment through the use of Bluetooth Low-Energy Beacons. Beacons are transmitters which send Bluetooth signals to be scanned continuously by mobile devices [1]. This will allow for a much more enjoyable and engaging learning experience, potentially also increasing the rate at which one is able to learn and memorise facts.

Such a platform will be achievable through the implementation of concepts referred to as gamification and alternate-reality. Gamification refers to the idea of introducing game concepts into non-game related contexts [2]. On the other hand, alternate-reality is a game genre with a story-driven approach and core mechanics mainly requiring collaborative problem-solving and storytelling [3].

As a result of reviewing and analysing various systems with their own implementations of gamification or alternate-reality, we were able to propose and develop a system of our own. For this project, we developed an Android application which takes into account aspects from both concepts. The application developed presents an alternate-reality game with a story-line based on the medieval times of the capital city of Malta, Valletta. Beacons act as an intermediary between factual and actual reality, portraying characters from the past which may be interacted and spoken with through the application itself by being in vicinity of the physical location of the Beacon.

The implementation, which was launched at the Faculty of Information and Communication Technology (at Level -1) in the University of Malta, provides the users with a single quest to be completed. Throughout the quest, users are taken through the process involved with the building of Valletta. This is achieved by referring to the steps indicated by the quest, which often finds the players having to travel from one position to another (within the Faculty) to either interact with different characters or pick up various objects. As they progress through, players are challenged on the facts which are being mentioned, putting more emphasis on the learning aspect. This resulted in a majority of the players learning new history facts, as backed up by the analysis conducted.

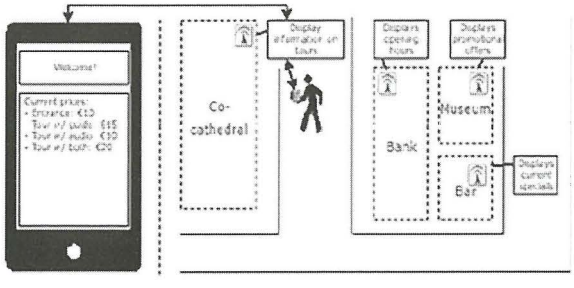


Figure 1: Example of an implementation using beacons - each location has a beacon set up, displaying its own information on the mobile devices of passers-by.

Figure 2: The main screen with the map and interactable entities on the left side, the steps required to complete the quest on the right.

Ultimately, the system provided may be downloaded and installed from the Google Play Store, by searching for 'Gamified Beacons'[1].

REFERENCES

[1] "Beacons: Everything you need to know. pointr blog." http://www.pointrlabs.com/blog/beacons-everything-you-need-to-know/. (Accessed on 12/13/2016.

[2] A. Marczewski, *Gamification: a simple introduction*. Andrej Marczewski, 2013.

[3] E. Bonsignore, D. Hansen, K. Kraus, J. Ahn, A. Visconti, A. Fraistat, and A. Druin, "Alternate reality games: platforms for collaborative learning," in *Proceedings of the 10th International Conference of the Learning Sciences*, ICLS 2012, vol. 1, pp. 251-258, 2012.

[1]https://play.google.com/store/apps/details?id=fyp.gamifiedbeacons\&hl=en

PROJECT NO. 56

Visualising and Interacting with Music-Theoretic Structures

SAMUEL MIZZI

SUPERVISED BY
DR KEVIN VELLA AND PROF. GORDON PACE

Music theory and mathematics originate from completely different points of view. However, it was shown that there are certain music-theoretic structures that are meaningful in science and mathematics [6]. This is one of the main reasons why there is an ongoing scientific study in music theory.

In the dissertation paper concerned, various musical terms and related properties were explained. In addition to this, particular focus was on musical scales, and exploring their geometric properties [3]. It was found [1] that some scales have maximal properties known as maximally even sets and maximal area sets [4]. Up till now, these properties were all demonstrated and computed manually on paper. The dissertation presents a software application tool that can help apprehend whoever is interested about this subject. The user of this application can input the desired parameters to visualise the said properties and results

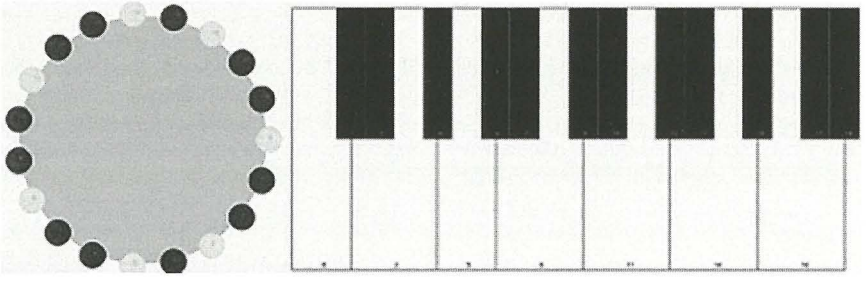


Figure 1: Software artefact visualisations of the clock diagram (semitone space) and piano keyboard

Moreover, this software application allows the user to choose between two different instruments: clock diagram [2] and piano keyboard. The two instruments can show the same results in a different visualisation (see also figure 1). In addition to this, another intriguing detail for the end user is the visualisation of microtonal music [5].

The software application was implemented for Android using mostly Java language (figure 2 show the design prior to the implementation of the Android application). Finally, the application was evaluated from both algorithmic and qualitative perspectives. The overall evaluation contributed to come up with ideas for future related work and improvements. One example of this is to make the Android application more entertaining for the user. This can be accomplished by injecting the aspect of gamification into the application [7].

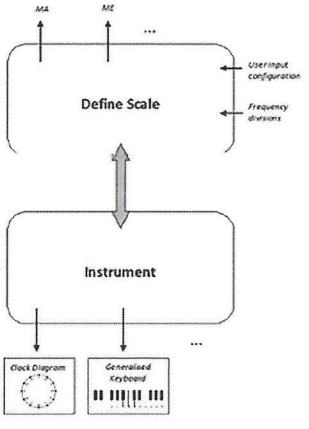


Figure 2: High level view of the Android application design stage

REFERENCES

[1] Jack Douthett. Maximally even sets. Journal of Music Theory, 35:93{173, 1991.
[2] Kevin Vella Gordon J. Pace. Tipping the scales...functionally! Computer Science Annual Workshop. University of Malta, 2014.
[3] David Rappaport. Musical scales, integer partitions, necklaces, and polygons. In BRIDGES: Mathematics, Music, Art, Architecture, Culture, pages 595{598, London, 2006.
[4] David Rappaport. Maximal area sets and harmony. Graphs and Combinatorics, 23(1):321{329, 2007.
[5] Thanassis Rikakis. A pratical microtonal pitch space; theoretical and psychoacoustic issues, compositional applications. https://ame2.asu.edu/faculty/thanassis/papers/microsystem.pdf, 2001 Microfest Conference and Festival. Harvey Mudd Collage, CA, April 2001
[6] Saloni Shah. An exploration of the relationship between mathematics and music. http://eprints.ma.man.ac.uk/1548/01/Saloni2010.pdf, 2010.
[7] Ge Wang. Game design for expressive mobile music. https://nime2015.lsu.edu/proceedings/143/0143-paper.pdf. Stanford University, 2015.

PROJECT NO. 61

An Assessment of the Impact of Human Computer Interface on Maltese Mobile Users

LARINOV MUSCAT

SUPERVISED BY
DR COLIN LAYFIELD

PURPOSE

With the development of new and improved information and communication technologies, the issue of Human Computer Interaction (HCI) has taken centre stage. Moving away from command-line and graphical user interfaces and more towards perceptual user interfaces, customer experience has faced over recent years significant challenges, particularly in terms of adapting to various new functionalities [1]. In view of this, this study sets out to critically assess the impact of human computer interaction on mobile phones amongst Maltese youth, given that this is the technology they often use to access the Web [2]. In addition, the study seeks to determine characteristics and qualities that need to be present in website mobile-based interfaces that make them effective to mobile youth and consequently make recommendations to developers based on the results obtained.

DESIGN

In order to reach the objectives identified, an e-commerce interface was selected to be the basis of three main iterations of focus groups. In the first iteration, paper prototyping and user sketching were used to assess youth's preferences amongst four primary designs developed by the researcher [3], [4]. A total of 18 participants from this target cohort took part at this stage. Based on the feedback obtained, a second iteration involving usability testing amongst 5 test subjects was conducted to iron out any issues that persisted [5]. An e-commerce mobile website was then developed according to the preferences expressed during the previous two iterations. A further 40 participants then took part in focus group sessions in which they were asked to compare and contrast the interface developed by the researcher with that of a popular Maltese e-commerce website – Maltapark.

FINDINGS

From the three iterations of usability testing conducted, it was clearly observed that Maltese youth give considerable importance to user experience when using mobile websites on their mobile phones. In particular, certain characteristics emerged as being key in their decision as to whether a particular mobile website has an optimal user experience or otherwise. To this end, Maltese youth participating in the sessions preferred mobile websites that did not include an overflow of information on screen, that were customised to the size of their phones' screen, that involved an easy and logical structure to navigate from one page to another and that were provided useful features such as real-time validation, progress bars and password strength indicators. Conclusion: The major conclusion drawn from this study is that a user-centric approach should be adopted for user interface design whereby end users are continuously engaged and asked for their feedback not after but during the development phase of a particular software [3], [6]. In addition, companies should not underestimate the significance attributed to user experience and should dedicate time and resources to ensure an optimal UX [6].

VALUE

This study has shed light on the importance of UX for the popularity and success (or otherwise) of a particular website. As users increasingly use their mobile phones to access the Internet, enterprises should well consider adapting their websites to incorporate more mobile-friendly features.

REFERENCES

[1] O. K. Shoukry and M. B. Fayek, "for Mobile Content Pre-fetching," no. May 2007, pp. 228–239, 2013.
[2] M. Sue and A. Scott, "NSO Household Internet Usage Statistics 2015," no. February, pp. 2–9, 2015.
[3] M. Tohidi, W. Buxton, R. Baecker, and A. Sellen, "Getting the Right Design and the Design Right : Testing Many Is Better Than One," *Proc. CHI 2006 Conf. Hum. Factors Comput. Syst.*, pp. 1243–1252, 2006.
[4] M. Tohidi, W. Buxton, R. Baecker, and A. Sellen, "User Sketches: A Quick, Inexpensive, and Effective way to Elicit More Reflective User Feedback," *NordiCHI*, no. October, pp. 105–114, 2010.
[5] J. R. Lewis, "Usability Testing," *Handbook of Human Factors Testing and Evaluation*. pp. 1275–1316, 2006.
[6] J. Moule, *Killer Ux Design*. 2012.

PROJECT NO. 4

Image Processing Techniques for Brain Haemorrhage Detection in Head CT Scans

JOHN NAPIER

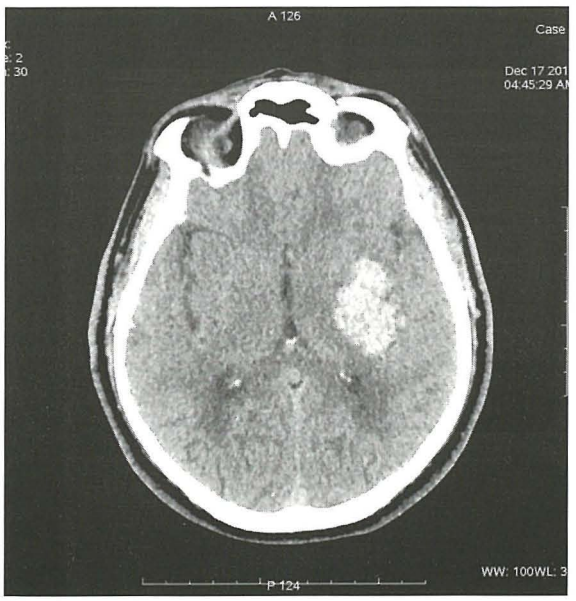
SUPERVISED BY
PROF. ING. CARL JAMES DEBONO, DR PAUL BEZZINA AND DR FRANCIS ZARB

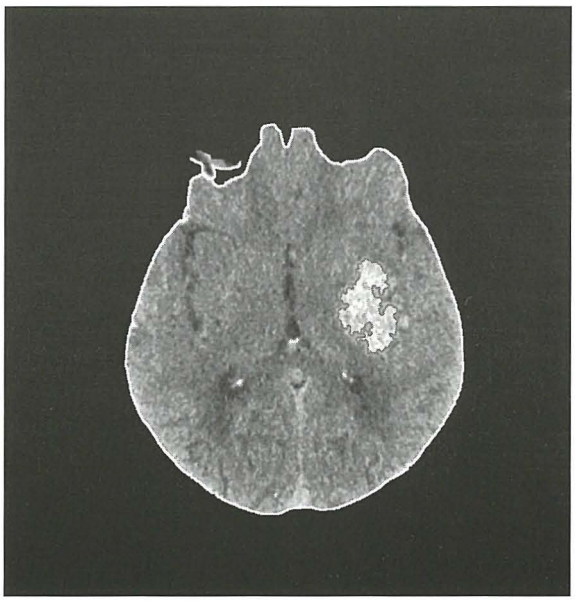
Medical imaging is an important tool used for obtaining visual information of the interior of the human body. There exist various imaging modalities such as Magnetic Resource Imaging, Ultrasound and Computed Tomography (CT). The latter is used extensively for detection and diagnosis of brain haemorrhage.

Computer Aided Detection/Diagnosis (CAD) systems are used by radiologists as a tool which helps them during the diagnosis phase. Increasingly, CAD is becoming a key component of routine clinical practice in several medical areas such as mammography and colonoscopy. Research regarding brain CAD systems has not progressed at the same pace as research in the mentioned areas. This provides a need as well as an opportunity to contribute to the research by developing a CAD system for brain haemorrhage detection and classification.

This study applies image processing techniques to brain CT scans with the aim of creating a CAD system which detects fresh bleeds. The system also includes a basic classification that identifies the haemorrhage type. The system distinguishes between an intra-axial haemorrhage and an extra-axial haemorrhage with the only limitation being subarachnoid haemorrhage (SAH), which is not always properly classified due to its complex structure. The techniques implemented in this study include noise reduction methods through the use of filters, morphological operations and segmentation algorithms, such as thresholding and clustering.

The designed CAD system was tested on 36 brain CT sets obtained from the general hospital in Malta. The results show that the system achieved a sensitivity of 94.4%, a specificity of 94.4%, a precision of 91.259% and a classification accuracy of 88.89%. The system performed the required operations in an average time of 0.26 seconds per slice.





PROJECT NO. 7

The Implementation of the SATA Protocol on an FPGA

MATTHEW SACCO

SUPERVISED BY
DR ING EDWARD GATT AND MR ANDRE MICALLEF

Medical imaging is an important tool used for obtaining visual information of the interior of the human body. There exist various imaging modalities such as Magnetic Resource Imaging, Ultrasound and Computed Tomography (CT). The latter is used extensively for detection and diagnosis of brain haemorrhage.

Computer Aided Detection/Diagnosis (CAD) systems are used by radiologists as a tool which helps them during the diagnosis phase. Increasingly, CAD is becoming a key component of routine clinical practice in several medical areas such as mammography and colonoscopy. Research regarding brain CAD systems has not progressed at the same pace as research in the mentioned areas. This provides a need as well as an opportunity to contribute to the research by developing a CAD system for brain haemorrhage detection and classification.

This study applies image processing techniques to brain CT scans with the aim of creating a CAD system which detects fresh bleeds. The system also includes a basic classification that identifies the haemorrhage type. The system distinguishes between an intra-axial haemorrhage and an extra-axial haemorrhage with the only limitation being subarachnoid haemorrhage (SAH), which is not always properly classified due to its complex structure. The techniques implemented in this study include noise reduction methods through the use of filters, morphological operations and segmentation algorithms, such as thresholding and clustering.

The designed CAD system was tested on 36 brain CT sets obtained from the general hospital in Malta. The results show that the system achieved a sensitivity of 94.4%, a specificity of 94.4%, a precision of 91.259% and a classification accuracy of 88.89%. The system performed the required operations in an average time of 0.26 seconds per slice.

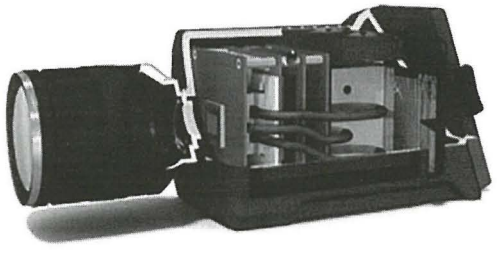


Figure 1: CAD model of the camera product (section view)

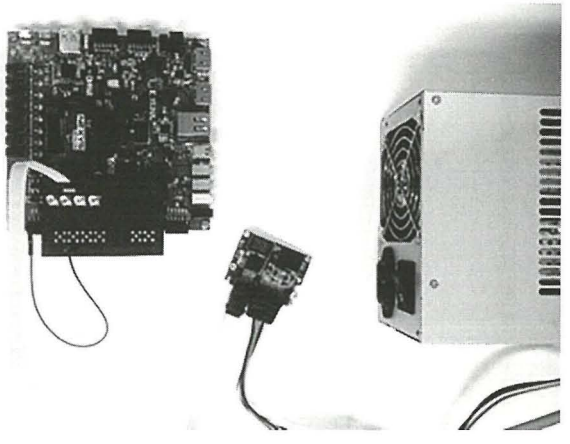


Figure 2: Hardware setup while testing

REFERENCES

[1] M. A. Azzopardi, "Method and apparatus for generating and transmitting synchronised video data". Malta Patent EP2332339 A2, 15 June 2011.

[2] SATA, Serial ATA Revision 3.0, Serial ATA International Organisation, 2009

PROJECT NO. 31

Investigating Search-based Test Generation for Web Applications

MALCOLM SALIBA

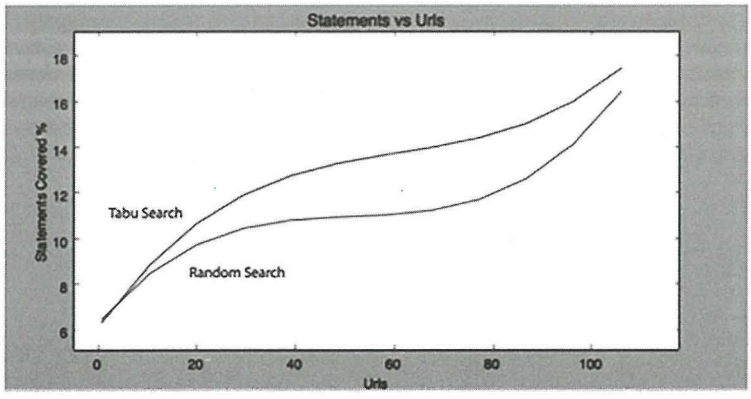
SUPERVISED BY
DR MARK MICALLEF

Medical imaging is an important tool used for obtaining visual information of the interior of the human body. There exist various imaging modalities such as Magnetic Resource Imaging, Ultrasound and Computed Tomography (CT). The latter is used extensively for detection and diagnosis of brain haemorrhage.

Computer Aided Detection/Diagnosis (CAD) systems are used by radiologists as a tool which helps them during the diagnosis phase. Increasingly, CAD is becoming a key component of routine clinical practice in several medical areas such as mammography and colonoscopy. Research regarding brain CAD systems has not progressed at the same pace as research in the mentioned areas. This provides a need as well as an opportunity to contribute to the research by developing a CAD system for brain haemorrhage detection and classification.

This study applies image processing techniques to brain CT scans with the aim of creating a CAD system which detects fresh bleeds. The system also includes a basic classification that identifies the haemorrhage type. The system distinguishes between an intra-axial haemorrhage and an extra-axial haemorrhage with the only limitation being subarachnoid haemorrhage (SAH), which is not always properly classified due to its complex structure. The techniques implemented in this study include noise reduction methods through the use of filters, morphological operations and segmentation algorithms, such as thresholding and clustering.

The designed CAD system was tested on 36 brain CT sets obtained from the general hospital in Malta. The results show that the system achieved a sensitivity of 94.4%, a specificity of 94.4%, a precision of 91.259% and a classification accuracy of 88.89%. The system performed the required operations in an average time of 0.26 seconds per slice.



REFERENCES

[1] M. Harman, "The current state and future of search based software engineering," Future of Software Engineering, vol. 2007, pp. 342–357, 2007.

[2] N. Alshahwan and M. Harman, "Automated web application testing using search based software engineering," Automated Software Engineering, vol. 2011 26th IEEE/ACM International Conference, pp. 3–12, 2011.

[3] P. McMinn, "Search-based software testing: Past, present and future," Interna- tional Conference on Software Testing, Verification and Validation Workshops (ICSTW), vol. 2011, pp. 153–163, March 2011

PROJECT NO. 58

E-Voting System using Blockchain Technology

DYLAN SCHEMBRI

SUPERVISED BY
DR JOSEPH VELLA AND DR ABDALLA KABLAN

Electronic voting encompasses all means which aid to speed up and increase the accuracy of the vote casting or vote counting process through automation. Various countries have already implemented an electronic voting system to be used in general elections; a case in point is Estonia who was one of the pioneers in introducing electronic voting through the use of the Internet [1].

Malta has been yearning for the introduction of an electronic voting system to help overcome the long and tedious vote counting and vote casting process. From last election's statistics it was deduced that there were over 4,000 votes which were deemed as invalid [2]. In an article published by the Malta Independent, it was mentioned that the 2017 elections would be the last to use manual counting and that an electronic vote counting procedure is to be introduced in the next election [3][4]. The way forward would then be to introduce an electronic vote casting system, so that certain inaccuracies introduced when writing down the vote by hand are completely eliminated. There would also be no need of transporting the votes from the polling booths to the counting hall because they would be in electronic format.

The system being proposed is an electronic vote casting system which uses a secure technology called blockchain, for storing and moving the electronic votes. This ensures that none of the ballots are tampered with and that the votes are delivered to the vote counting hall in a safe and secure manner whilst guaranteeing voting anonymity – see figure 1. Initially a blockchain was created to serve as an environment where financial transactions for the bitcoin cryptocurrency can be executed while eliminating the problem of double spending [5]. However, it has now evolved to be used beyond its financial application, attributed to the fact the data is shared between all the users in the blockchain rather than having a single central server, giving a small window of opportunity for tampering with the system.

Therefore, a blockchain based e-voting system makes a good candidate amongst the other e-voting systems in providing a good alternative to traditional voting methods.

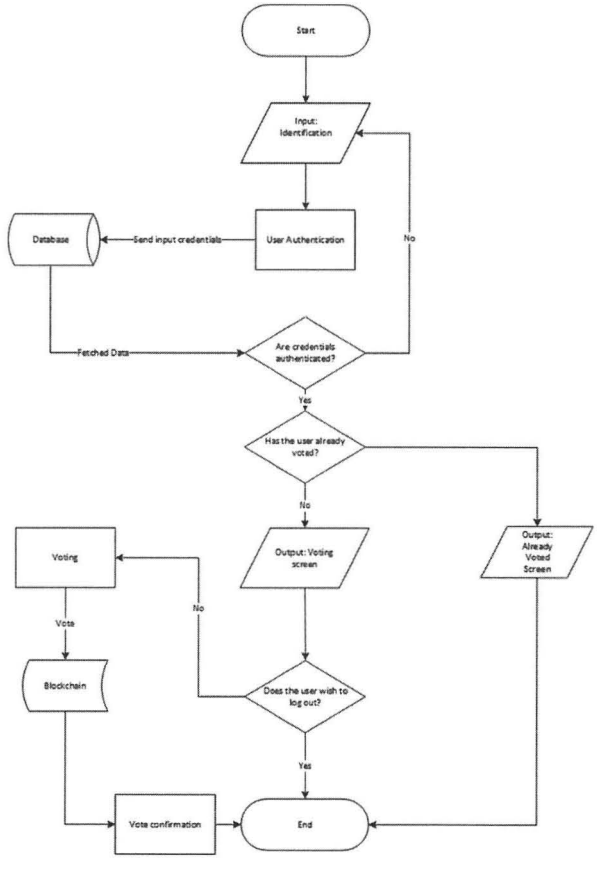


Figure 1: A diagram illustrating a high level visualisation of the system.

REFERENCES

[1] "Estonia pulls off nationwide Net voting - CNET News." [Online]. Available: http://archive.is/20120713045721/http://news.com.com/Estonia+pulls+off+nationwide+Net+voting/2100-1028_3-5898115.html. [Accessed: 22-May-2017].

[2] "The Voters - Malta Elections." [Online]. Available: http://www.um.edu.mt/projects/maltaelections/maltesepolitics/voters. [Accessed: 23-May-2017].

[3] "2017 election will be last one with manual counting - The Malta Independent." [Online]. Available: http://www.independent.com.mt/articles/2017-05-31/local-news/2017-election-will-be-last-one-with-manual-counting-6736174953. [Accessed: 04-Jun-2017].

[4] "Electronic vote counting for all elections from 2019 - The Malta Independent." [Online]. Available: http://www.independent.com.mt/articles/2017-04-16/local-news/Electronic-vote-counting-for-all-elections-from-2019-6736173045. [Accessed: 04-Jun-2017].

[5] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," www.bitcoin.org, p. 9, 2008.

PROJECT NO. 79

Question Answering using Wikipedia

YURGEN SCHEMBRI

SUPERVISED BY
DR JOEL AZZOPARDI AND DR CHRIS STAFF

Unstructured and semi-structured information sources written in natural language, such as Wikipedia[1], make available huge amounts of factual information. With such information sources and the availability of powerful search engines, a user may easily find documents relevant to a query. However, retrieving a particular piece of information from these documents may be time consuming. This is tackled in the question answering problem; the process of using computers to find succinct answers to questions posed in a natural language [1].

In this work, a scalable open-domain question answering system is designed and developed. While this task has been widely studied in the past, we constrain our approach to using Wikipedia, the free encyclopedia, as the exclusive source of information. We present a robust architecture (Figure 1) which can be expanded beyond a few types of questions.

The focus is mainly on factoid questions and yes/no questions, but other types, such as decision questions, are also studied. This goes beyond the usual research methods in question answering as, commonly, a study focuses on a single question type. We use only the documents' plain text to answer questions, but the implementation allows for other features to be added. These features, such as utilising the categories or infoboxes of Wikipedia pages, are usually studied in isolation.

We deal with several sub-tasks in the question answering problem, namely question focus extraction, question classification, information retrieval and answer extraction and selection. The question focus describes the main information required by the question [2]. In question classification, the question is assigned a class in a hierarchical taxonomy which is later used to filter out unreasonable answers [3]. Keyword queries are then generated from the question and Wikipedia is searched for relevant pages through the MediaWiki API[2]. The top-ranking document is used to find the answer. This is divided in paragraphs which are then ranked and the segment most likely containing the answer is chosen [4]. The rest of the process depends on the type of question at hand. Factoid and decision questions require extracting possible answers from text, and selecting the best one after filtering and ranking [1, 4]. When answering yes/no questions, on the other hand, the system attempts to find pieces of the text which support the question [5].

The developed modules are evaluated separately. The results show that the focus extraction and document retrieval modules can be improved further as their performance is not satisfactory. The results for answer selection show that the proposed methods achieve adequate performance on simple questions and yes/no questions. In addition, extra effort put into the question classification task contributed to achieving state-of-the-art performance with room for further improvement.

1 https://en.wikipedia.org

2 https://www.mediawiki.org/wiki/API:Main_page

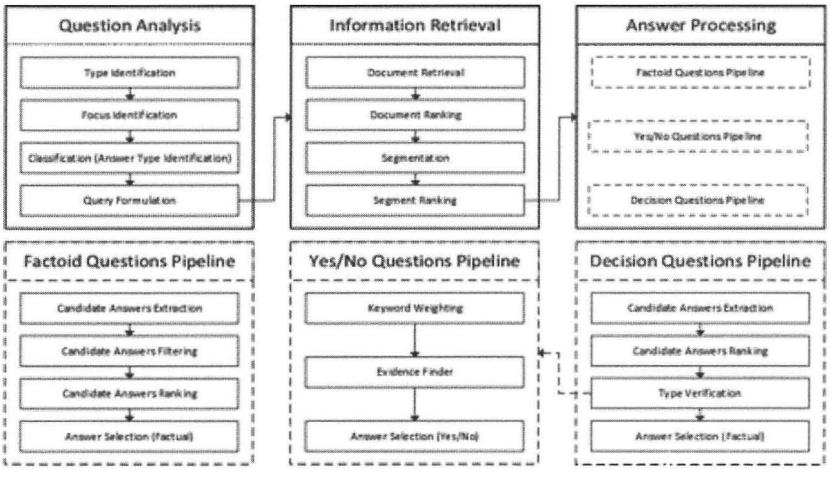


Figure 1 The global architecture (top) and the pipelines specifically designed for three different question types (bottom). Integration between the three question type solvers is also shown: the decision questions solver depends on the yes/no question solver.

REFERENCES

[1] D. Jurafsky and J. H. Martin, "Question answering," in *Speech and Language Processing*, ch. 28, Pearson Education India, 3 ed. Draft of November 6, 2016.

[2] D. Moldovan, S. Harabagiu, M. Pasca, R. Mihalcea, R. Girju, R. Goodrum, and V. Rus, "The structure and performance of an open-domain question answering system," in *Proceedings of the 38th Annual Meeting on Association for Computational Linguistics*, pp. 563-570, 2000.

[3] X. Li and D. Roth, "Learning question classifiers," in *Proceedings of the 19th International Conference on Computational Linguistics*, pp. 1-7, 2002.

[4] C. Kwok, O. Etzioni, and D. S. Weld, "Scaling question answering to the web," *ACM Transactions on Information Systems*, vol. 19, no. 3, pp. 242-262, 2001.

[5] Y. Kano, "Answering yes-no questions by penalty scoring in history subjects of university entrance examinations," OKBQA 2016, pp. 91-96, 2016.

PROJECT NO. 60

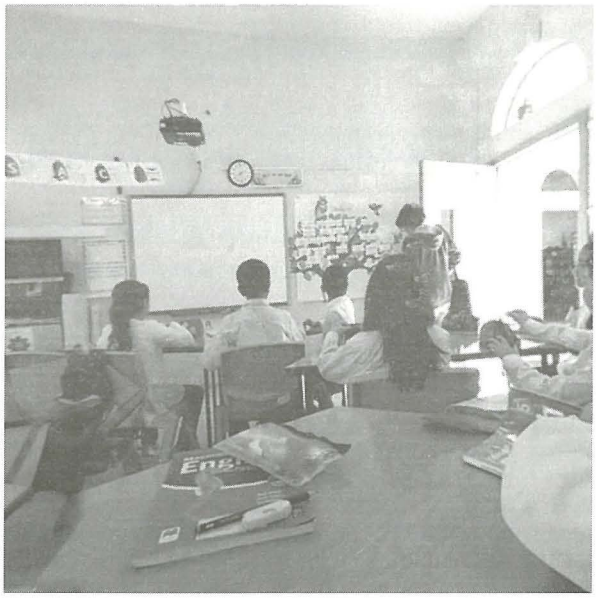
Addressing Cross Cultural Integration Through a Virtual Reality Experience

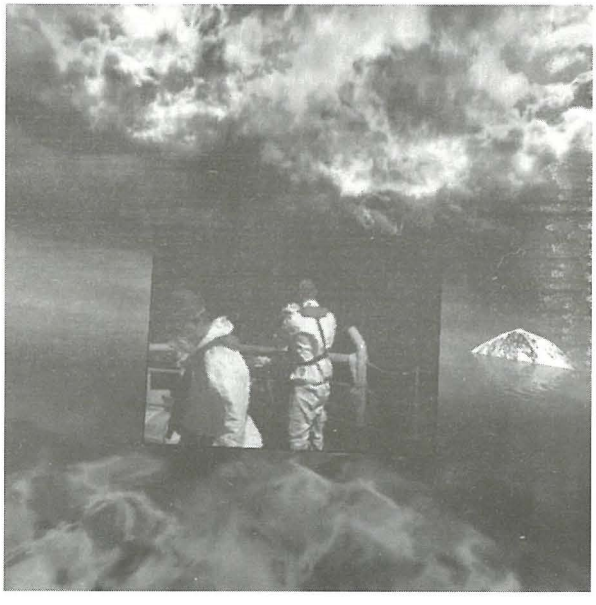
DAVID SCICLUNA

SUPERVISED BY
PROF. ALEXIEI DINGLI, DR VANESSA CAMILLERI AND DR COLIN LAYFIELD

Classes are becoming increasingly multi ethnic and it is often the case that teachers may not be effectively trained to handle such new multiethnic classes. Our project involves the development of an immersive virtual reality experience for teachers. This will take the user through actual experiences of a child from a migrant background with the aim of supporting empathy. Therefore, the VR headset can easily become an empathy machine capable of influencing the perception of the users. The success of such systems is centered around the principles of immersion and interaction. The adequate use of both will determine the overall outcome and effectively define the impact of the virtual reality experience. Through using the appropriate methodologies alongside the necessary hardware, it was envisioned that by the end of the project a fully working prototype would be created. This could then be evaluated to see whether any effect was present on the user.

The main objectives of the project were to find new innovative techniques to combine interaction with different types of user interfaces, which would essentially recreate empathy. The latter were mainly achieved throughout the project Evaluation of the latter yielded very interesting results. It could be seen that the presented experience provided the necessary level of immersion that it altered the user's emotions and perception. This was an indicator that the hypothesis can be proven through further testing and analysis. All possible future methods for testing were duly listed and their relevance explained. Alongside this all possible future improvements to the experience were clearly listed. These improvements would increase the overall quality of the experience and improve the effect it can have on users. Further to this all it was concluded that virtual reality can have effect on user perceptions but further study must be done into the extent of such effect.





PROJECT NO. 1

FPGA Based Gravitational Multi-Particle Simulator

GLENN SCIORTINO

SUPERVISED BY
DR ING. OWEN CASHA

The Multi-Particle or N-body simulation problem has been pursued by numerous scholars in the last century. Nowadays particle simulators are considered to be highly valuable tools in several fields of Physics, particularly that of Astronomy. The idea behind the N-body problem is to find the most time and resource efficient method to simulate the gravitational interaction of a number of particles in a given space. In particular, the position of each particle is continuously predicted by considering the particle's current velocity and the forces exerted on that same particle by the other particles. There are various ways to achieve this, with the most basic approach being the brute-force approach which compares each particle with every other particle in the given space for each time-step. An optimised approach towards solving the N-body problem is the Barnes-Hut algorithm that makes use of a quadtree data-structure.

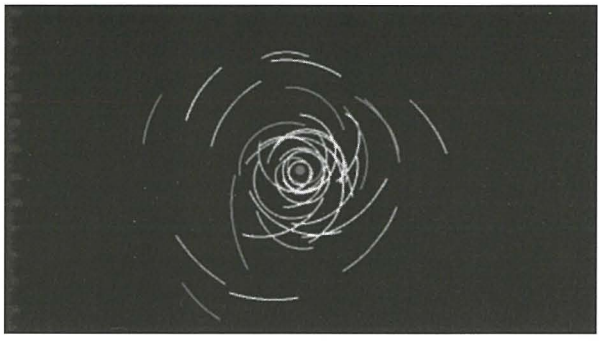


Figure 1: Screen shot of the brute-force simulation with 50 particles.

The multi-particle problem in this dissertation has been tackled from two different perspectives: a software and a hardware perspective. For the scope of this project the brute-force approach was implemented as both a software and a hardware version, while the Barnes-Hut was only implemented as a software version. The software implementation consisted of a simulator built from scratch in Python, utilising custom packages such as the quad-tree data-structure. The basic brute-force algorithm compares each particle with every other particle while the Barnes-Hut makes use of the quad-tree, to improve computational times at the expense of losing some degree of accuracy since it only considers particles which are in the close neighbourhood. Figure 1 shows a screen shot of the brute-force simulation with 50 particles, where the colours of the orbiting paths are indicative of the mass of the respective particle. On the other hand the hardware implementation consists of a brute-force simulator implemented on an FPGA platform using VHDL.

The idea behind this project was to investigate the difference between the two implementations in terms of computational efficiency. Each implementation was approached using a different mind-set. The software implementation required a procedural and structural approach even though object oriented principles were applied, whilst the hardware implementation required a more state driven and concurrent approach. In the software implementation the N-body problem was solved using a rather highly abstracted understanding of the logic with minimal to no particular interest of how it is implemented at a lower level. This is in stark contrast with the hardware implementation, where the hardware dictated how the simulator was to be built, so a low level understanding was mandatory and a priority.

The results for the software implementation showed that the Barnes-Hut algorithm offers significant improvement over the basic brute-force approach in terms of computational speed. This is primarily due to the efficient quad-tree data-structure implementation, with fast traversal times. It was also concluded that the FPGA is not necessarily the best option in terms of platform, even though it might sound counter intuitive. This is because if parallelisation is not utilised, the problem becomes quadratic and computational times become large in a very short amount of time. In the hardware implementation computation times are dependent on the propagation of the values through the logic and the intensiveness of the calculations, rather than the inputs themselves.

REFERENCES

[1] E. Holmberg, "On the clustering tendencies among the nebulae. ii. a study of encounters between laboratory models of stellar systems by a new integration procedure." *The Astrophysical Journal*, vol. 94, p. 385, 1941.

[2] S. J. Aarseth, *Gravitational N-body simulations*. Cambridge University Press, 2009.

[3] J. Barnes and P. Hut, "A hierarchical o(n log n) force-calculation algorithm," *Nature*, vol. 324, no. 6096, pp. 446–449, 1986.

[4] R. Hockney and J. Eastwood, "Computer simulation using particles," 1988.

[5] S. L.W. McMillan and S. J. Aarseth, "An o(nlogn) integration scheme for collisional stellar systems," *The Astrophysical Journal*, vol. 414, p. 200, 1993.

[6] "FPGA market size, share, growth." 2017. [Online]. Available: http://www.grandviewresearch.com/industry-analysis/fpga-market

[7] "Introduction to FPGA technology: top 5 benefits - national instruments 2017," 2017. [Online]. Available: http://www.ni.com/white-paper/6984/en/

[8] P. J. Ashenden, *The designer's guide to VHDL*, 1st ed. Morgan Kaufmann Publishers, 2011.

[9] "VGA simulator," 2017. [Online]. Available: http://ericeastwood.com/lab/vga-simulator/

PROJECT NO. 8

Sound Localisation

MARC ALAN SPITERI STAFRACE

SUPERVISED BY
PROF. ING. VICTOR BUTTIGIEG

This dissertation aims to implement a solution for sound localisation of deterministic signals, trading off between cost and result accuracy, by using industry-standard equipment and setup. The solution must provide a real-time result: the code behind the implementation must be light enough to be able to keep up with this requirement. The system consists of four microphones which are pre-amped, and fed through a DAC, where MATLAB analyses the sound windows captured and performs tests based on the time delay between microphone pairs. Direction and distance values are derived using two sorts of tests, geometric and Cartesian, over two topologies: closely grouped microphone pairs, and an array of equally spaced four microphones which are then paired using all possible permutations as shown in Figure 1.

The best-performing test is the one using closely-located pairs and a Cartesian analysis, giving a result in 2.107s with fairly lightweight code. Using a BPSK-modulated Barker-13 sequence, the system returns values as shown in Figure 2, with mean percentage errors of 28.5% on directional values and 16.04% on distance values. Going a step further and testing using a real-life signal, the Cartesian analysis on both topologies interestingly return almost equal result as shown. The inaccuracies in the results obtained are mostly due to the presence of reverberations as well as human errors during direction and distance measurement.

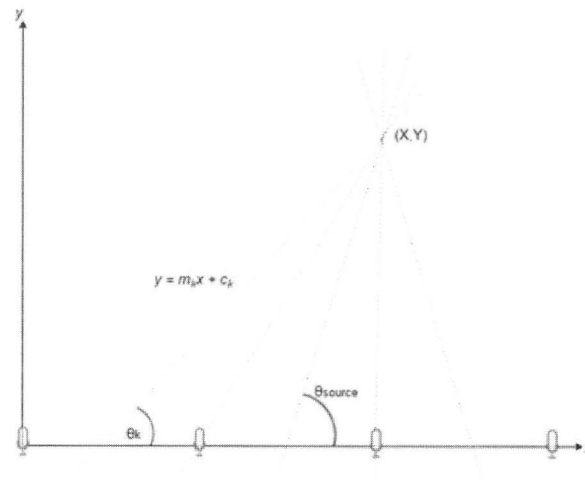
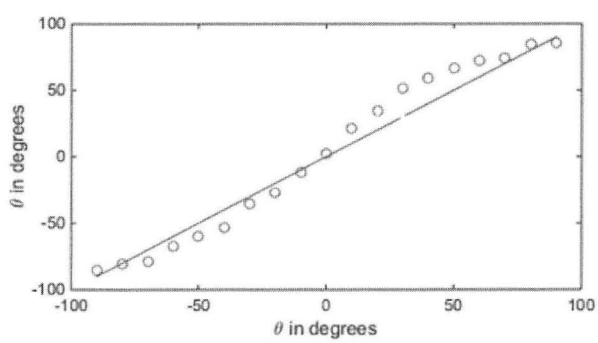


Figure 1: Setup for a generally-spaced microphone topology, mapped over Cartesian 2D-space



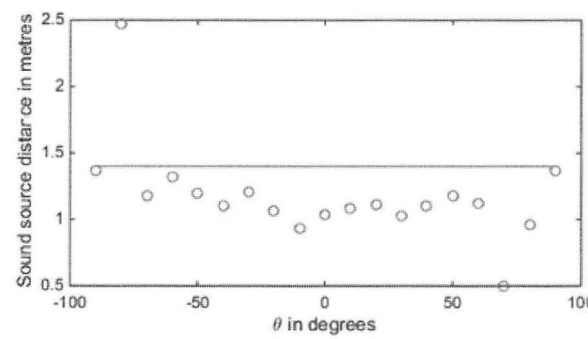


Figure 2: Results for the first topology using Cartesian mapping

PROJECT NO.
78

Football Scouting System

SANJAY SWAMY PRASAD

SUPERVISED BY
DR CHARLIE ABELA

The research carried out focused on the development of a weighted ranking mechanism for football players. The purpose of this was to find out how good a player is, by analysing the player's abilities according to his playing position. Since all players are unique and have different strengths, the players were ranked based on their respective qualities by making use of performance data.

Through the use of the FIFA dataset [1], an online tool was implemented to allow football scouts to easily analyse and discover players who may be of interest. This dataset was collected by EA Sports [2] through factual data and is used in the FIFA video game series on a yearly basis [1]. The dataset is very rich and also consists of player data from 2007 up till 2016. All the players were categorised according to their playing position obtained through an additional source, Futwiz [3]. This source was also used to obtain other attributes which were missing from the dataset for each player, such the team and nationality.

Data analysis was performed by analysing a variety of attributes found in the dataset, such as dribbling and stamina. All attributes which were found to be important, for each position based on relevant research, were used for the ranking mechanism by performing a weighted average on the selected attributes. Additional attributes relevant to each position were discovered and selected by analysing positive associations through the Pearson correlation coefficient. This was applied by analysing all correlation coefficient of over 0.70 and then including them for the weighted average. The ranking mechanism itself was developed by assigning a weighting to all the selected attributes. The weightings reflected the attribute's respective importance for a particular position and was derived by calculating the geometric mean for each attribute and then normalising each weighting out of 100.

An overall ranking out of 100 for each player was generated by applying the attribute weightings in the weighted average method. The concept of weighted average was applied also in [2] to rank cricket players based on raw player data collected. Similar research conducted by [3] used a weighted average to find the most important attributes for playing positions, whilst [4] used a similar approach to rank Premier League players on a weekly basis through the use of attribute weightings.

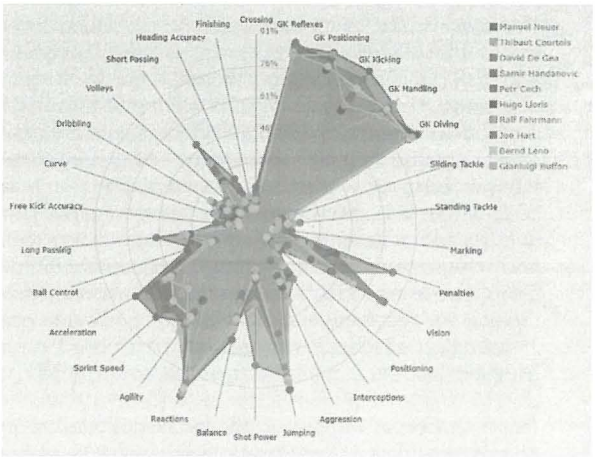


Figure 1: Player comparisons through a radar chart

Data visualisations such as radar charts (Figure 1) and multi-line charts were used to visualise the players, their ranking and their performance. We also include several visualisations in relation to player demographics. The aspect of real-time data was introduced by integrating statistics from Opta's [4] Twitter feeds into the system.

The player rankings were evaluated for each position using Root-Mean Square Error with the FIFA and Football Manager [5] datasets. A comparison was performed between the system's ratings and the ratings of FIFA and Football Manager. The results in comparison with the FIFA dataset were quite satisfactory for all positions, whilst the results for the Football Manager dataset varied accordingly for different positions.

1 https://www.kaggle.com/hugomathien/soccer
2 https://www.easports.com/fifa
3 http://www.futwiz.com
4 http://www.optasports.com
5 http://www.footballmanager.com

REFERENCES

[1] L. Cotta, P. O. V. de Melo, F. Benevenuto, and A. A. Loureiro, "Using fifa soccer video game data for soccer analytics," in *Workshop on Large Scale Sports Analytics*, 2016.

[2] S. Akhtar, P. Scarf, and Z. Rasool, "Rating players in test match cricket," *Journal of the Operational Research Society*, vol. 66, no. 4, pp. 684-695, 2015.

[3] H. S. Hossain, M. A. A. H. Khan, and N. Roy, "Soccermate: A personal soccer attribute profiler using wearables," in Pervasive Computing and Communications Workshops (PerCom Workshops), 2017 IEEE International Conference on, pp. 164-169, IEEE, 2017.

[4] I. G. McHale, P. A. Scarf, and D. E. Folker, "On the development of a soccer player performance rating system for the english premier league," *Interfaces*, vol. 42, no. 4, pp. 339-351, 2012.

GUI Toolset for a NoSQL DBMS

PROJECT NO. 30

LUKE THEUMA

SUPERVISED BY
DR JOSEPH VELLA

This dissertation presents a tool whose aim is to facilitate usage, control and management of a graph based data model maintained by Neo4j [1]. The tool's objective is to allow a developer using Neo4j to effectively and efficiently adopt it by attenuating the latter's reputed steep learning curve. This learning curve is mentioned in a paper by Peinl and Holzschuher [2].

The strengths and weaknesses of graph databases in general and those particular of Neo4j are given. In particular data and querying modelling details are presented and evaluated. A case is also made for graph visualisation especially since a graph database can be schema less [3]. Also an analysis of known tools, mainly from RDBMSs, that have similar aims to ours was undertaken and their essential features are elucidated. Software development tools and libraries that offer a contribution to the development of our tool's functionality in handling graphs and interactive visualization of graphs are surveyed.

After stating the project's functional and non-functional requirements of the solution, an analysis exercise yielded a number of decisions. The most important being to use a wizard to generate Neo4j Cypher code, and also to tie graph visualization with the tool's activities. Also elaborate import and export facilities have been identified and implemented. Much of the software development required here uses Neo4j and the other tools surveyed in the literature review.

Five main components were then designed and implemented successfully (see Figure 1), namely: creating an environment for access to Neo4j; executing queries and visualizing their results; importing and exporting of graph data; an interactive visualization tool; and a query builder with additional wizard option.

Finally the solution was tested by methods of verification and validation. Verification was carried out by gathering a set of known queries, generating them using our solution, and checking them for correct syntax and output. On the other hand, validation was performed by looking up a list of known types of queries and checking how many of these types our solution covers.

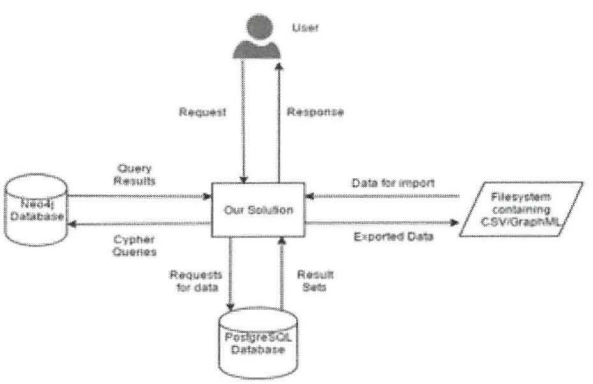


Figure 1: Overall design of the solution

Figure 2: Example of a visualization

REFERENCES

[1] Neo4j, "Neo4j," [Online]. Available: https://neo4j.com/.

[2] F. Holzschuher and P. D. R. Peinl, "Performance of Graph Query Langauges," in *EDBT/ICDT*, Genoa, 2013.

[3] I. Robinson, J. Webber and E. Eifrem, "The Power of Graph Databases," in *Graph Databases*, O'Reilly, 2015, p. 9.

PROJECT NO. 57

Fast Video Compression

JAMIE VELLA

SUPERVISED BY
PROF. ING. CARL JAMES DEBONO

Video traffic has become one of the biggest loads on communication networks and data storage worldwide due to ubiquity of video content along with the increased desire for higher video quality, such as increased video resolution [1]. Uncompressed video signals generate a significant amount of data which would not be feasible for transmission or storage. This uncompressed data generated by video signals contains a lot of redundant information that can be removed to compact the amount of data required to store or transmit.

The latest video compression standard, High Efficiency Video Coding (HEVC), was developed with the interest of creating a new standard to improve compression as much as possible but to also enable deployment of services which would have not been practical with previous standards, such as the introduction of Ultra-HD (UHD) and High Dynamic Range (HDR) video content [1]. HEVC is designed on the same structure of the previous standard called H.264/AVC (Advanced Video Coding) but has incorporated multiple improvements, such as decreased bit rate and better compression efficiency, but at the cost of increased complexity, due to the increased number of predictions and searches that need to be performed [2].

The reference codec, called the HEVC test model (HM), is provided as a common reference implementation of the HEVC standard but is only made available to demonstrate the functionality of the codec and is not particularly optimised for speed or efficiency [2]. To speed up the encoding process of the reference codec, functions consuming a large part of the encoding time are modified to utilise the entire CPU through multi-threading provided by the OpenMP API. The functions consuming the most of the encoding time are identified by profiling the unmodified HM encoder. Profiling results give various performance metrics that help to identify where the most work is being done by the encoder. Parallelisation is implemented into the HM encoder based on the results obtained from the profiling stage.

The unmodified and modified encoders are compared with each other by encoding various video sequences using the same encoder configuration throughout. The implemented parallelisation resulted in a reduction of encoding time that ranges between ~10.5% to ~20.9%, with a negligible decrease in video quality of up to ~0.7% and a small increase in bit rate of up to ~4.4%.

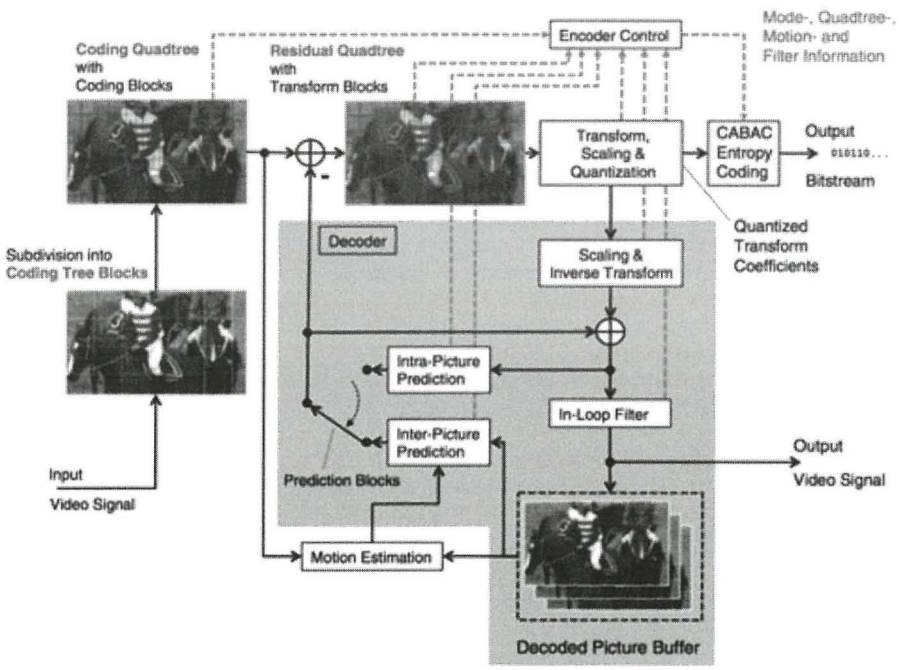


Figure 1: HEVC Encoder Block Diagram [3]

REFERENCES

[1] G. J. Sullivan. Introduction. In V. Sze, M. Budagavi, and G. J. Sullivan, editors, High Efficiency Video Coding (HEVC), chapter 1, pages 1–3. Springer, 2014.

[2] F. Bossen, B. Bross, K. Suhring, and D. Flynn. HEVC Complexity and Implementation Analysis. IEEE Transactions on Circuits and Systems for Video Technology, 22(12):1685–1696, Dec 2012.

[3] H. Schwarz, T. Schierl, and D. Marpe. Block Structures and Parallelism Features in HEVC. In V. Sze, M. Budagavi, and G. J. Sullivan, editors, High Efficiency Video Coding (HEVC), chapter 3. Springer, 2014.

PROJECT NO. 37

Profiles for Predicting Financial Distress using Company Final Accounts

VANESSA VELLA

SUPERVISED BY
MR JOSEPH BONELLO

Financial distress prediction is of interest to many stakeholders who strive to understand the financial position of a business often with limited knowledge about how it operates. This study aims to improve the current methods of predicting corporate distress through the development of an integrated decision support system based on the application of Machine Learning techniques, namely Decision Trees, Naïve Bayes and Artificial Neural Networks (Shalev-Shwartz & Ben-David, 2014).

Signs of business failure are in most cases evident long before official bankruptcy occurs. Through the use of 96 indicators, including financial ratios, industry-related variables and fraud red-flags, this study analyses a company's financial records in order to predict the sustainability of a given firm. Key profitability ratios were also compared to market averages. One important result which emerged was that the industry a company forms part of highly affects the financial stability of the firm. Therefore, whilst some markets may be performing well and boosting company growth, other industries may be the cause of company status deterioration. Two pattern analysis techniques were used: comparison to the previous year, and comparison to the first year of the available financial records. This analysis, combined with the proposed model for checking both the nature and magnitude of change across the comparisons yielded an F_1 Score (i.e. a score which conveys the balance between the precision and the recall) of up to 88.7%.

The research findings over four real-life datasets confirmed the strength and ability of the proposed model to predict imminent business failure. The proposed system compares three machine learning techniques. All three techniques are tested with the same datasets. Each technique uses a history span from three to five years.

The study finds that a model based on previous-year analysis performs better than a model based on base-year analysis. It was also found that in most cases, the inclusion of different aspects of a company's upkeep (profitability, solvency, leverage, management efficiency, industry specifics) lead to more accurate results (Refer to Figure 1).

Results have also shown that the Decision Tree outperforms the other two techniques, especially in previous-year comparison. One issue to note is that the nature of the model created in this study, that is dealing with the differences between calculated ratios, rather than the raw ratios themselves, has provided some reliable results with quite an adequate accuracy rate. With this implementation, the models were able to predict records with up to 94.7% accuracy. Research has shown that accuracy rates reached in the past were at most 94.2% (Mehrazin, Taghipour, Froutan, Ghabdian, & Soleimani, 2013).

Finally, the study can be extended by predicting the degree to which a company is failing. This would be quite advantageous as it would indicate to what extent firms have to go, to improve their financial status.

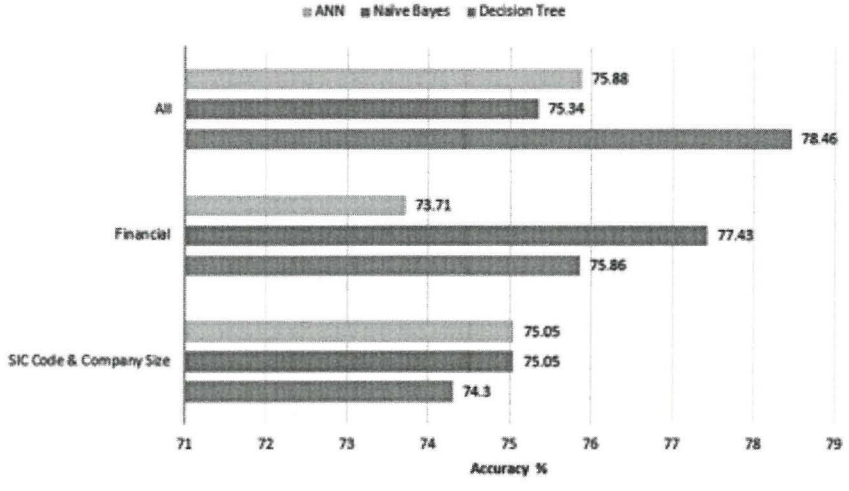


Figure 1: Previous Year Analysis on the American Dataset

REFERENCES

[1] Mehrazin, A., Taghipour, M., Froutan, O., Ghabdian, B., & Soleimani, H. (2013). Radial Basis Function in Artificial Neural Network for Prediction of Bankruptcy. International Business Research, 6(8), 121–129.

[2] Shalev-Shwartz, S., & Ben-David, S. (2014). Understanding Machine Learning.

PROJECT NO. 38

Generating Datasets through Data Source Analysis using ADaGe

MATTHEW XUEREB

SUPERVISED BY
MR JOSEPH BONELLO AND DR CONRAD ATTARD

The generation of a synthetic dataset involves the creation of fictitious data that follows some set of rules. The aim of this project is the generation of synthetic data that simulates the properties of an existing dataset by retaining the characteristics of its real-world counterpart. Using synthetic datasets as opposed to real-world data has various advantages, such as the ability to safeguard sensitive information by replacing it with fictitious subject matter (Wu et al., 2016). This prevents personal and identifiable information from being shared, thus retaining privacy for the individuals involved and enforcing data protection. It can also generate conditions which cannot be found in real-world datasets by modifying certain field properties. Changing the distributions of values within a dataset, for example, can be useful to improve the robustness of new algorithms (Ayala-Rivera, McDonagh, Cerqueus, & Murphy, 2013). This has the potential to reveal more software bugs within a product which is a highly desirable outcome of testing.

The aim of this project was to semi-automatically detect characteristics and properties between the fields of a MySQL datasource which were then replicated by a synthetic dataset. The program inferred information on the data properties from the source database and predicted their use. Once these results were verified and, where necessary, modified by the user, the tool identified the relationships between the variables, along with other useful information to generate semantically similar data. By automating Exploratory Data Analysis (EDA) techniques, relevant information on the relationship between different fields can be acquired as shown in Figure 1. In contrast to existing data generation tools which generally generate data on a column-by-column basis, this algorithm replicates data on a row-by-row basis maintaining the meaning between the different properties. This framework extends ADaGe, a synthetic data generator developed at the University of Malta by (Camilleri, Bonello, & Attard, 2016).

Statistical tests on the generated data, such as percentage frequency distribution statistics, showed that ADaGe retained relationships between categorical fields, whereas the binning techniques were effective in preserving the distribution in quantitative fields.

The correlation between multiple date fields in a table was maintained by preserving the range of the dates generated between date fields. Each date field also considered the days of the week, to cater for cases that exclude certain weekdays (such as omitting weekends in the simulation of the sales of a small grocery store).

Regular expressions were used to define values for text fields which are known to follow a pre-specified pattern. In trying to recreate numerical distribution patterns from normally distributed data, results showed mean and variance remained comparable when generating large volumes of data.

Future work includes enabling ADaGe to connect to other data sources such as Web Services. With regards to categorical fields, the percentage distribution observed for each of its possible values could be shown in parameterised fields such that the user is able to choose a different probability distribution for the data. For fields which follow a specific pattern, an automatic regular expression generator could be used to suggest patterns to the user.

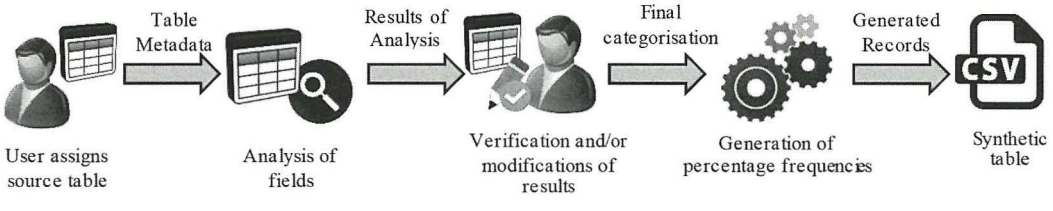


Figure 1: System Process

REFERENCES

[1] Ayala-Rivera, V., McDonagh, P., Cerqueus, T., & Murphy, L. (2013). Synthetic Data Generation using Benerator Tool. Retrieved from https://arxiv.org/pdf/1311.3312.pdf

[2] Camilleri, S., Bonello, J., & Attard, C. (2016). Automated Dataset Generator, (May), 60.

[3] Wu, H., Ning, Y., Chakraborty, P., Vreeken, J., Tatti, N., & Ramakrishnan, N. (2016). Generating Realistic Synthetic Population Datasets. Retrieved from http://arxiv.org/abs/1602.06844

PROJECT NO. 22

Enhancing Handling Precision of Virtual Objects in VR Experiences.

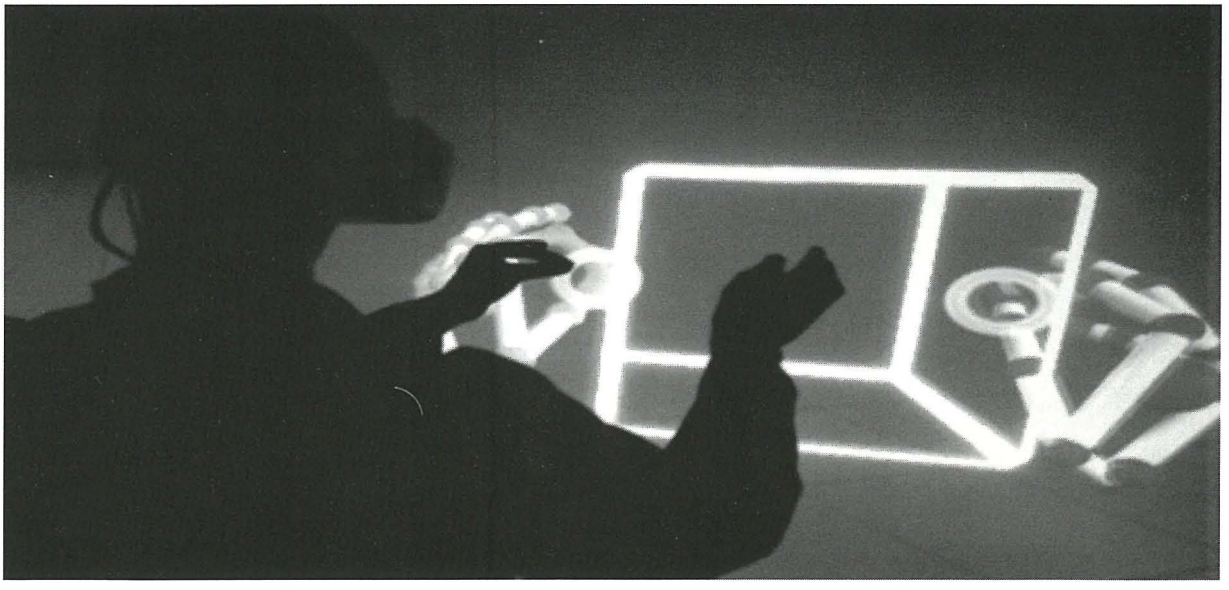
GARY ZAMMIT

SUPERVISED BY
DR SANDRO SPINA AND DR CHRISTOPHER PORTER

Interaction with virtual objects in a virtual reality environment is still not as realistic as one would expect. In virtual reality, there are two main modes to interact with virtual objects: *indirect interaction* and *direct interaction* [1, 2]. Indirect interaction revolves around the use of peripherals to interact with objects, whereby the user is required to make use of hand-held joysticks/controllers to manipulate objects. On the other hand, direct manipulation of objects in immersive virtual environment allows users to interact with objects without any physical device, but through a representation of the user's hands and fingers in a virtual space. The latter mode is more challenging because users must interact with objects without receiving any kind of haptic feedback arising from contact with the object itself (e.g. resistance, weight, texture, material density). Nonetheless, direct manipulation affords a more lifelike and natural method of interaction when compared to indirect peripheral-based interaction.

Launched in summer 2013 [2], the Leap Motion Controller, is a small optoelectronic sensor device, which enables computers to constantly track the user's bare hands through computer vision algorithms. This in turn allows for programmatic rendering of hand-data in a three-dimensional space with sub-millimeter accuracy [3]. However, challenges of precision exist when sensors are not able to read data correctly (e.g. because of finger occlusion), often making it hard to interact with objects in a realistic fashion (e.g. picking small object using the opposable thumb and index finger).

The main aim of this dissertation was to explore and attempt to improve the user experience when directly interacting with virtual objects of different shapes and sizes as well as actions and gestures. A novel interaction technique - *state-driven interaction* (SDI) - has been developed for this research project, building upon the state of the art in terms of hand-detection accuracy while predicting hand positions and user's intentions where necessary. This enabled the sensor to be used in scenarios where high hand movement precision (e.g. handling delicate or small objects) is required.



REFERENCES

[1] J. Lin, "Towards Naturally Grabbing and Moving Objects in VR.", San Diego's Qualcomm Institute, 2015.

[2] Oropesa, I., De Jong, T. L., Sánchez-González, P., Dankelman, J., & Gómez, E. J. (2016). Feasibility of tracking laparoscopic instruments in a box trainer using a Leap Motion Controller. Measurement: Journal of the International Measurement Confederation.

[3] Weichert, F., Bachmann, D., Rudak, B., & Fisseler, D. (2013). Analysis of the accuracy and robustness of the Leap Motion Controller. Sensors (Switzerland), 13(5), 6380–6393.

PROJECT NO. 6

Indoor Location System

JOHN ZAMMIT

SUPERVISED BY
PROF. SAVIOUR ZAMMIT

The objective of this project was to design and develop a mobile, indoor location technique. Existing outdoor triangulation-based positioning techniques like the Global Positioning System and GLONASS, are generally not suitable or accurate enough for indoor use, particularly in multilevel buildings. Developing a positioning technique for indoor spaces, opens possibilities for the deployment of various next generation intelligent systems that provide the user with a plethora of location based applications. These range from step-by-step indoor navigation, to product-level retail customer engagement, to proximity-based social networking.

The system was designed to utilize the signals from a single Wi-Fi access point to locate the client device, without support from any additional infrastructure. Furthermore, no prior radio-frequency finger-printing procedure is required, thus reducing the implementation complexity significantly.

The novel positioning approach developed in this project is based on a hybrid angle of arrival and received signal strength technique, which takes advantage of explicit beamforming (figure 1) and multiple-input and multiple-output (figure 2) techniques to achieve a fairly accurate location estimation.

To test the system, ray-tracing (figure 3) was used to characterise the propagation in an indoor, multipath-rich environment, and build appropriate indoor radio propagation models. Simulation was then used to show that accuracy improves with the number of transmitter and receiver antennas and was on par with more sophisticated techniques reported in the literature.

The structure of our final year project report was to first review the literature on indoor location services. The methodologies used to implement the system were then presented, followed by a preliminary analysis of the necessary calculations, a description of the simulation environment and an appraisal of the accuracy achieved by the suggested algorithm. Conclusions were then drawn and further work suggested.

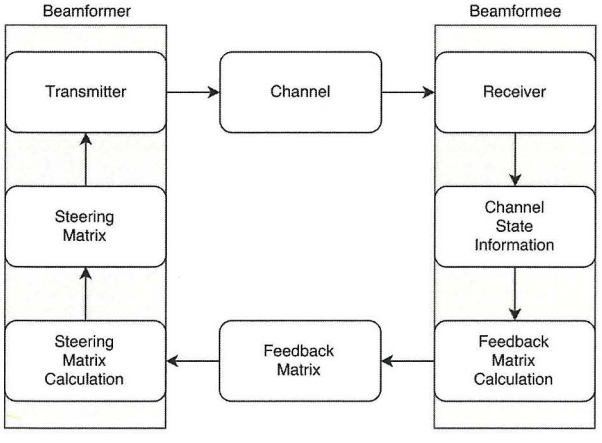


Figure 1: Explicit beam forming procedure

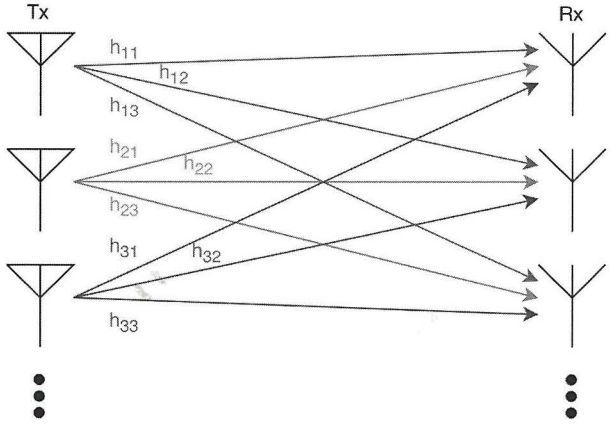


Figure 2: MIMO channel model

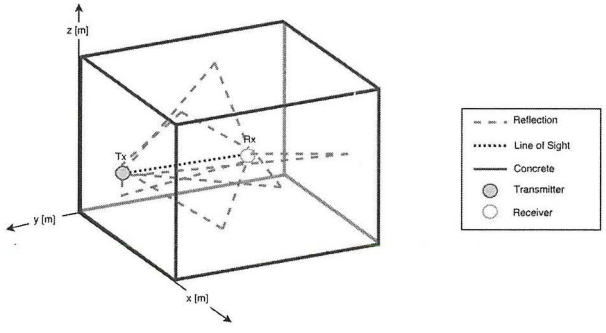


Figure 3: Three dimensional view of the measurement environment with ray traces

A Gamified Beacon Based Conference Management System

KENNETH ZERAFA

SUPERVISED BY
PROF. ALEXIEI DINGLI

Conferences are events with the aim of bringing together people with a shared interest. The ever-increasing dependence on smart mobile devices is leading to the development of new technologies which can improve attendees' experience at conferences. Mobile applications are nowadays expected to derive context and provide useful information without being prompted to. The concept of intuitively and contextually providing users with the right information at the appropriate time is crucial to Artificial Intelligence.

This project aims to improve attendees' experience at conferences by using Indoor Positioning technologies to provide contextual information, while offering a gamified experience to motivate the user to explore the conference. The developed mobile app is designed to make use of commonly available technology such as Bluetooth to satisfy the aims of the project.

For Indoor Positioning purposes, Bluetooth Low Energy (BLE) beacons were used. Beacons continuously transmit a signal which is picked up by mobile devices, which can in turn estimate the distance from a beacon by signal strength. Beacons are deployed in strategic points in the conference space and by estimating the distance from every beacon, the mobile device can approximate its location within the conference. The approximate location is used to inform the user of the closest attractions.

Gamification is defined as the use of game design elements in nongaming contexts [1]. One of the most prevalent problems of conferences is to keep attendees interested in the various attractions at specific days of the conference. For this reason, gamification techniques such as points, badges and milestones which can be shared to social media were implemented into the system. Gamification aims to increase user engagement by providing a more challenging, yet rewarding experience.

The system was used in two separate conferences, with a total of 70 invited participants for evaluation. Due to circumstances such as Bluetooth interference, one of the identified weaknesses of the system is that the response time to identify the closest beacon is not optimal. Notwithstanding the limitations of the artefact, the approach helped to demonstrate that the use of nowadays ubiquitous mobile technology for conference management can make a significant difference, especially when the application is tailored for the needs of every conference.

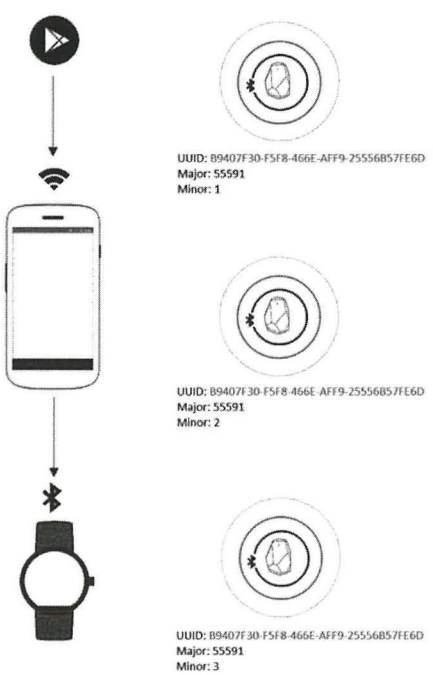


Figure 1: System Architecture

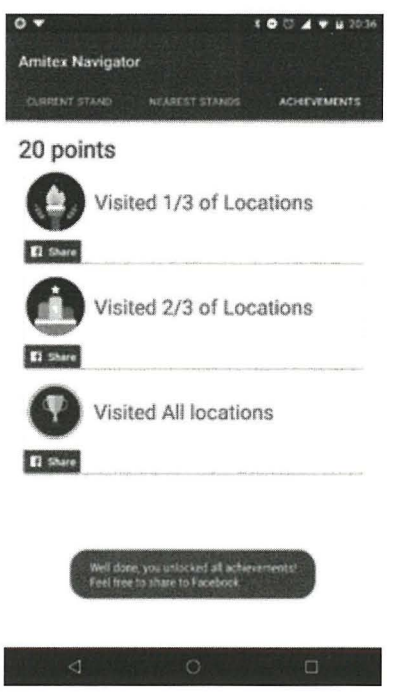


Figure 2: Gamification Aspect

REFERENCES

[1] S. Deterding, R. Khaled, L. Nacke, and D. Dixon, "Gamification: toward a definition," Chi 2011, pp. 12–15, 2011.


SUM
MER
Enrolment
Now Open
Achieve a worldwide recognised
ICT cerification. Back up your
University studies.
Raise your game.
icemalta.com
AUTODESK
Authorized Training Center
Authorized Certification Center
Adobe
digital
marketing
institute
Linux
Professional
Institute
Microsoft Partner
Silver Learning
cisco
Networking
Academy

Parallels®

Worldwide Academic Program

Parallels Worldwide Academic Program

Malta Student Partnership Experience

- Project work
- Practical R&D tasks
- Possibility to work with leading specialists of the best IT company in Malta, leading professionals of the world IT-Industry
- Real life case studies and workshops
- Professional mentors from acting team of developers

Advantages and benefits

- Real life projects
- Variety of themes and projects
- Research experience exchanges
- Team work experience
- Extra money bonus for achievements in research and study
- Job offer for the best students

Achievements of Parallels Worldwide Academic Program

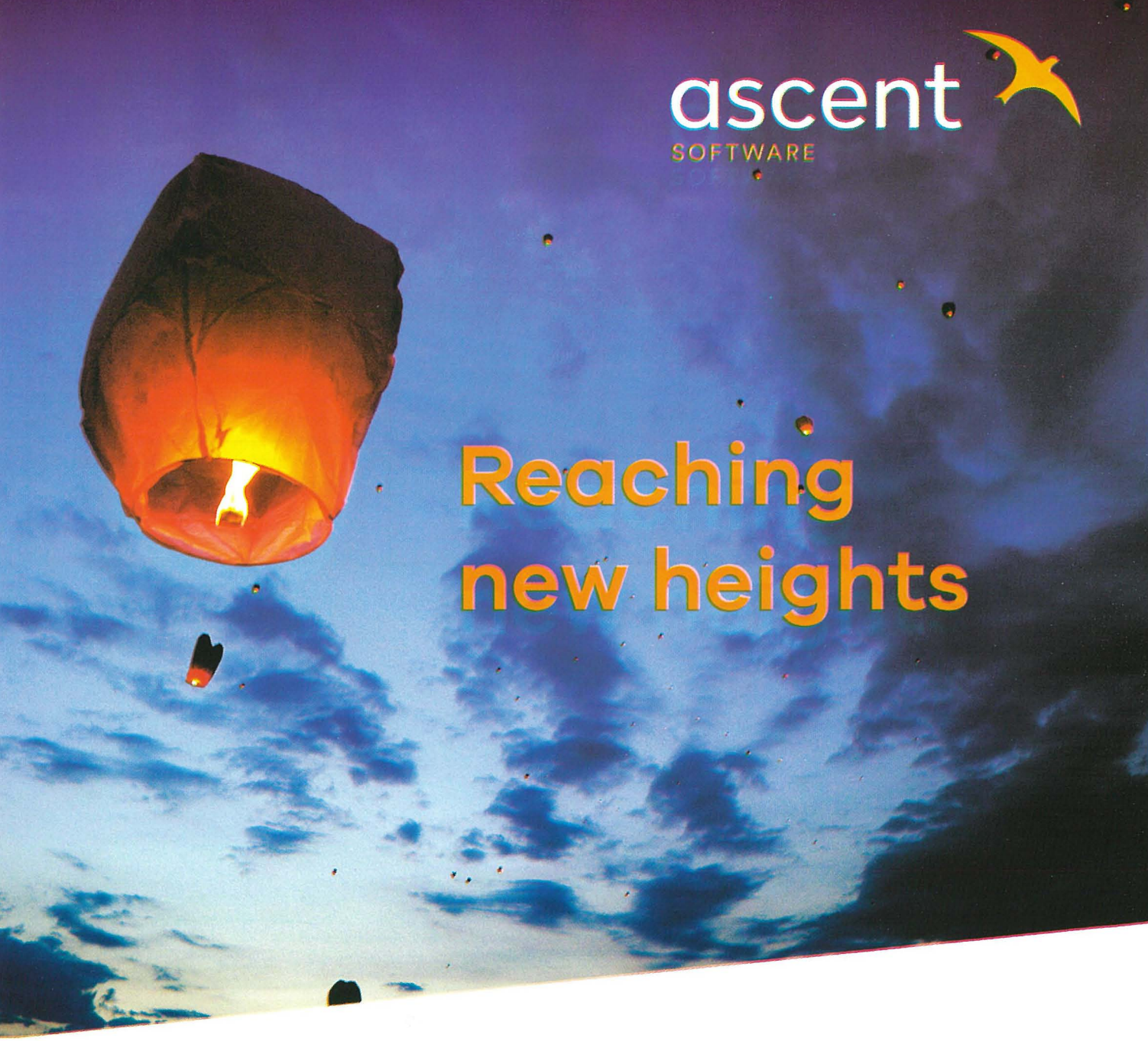
- More than 600 graduates
- 80% of graduates are now employees of the best IT-Companies of the world, other 20 % - academicians of leading universities in different countries
- Connections with the best universities in Europe

Our Academic Program is a big chance for prominent students to become part of global IT-Company. This time hundreds of our employees in different locations develop state-of-the-art modern and functional software for personal and business use all over the world. Will be glad to meet in the Company bold, talented, hard thinking, hardworking and ambitious young professionals.

Jack Zubarev, President of Parallels

About Parallels

Parallels Inc., a global leader in cross-platform solutions, makes it simple for customers to use and access the applications and files they need on any device or operating system. We help businesses and individuals securely and productively use their favorite devices and preferred technology, whether it's Windows®, Mac®, iOS, Android™, Chromebook, Windows Phone, Linux, Raspberry Pi or the Cloud. Parallels solutions enable seamless delivery of virtual desktops and applications to any device, running Windows on a Mac, Mac management with Microsoft SCCM, and remote access to PCs and Mac computers from any device. Parallels has offices in North America, Europe, Australia and Asia.



We are looking for talent, skills and passion

We offer summer internships and full-time positions to current students and fresh graduates.
Do you think you have what it takes to become an Ascentian?
Apply at careers@ascent.software

www.ascent.software
E. careers@ascent.software T. +356 2596 3000
92/3, Alpha Centre, Tarxien Road, Luqa LQA 1815
U.K. | U.S. | **Malta** | Middle East | Channel Islands

PROJECT NO: 37

Profiles for Predicting Financial Distress using Company Final Accounts

VANESSA VELLA

SUPERVISED BY
MR JOSEPH BONELLO

Financial distress prediction is of interest to many stakeholders who strive to understand the financial position of a business often with limited knowledge about how it operates. This study aims to improve the current methods of predicting corporate distress through the development of an integrated decision support system based on the application of Machine Learning techniques, namely Decision Trees, Naïve Bayes and Artificial Neural Networks (Shalev-Shwartz & Ben-David, 2014).

Signs of business failure are in most cases evident long before official bankruptcy occurs. Through the use of 96 indicators, including financial ratios, industry-related variables and fraud red-flags, this study analyses a company's financial records in order to predict the sustainability of a given firm. Key profitability ratios were also compared to market averages. One important result which emerged was that the industry a company forms part of highly affects the financial stability of the firm. Therefore, whilst some markets may be performing well and boosting company growth, other industries may be the cause of company status deterioration. Two pattern analysis techniques were used: comparison to the previous year, and comparison to the first year of the available financial records. This analysis, combined with the proposed model for checking both the nature and magnitude of change across the comparisons yielded an F_1 Score (i.e. a score which conveys the balance between the precision and the recall) of up to 88.7%.

The research findings over four real-life datasets confirmed the strength and ability of the proposed model to predict imminent business failure. The proposed system compares three machine learning techniques. All three techniques are tested with the same datasets. Each technique uses a history span from three to five years.

The study finds that a model based on previous-year analysis performs better than a model based on base-year analysis. It was also found that in most cases, the inclusion of different aspects of a company's upkeep (profitability, solvency, leverage, management efficiency, industry specifics) lead to more accurate results (Refer to Figure 1).

Results have also shown that the Decision Tree outperforms the other two techniques, especially in previous-year comparison. One issue to note is that the nature of the model created in this study, that is dealing with the differences between calculated ratios, rather than the raw ratios themselves, has provided some reliable results with quite an adequate accuracy rate. With this implementation, the models were able to predict records with up to 94.7% accuracy. Research has shown that accuracy rates reached in the past were at most 94.2% (Mehrazin, Taghipour, Froutan, Ghabdian, & Soleimani, 2013).

Finally, the study can be extended by predicting the degree to which a company is failing. This would be quite advantageous as it would indicate to what extent firms have to go, to improve their financial status.

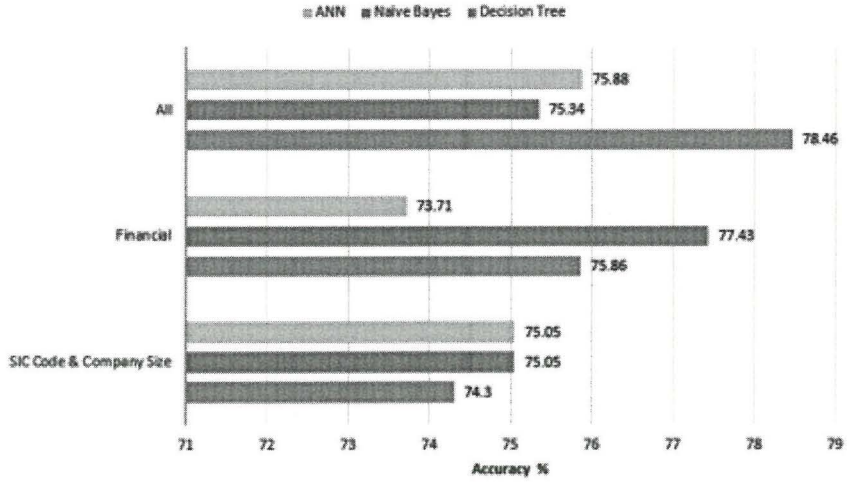


Figure 1: Previous Year Analysis on the American Dataset

REFERENCES

[1] Mehrazin, A., Taghipour, M., Froutan, O., Ghabdian, B., & Soleimani, H. (2013). Radial Basis Function in Artificial Neural Network for Prediction of Bankruptcy. International Business Research, 6(8), 121–129.

[2] Shalev-Shwartz, S., & Ben-David, S. (2014). Understanding Machine Learning.

PROJECT NO. 38

Generating Datasets through Data Source Analysis using ADaGe

MATTHEW XUEREB

SUPERVISED BY
MR JOSEPH BONELLO AND DR CONRAD ATTARD

The generation of a synthetic dataset involves the creation of fictitious data that follows some set of rules. The aim of this project is the generation of synthetic data that simulates the properties of an existing dataset by retaining the characteristics of its real-world counterpart. Using synthetic datasets as opposed to real-world data has various advantages, such as the ability to safeguard sensitive information by replacing it with fictitious subject matter (Wu et al., 2016). This prevents personal and identifiable information from being shared, thus retaining privacy for the individuals involved and enforcing data protection. It can also generate conditions which cannot be found in real-world datasets by modifying certain field properties. Changing the distributions of values within a dataset, for example, can be useful to improve the robustness of new algorithms (Ayala-Rivera, McDonagh, Cerqueus, & Murphy, 2013). This has the potential to reveal more software bugs within a product which is a highly desirable outcome of testing.

The aim of this project was to semi-automatically detect characteristics and properties between the fields of a MySQL datasource which were then replicated by a synthetic dataset. The program inferred information on the data properties from the source database and predicted their use. Once these results were verified and, where necessary, modified by the user, the tool identified the relationships between the variables, along with other useful information to generate semantically similar data. By automating Exploratory Data Analysis (EDA) techniques, relevant information on the relationship between different fields can be acquired as shown in Figure 1. In contrast to existing data generation tools which generally generate data on a column-by-column basis, this algorithm replicates data on a row-by-row basis maintaining the meaning between the different properties. This framework extends ADaGe, a synthetic data generator developed at the University of Malta by (Camilleri, Bonello, & Attard, 2016).

Statistical tests on the generated data, such as percentage frequency distribution statistics, showed that ADaGe retained relationships between categorical fields, whereas the binning techniques were effective in preserving the distribution in quantitative fields.

The correlation between multiple date fields in a table was maintained by preserving the range of the dates generated between date fields. Each date field also considered the days of the week, to cater for cases that exclude certain weekdays (such as omitting weekends in the simulation of the sales of a small grocery store).

Regular expressions were used to define values for text fields which are known to follow a pre-specified pattern. In trying to recreate numerical distribution patterns from normally distributed data, results showed mean and variance remained comparable when generating large volumes of data.

Future work includes enabling ADaGe to connect to other data sources such as Web Services. With regards to categorical fields, the percentage distribution observed for each of its possible values could be shown in parameterised fields such that the user is able to choose a different probability distribution for the data. For fields which follow a specific pattern, an automatic regular expression generator could be used to suggest patterns to the user.

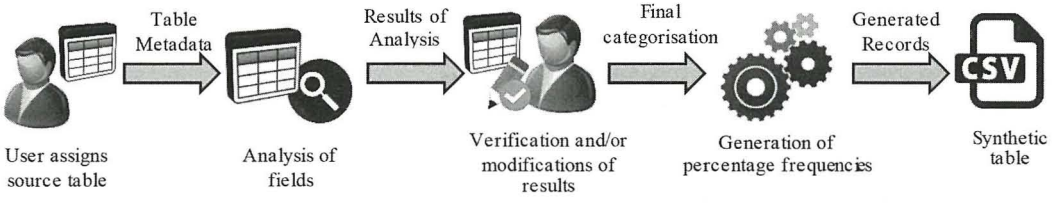


Figure 1: System Process

REFERENCES

[1] Ayala-Rivera, V., McDonagh, P., Cerqueus, T., & Murphy, L. (2013). Synthetic Data Generation using Benerator Tool. Retrieved from https://arxiv.org/pdf/1311.3312.pdf

[2] Camilleri, S., Bonello, J., & Attard, C. (2016). Automated Dataset Generator, (May), 60.

[3] Wu, H., Ning, Y., Chakraborty, P., Vreeken, J., Tatti, N., & Ramakrishnan, N. (2016). Generating Realistic Synthetic Population Datasets. Retrieved from http://arxiv.org/abs/1602.06844

PROJECT NO. 22

Enhancing Handling Precision of Virtual Objects in VR Experiences.

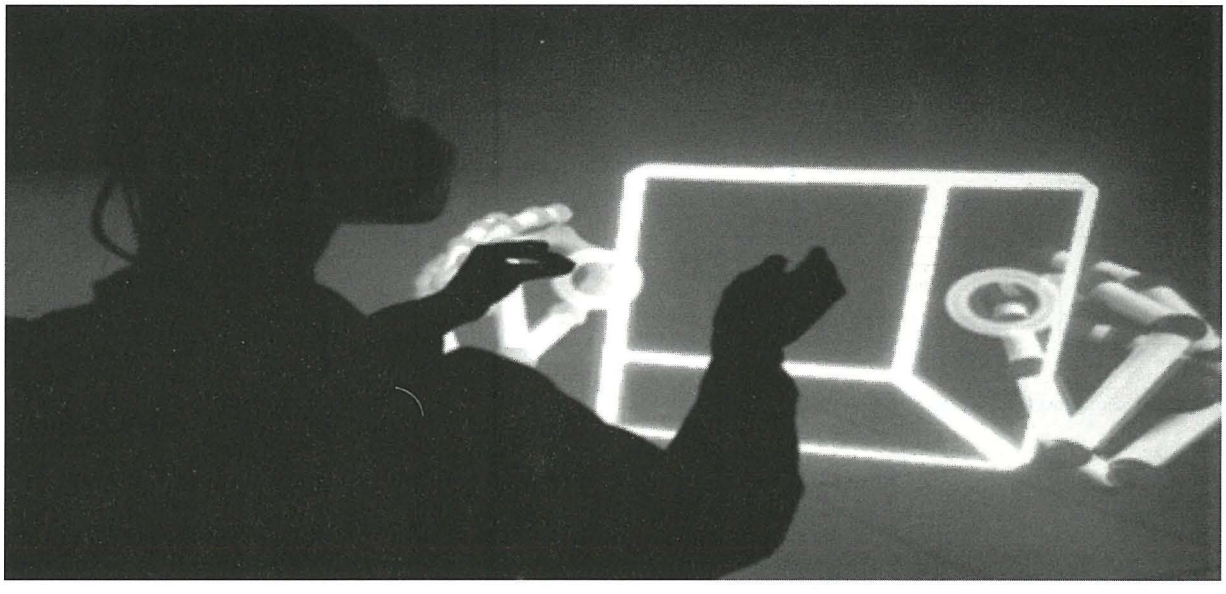
GARY ZAMMIT

SUPERVISED BY
DR SANDRO SPINA AND DR CHRISTOPHER PORTER

Interaction with virtual objects in a virtual reality environment is still not as realistic as one would expect. In virtual reality, there are two main modes to interact with virtual objects: *indirect interaction* and *direct interaction* [1, 2]. Indirect interaction revolves around the use of peripherals to interact with objects, whereby the user is required to make use of hand-held joysticks/controllers to manipulate objects. On the other hand, direct manipulation of objects in immersive virtual environment allows users to interact with objects without any physical device, but through a representation of the user's hands and fingers in a virtual space. The latter mode is more challenging because users must interact with objects without receiving any kind of haptic feedback arising from contact with the object itself (e.g. resistance, weight, texture, material density). Nonetheless, direct manipulation affords a more lifelike and natural method of interaction when compared to indirect peripheral-based interaction.

Launched in summer 2013 [2], the Leap Motion Controller, is a small optoelectronic sensor device, which enables computers to constantly track the user's bare hands through computer vision algorithms. This in turn allows for programmatic rendering of hand-data in a three-dimensional space with sub-millimeter accuracy [3]. However, challenges of precision exist when sensors are not able to read data correctly (e.g. because of finger occlusion), often making it hard to interact with objects in a realistic fashion (e.g. picking small object using the opposable thumb and index finger).

The main aim of this dissertation was to explore and attempt to improve the user experience when directly interacting with virtual objects of different shapes and sizes as well as actions and gestures. A novel interaction technique - *state-driven interaction* (SDI) - has been developed for this research project, building upon the state of the art in terms of hand-detection accuracy while predicting hand positions and user's intentions where necessary. This enabled the sensor to be used in scenarios where high hand movement precision (e.g. handling delicate or small objects) is required.



REFERENCES

[1] J. Lin, "Towards Naturally Grabbing and Moving Objects in VR.", San Diego's Qualcomm Institute, 2015.

[2] Oropesa, I., De Jong, T. L., Sánchez-González, P., Dankelman, J., & Gómez, E. J. (2016). Feasibility of tracking laparoscopic instruments in a box trainer using a Leap Motion Controller. Measurement: Journal of the International Measurement Confederation.

[3] Weichert, F., Bachmann, D., Rudak, B., & Fisseler, D. (2013). Analysis of the accuracy and robustness of the Leap Motion Controller. Sensors (Switzerland), 13(5), 6380–6393.

PROJECT NO. 6

Indoor Location System

JOHN ZAMMIT

SUPERVISED BY
PROF. SAVIOUR ZAMMIT

The objective of this project was to design and develop a mobile, indoor location technique. Existing outdoor triangulation-based positioning techniques like the Global Positioning System and GLONASS, are generally not suitable or accurate enough for indoor use, particularly in multilevel buildings. Developing a positioning technique for indoor spaces, opens possibilities for the deployment of various next generation intelligent systems that provide the user with a plethora of location based applications. These range from step-by-step indoor navigation, to product-level retail customer engagement, to proximity-based social networking.

The system was designed to utilize the signals from a single Wi-Fi access point to locate the client device, without support from any additional infrastructure. Furthermore, no prior radio-frequency finger-printing procedure is required, thus reducing the implementation complexity significantly.

The novel positioning approach developed in this project is based on a hybrid angle of arrival and received signal strength technique, which takes advantage of explicit beamforming (figure 1) and multiple-input and multiple-output (figure 2) techniques to achieve a fairly accurate location estimation.

To test the system, ray-tracing (figure 3) was used to characterise the propagation in an indoor, multipath-rich environment, and build appropriate indoor radio propagation models. Simulation was then used to show that accuracy improves with the number of transmitter and receiver antennas and was on par with more sophisticated techniques reported in the literature.

The structure of our final year project report was to first review the literature on indoor location services. The methodologies used to implement the system were then presented, followed by a preliminary analysis of the necessary calculations, a description of the simulation environment and an appraisal of the accuracy achieved by the suggested algorithm. Conclusions were then drawn and further work suggested.

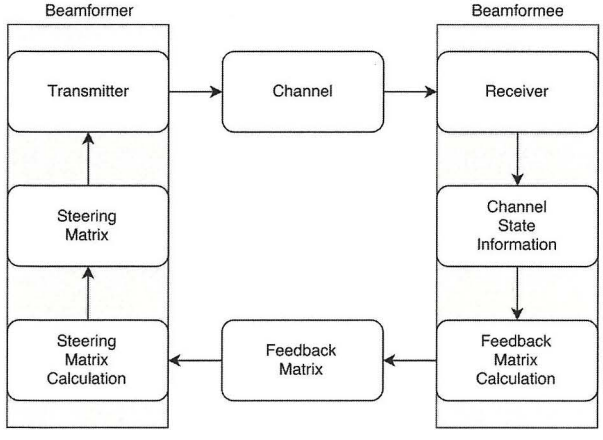


Figure 1: Explicit beam forming procedure

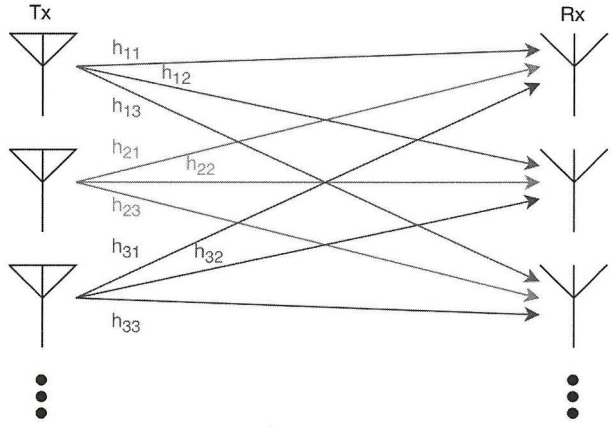


Figure 2: MIMO channel model

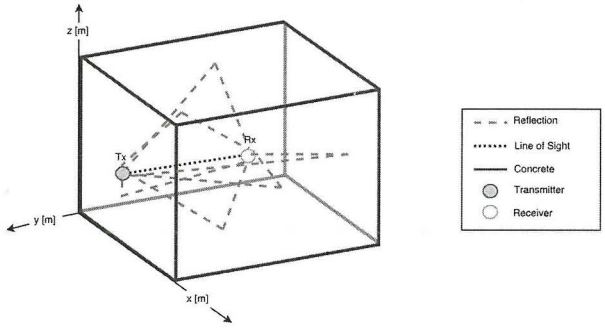


Figure 3: Three dimensional view of the measurement environment with ray traces

PROJECT NO. 81

A Gamified Beacon Based Conference Management System

KENNETH ZERAFA

SUPERVISED BY
PROF. ALEXIEI DINGLI

Conferences are events with the aim of bringing together people with a shared interest. The ever-increasing dependence on smart mobile devices is leading to the development of new technologies which can improve attendees' experience at conferences. Mobile applications are nowadays expected to derive context and provide useful information without being prompted to. The concept of intuitively and contextually providing users with the right information at the appropriate time is crucial to Artificial Intelligence.

This project aims to improve attendees' experience at conferences by using Indoor Positioning technologies to provide contextual information, while offering a gamified experience to motivate the user to explore the conference. The developed mobile app is designed to make use of commonly available technology such as Bluetooth to satisfy the aims of the project.

For Indoor Positioning purposes, Bluetooth Low Energy (BLE) beacons were used. Beacons continuously transmit a signal which is picked up by mobile devices, which can in turn estimate the distance from a beacon by signal strength. Beacons are deployed in strategic points in the conference space and by estimating the distance from every beacon, the mobile device can approximate its location within the conference. The approximate location is used to inform the user of the closest attractions.

Gamification is defined as the use of game design elements in nongaming contexts [1]. One of the most prevalent problems of conferences is to keep attendees interested in the various attractions at specific days of the conference. For this reason, gamification techniques such as points, badges and milestones which can be shared to social media were implemented into the system. Gamification aims to increase user engagement by providing a more challenging, yet rewarding experience.

The system was used in two separate conferences, with a total of 70 invited participants for evaluation. Due to circumstances such as Bluetooth interference, one of the identified weaknesses of the system is that the response time to identify the closest beacon is not optimal. Notwithstanding the limitations of the artefact, the approach helped to demonstrate that the use of nowadays ubiquitous mobile technology for conference management can make a significant difference, especially when the application is tailored for the needs of every conference.

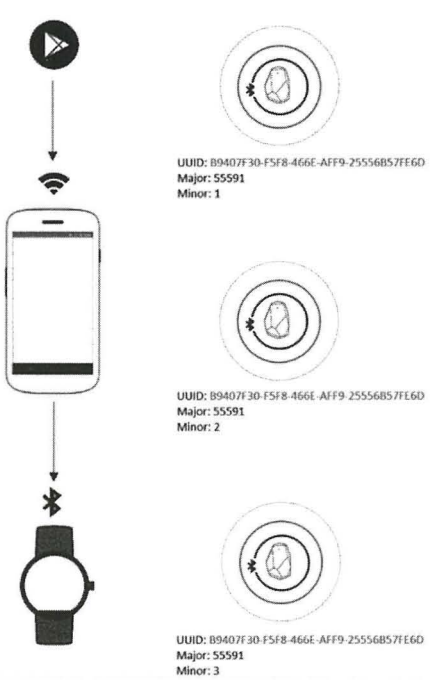


Figure 1: System Architecture

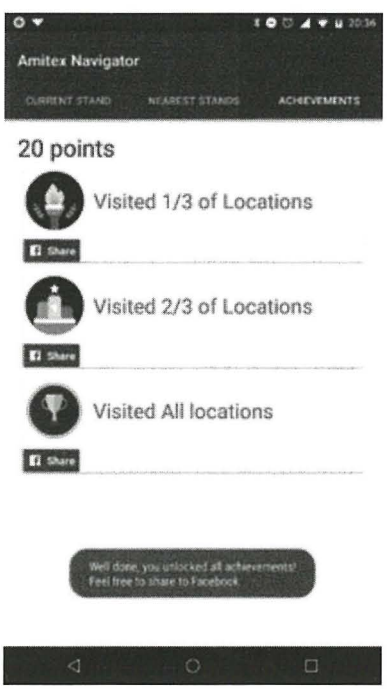


Figure 2: Gamification Aspect

REFERENCES

[1] S. Deterding, R. Khaled, L. Nacke, and D. Dixon, "Gamification: toward a definition," Chi 2011, pp. 12–15, 2011.


SUM
MER
Enrolment
Now Open
Achieve a worldwide recognised ICT cerification. Back up your University studies.
Raise your game.
icemalta.com
AUTODESK
Authorized Training Center
Authorized Certification Center
Adobe
digital marketing institute
Linux Professional Institute
Microsoft Partner
Silver Learning
cisco
Networking Academy

Parallels®

Worldwide Academic Program

Parallels Worldwide Academic Program

Malta Student Partnership Experience

- Project work
- Practical R&D tasks
- Possibility to work with leading specialists of the best IT company in Malta, leading professionals of the world IT-Industry
- Real life case studies and workshops
- Professional mentors from acting team of developers

Advantages and benefits

- Real life projects
- Variety of themes and projects
- Research experience exchanges
- Team work experience
- Extra money bonus for achievements in research and study
- Job offer for the best students

Achievements of Parallels Worldwide Academic Program

- More than 600 graduates
- 80% of graduates are now employees of the best IT-Companies of the world, other 20 % - academicians of leading universities in different countries
- Connections with the best universities in Europe

Our Academic Program is a big chance for prominent students to become part of global IT-Company. This time hundreds of our employees in different locations develop state-of-the-art modern and functional software for personal and business use all over the world. Will be glad to meet in the Company bold, talented, hard thinking, hardworking and ambitious young professionals.

Jack Zubarev, President of Parallels

About Parallels

Parallels Inc., a global leader in cross-platform solutions, makes it simple for customers to use and access the applications and files they need on any device or operating system. We help businesses and individuals securely and productively use their favorite devices and preferred technology, whether it's Windows®, Mac®, iOS, Android™, Chromebook, Windows Phone, Linux, Raspberry Pi or the Cloud. Parallels solutions enable seamless delivery of virtual desktops and applications to any device, running Windows on a Mac, Mac management with Microsoft SCCM, and remote access to PCs and Mac computers from any device. Parallels has offices in North America, Europe, Australia and Asia.

We offer summer internships and full-time positions to current students and fresh graduates.
Do you think you have what it takes to become an Ascentian?
Apply at careers@ascent.software

E. careers@ascent.software T. +356 2596 3000
92/3, Alpha Centre, Tarxien Road, Luqa LQA 1815
U.K. | U.S. | **Malta** | Middle East | Channel Islands

CCBill

why be **bored** @ work?

when you can have **fun** @ **CCBILL**

WAKE UP & BE AWESOME

interested? we're recruiting!

recruitment@ccbill.com

Find us on Facebook