

**INTERNATIONAL DOCUMENT
IMAGE PROCESSING
SUMMER SCHOOL 2015
21 - 26 JUNE 2015
KEFALONIA ISLAND - GREECE**

Informations

INDEX

ABOUT THE SUMMER SCHOOL.....	4
VENUE	5
PEOPLE	7
SCHOOL PROGRAM.....	8
SCIENTIFIC LECTURE	9
COMPANY PRESENTATIONS.....	21
POSTER ABSTRACTS	22
AWARDS	25
MAP OF POROS	26
USEFUL TELEPHONE NUMBERS	27

ABOUT THE SUMMER SCHOOL

After the previous successful Summer Schools IDIPS 2013 & IDIPS2014 (see IAPR reports <http://www.iapr.org/docs/newsletter-2013-03.pdf> and <http://www.iapr.org/docs/newsletter-2014-03.pdf>), the third Summer School (IDIPS2015) aims to provide among other, an objective overview and an in-depth analysis of the state-of-the-art of the Camera-Based Document Images, and their applications. Lectures will be delivered by world-renowned researchers of the field.

This year, the school aims to provide an attractive opportunity for 40 young researchers, post-graduate and PhD students. The participants will benefit from direct interaction and discussions with the renowned speakers and practice on real problems. The researchers will also have the opportunity to present the results of their scientific research, and interact with colleagues in a friendly and constructive atmosphere.

The Summer School is hosted on the beautiful island of Kefalonia, in the village of Poros, 21-26 June.

VENUE

The magnificent sights, amazing beaches, rich cultural heritage, great monuments, mountains, castles, remote monasteries and cheerful, hospitable people are the treasures making Kefalonia one of the most attractive destinations.

It is the biggest island of the Ionian Sea, at the outlet of Patraikos Bay, between Zante and Lefkada. It stretches over an area of 904 km² with a coastline length of 250 km.

The island has a splendid natural beauty and a diversified geological structure, with innumerable small bays and rich vegetation; the mountain Aenos (National Park, since 1962, supporting rich flora and fauna) looks impressive from far away.

During the Byzantine period Kefalonia formed part of the province of Achaia and the prefecture of East Illyria. It was conquered, by the Normans, the Franks, the Venetians and the Turks.

It joined the rest of Greece along with the other Ionian Islands, on May 21st, 1864. The inhabitants' resistance against the Italians and the Germans during World War II was considerable.

The island of Kefalonia has escaped mass tourism and recently was used to film Captain Corelli's Mandolin, set in war-time Kefalonia. The film, with Nikolas Cage, was a huge success and it put Kefalonia firmly on the 'must visit' list of destinations.

Kefalonia has one airport, Kefalonia Island International Airport, with a runway around 2.4 km (1.5 mi) in length, located about 10 km (6 mi) south of Argostoli. In summer, the airport handles a number of charter flights from all over Europe.

There are five harbors and ports in the prefecture: four main harbors on the island: Same or Sami, a major port with links to Patras and Ithaca. Poros, in the south, has ferry routes to Kyllini. Argostoli, in the west, is the largest port, for local boats and ferries to Zante and regularly to Lixouri. Fiscardo, in the north, has links to Lefkas and Ithaca.

The summer school takes place in Poros. Poros is effectively divided into three parts: Poros port, with its couple of tavernas and bars, connects the island with Kyllini on the Peloponnese area of mainland Greece via regular year-round ferry service. During the summer months, a ferry may connect Poros with Zakynthos and a couple of tourist caiques offer cruises to Ithaca and the Blue Caves of Zakynthos. The harbour is also home to the local fishing boats which supply the area with fresh fish.

Separated from Poros port by a small hillock, lies the shingle town beach, backed by a taverna-fringed square (plateia) and main services: bank, chemist (pharmacy), doctor's surgery, post office, police station and local shops. More tavernas and bars lie along the award-winning 'Blue Flag' beach, Aragia, separated from the center of Poros by the Vohinas river which, although used as a car park in summer, can be a raging torrent in winter.

The Vohinas, almost certainly, gave Poros its name. The word 'poros' in Greek means 'ford' or 'crossing' and where the Vohinas enters, the sea is the only natural crossing point, although two bridges now cross the river further upstream, where the Vohinas exits the impressive Poros gorge, an 80-metre deep rugged ravine which, until a road was cut through to Skala around 1996, was the only land route out of Poros. Local legend supports that the gorge was 'cut' by Heracles when he stood on, and flattened, this part of the mountain.

PEOPLE

General Chairs

[Athanasios Skodras](#), *University of Patras, Greece*

[Ergina Kavallieratou](#), *University of the Aegean, Greece*

Advisory Board

[Daniel P. Lopresti](#), *Lehigh University, USA*

[David Doermann](#), *University of Maryland, USA*

[Seiichi Uchida](#), *Kyushu University, Japan*

Sponsorship Chair

[Jean-Marc Ogier](#), *Université de La Rochelle, France*

Local Organizers

[Alexandros Papandreou](#), *National Center for Scientific Research "Demokritos", Greece*

[Giorgos Louloudis](#), *National Center for Scientific Research "Demokritos", Greece*

[Paraskevas Diamantatos](#), *University of the Aegean, Greece*

Invited Lecturers

[Dimosthenis Karatzas](#), *Universitat Autònoma de Barcelona, Spain*

[Efstathios Stamatatos](#), *University of the Aegean, Greece*

[Elisa H. Barney Smith](#), *Boise State University, USA*

[Enrique Vidal](#), *Universidad Politècnica de València, Spain*

[Ergina Kavallieratou](#), *University of the Aegean, Greece*

[Franck Lebourgeois](#), *INSA de Lyon, France*

[Jean-Marc Ogier](#), *Université de La Rochelle, France*

[Seiichi Uchida](#), *Kyushu University, Japan*

[Umapada Pal](#), *Indian Statistical Institute, India*

	Sunday 21/6	Monday 22/6	Tuesday 23/6	Wednesday 24/6	Thursday 25/6	Friday 26/6
9.30 -	E X C U R S I O N	<u>Preprocessing</u> Ergina Kavallieratou	<u>Historical Handwritten Documents</u> Enrique Vidal	<u>Business Documents</u> Frank Lebourgeois	<u>Indexing</u> Jean-Marc Oqier	<u>Robust Reading</u> Dimosthenis Karatzas
11.00 -		Coffee Break				
11.30		<u>Binarization</u> Elisa Barney Smith	<u>Segmentation</u> Jean-Marc Oqier	<u>Multi-Script Document Processing</u> Umapada Pal	<u>Performance Evaluation</u> Efstathios Stamatatos	<u>Camera-based OCR</u> Seiichi Uchida
11.30 -		Coffee Break				
13.00		Coffee Break				
-		Coffee Break				
13.30		<u>Advice For a Successful PhD Experience</u> Daniel Lopresti	<u>From Laboratory to Market</u> Adolfo Santoro	<u>The Historical Archives of Cephalonia</u> Dora Zafeiratou	<u>Case Study of a Commercial DIA System</u> Anagnostopoulos Tassos	Poster Session & Evaluation
15.00		Workshop	Workshop	Workshop	Workshop	Workshop Presentation
-		Coffee Break				
16.00		Workshop	Workshop	Workshop	Workshop	Workshop Presentation
20.00 - ...	Dinner at "Agrapidos"	Dinner at "Memas"	Dinner at "Fotis Family"	Dinner at "Pantelis"	Dinner at "To Steki"	Dinner at the Lecture Hall & Closing Ceremony

SCIENTIFIC LECTURE

Preprocessing

Ergina Kavallieratou

By preprocessing, we refer to all the procedures before the main task of a Document Image Processing system. It is a very wide area starting from image binarization, extending to noise removal, Geometric and any other transformations and could reach to segmentation, normalization, etc. Such preprocessing tasks are included in almost every Document Image Processing system. In this presentation, the evolution through time of several popular tasks is going to be presented. The combination of various tasks and their dependence on the system will be analyzed and evaluated. Our intention is to demonstrate how any given system could impose its preprocessing requirements, analyzing its involved problems and characteristics. Moreover, we will examine how the same preprocessing tasks can be shared between similar or related systems.

Ergina Kavallieratou was born in Kefalonia, Greece, in 1973. She received her Diploma in Electrical and Computer Engineering in 1996 from the Polytechnic School of the University of Patras and her PhD in Handwritten Optical Character Recognition and Document Image processing from the same department in 2000. She has worked as guest researcher in the Signals, Systems and Radiocomunications Laboratory of the Dept. of Telecommunications Engineering of the Polytechnic School of Madrid (1997-1998), in the Institute of Communication Acoustics of Ruhr-Universitaet Bochum, Germany (2000, 2001), Computer Science & Engineering, Lehigh University - USA (December 2009), CVC, Universidad de Barcelona (June 2011). During the years 2002-2004, she was an Assistant Professor of Audio Processing in Dept. of Audio and Musical Instruments Technology in the Technological Educational Institute of Ionian Islands, Greece. She teaches in Greek Open University, since 2001. Since September 2004, she is a member of the teaching staff of the department of Information and Communication System Engineering, University of the Aegean, as Assistant Professor since 2013. Her research interests include Optical Character Recognition, Document Image Analysis, Computer Vision and Pattern Recognition.

<http://samosweb.aegean.gr/cv/ekavallieratou.php>

Binarization

Elisa Barney Smith

Image binarization is the first step in the Document analysis pipeline, usually followed by skew detection and correction, line and word segmentation then recognition. Unless the subsequent algorithms process directly on gray scale images, the quality of the image binarization will affect the other algorithms. This lecture will discuss several image binarization algorithms, and then talk about how they can be evaluated and integrated with the rest of the document processing work flow.

Elisa Barney Smith is a professor in the Electrical & Computer Engineering department at Boise State University in Boise Idaho USA, where she has worked since 1999. She received a B.S. in Computer Science and the M.S. and Ph.D. degrees in Electrical, Computer and Systems Engineering all from Rensselaer Polytechnic Institute, Troy, NY, USA. Elisa's main research interests are image processing and machine learning. She applies these to document imaging as well as to biomedical image processing and image processing for materials science research and soil remediation evaluation. She is also PI on a new NSF funded project to design computer chips with an architecture inspired by the human brain.

Document analysis research has been primarily focused around developing models of the degradations produced during document image acquisition, and analyzing the defects that can be produced by the models.

Other work in this area includes image comparison for defect detection, ballot image processing, and non-linear image pre-processing to improve binarization and recognition. She has worked on multiple multi-institutional and cross-disciplinary projects, and across wide geographic expanses, including internationally. Her research has been funded by NSF, including a CAREER grant, NASA, NIH, and HP. She was a program chair for ICDAR 2013, and for DRR 2003-2005, and serves on many other conference program committees. She is currently an associate editor for Springer's International Journal on Document Analysis and Recognition.

<http://coen.boisestate.edu/EBarneySmith/>

Advice for a Successful PhD Experience

Daniel Lopresti

Performing original research is exhilarating. You have an opportunity to examine

basic questions facing our field that can be addressed through new ideas in document image analysis. You will become an international expert in a particular area. Others will seek out your advice on important topics. In the process, not only will you learn a great deal, but you will also have a chance to make the world a better place. With this opportunity comes significant responsibility, however. In this talk, I will discuss some of the expectations held by established researchers for graduate students just beginning their careers. My goal is to offer advice that I hope is helpful, and to outline some of the things you should keep in mind to increase your chances of having a successful and rewarding PhD experience.

Daniel Lopresti received his bachelor's degree from Dartmouth in 1982 and his Ph.D. in computer science from Princeton in 1987. After completing his doctorate, he joined the Department of Computer Science at Brown and taught courses ranging from VLSI design to computational aspects of molecular biology and conducted research in parallel computing and VLSI CAD. He went on to help found the Matsushita Information Technology Laboratory in Princeton, and later also served on the research staff at Bell Labs where his work turned to document analysis, handwriting recognition, and biometric security.

In 2003, Dr. Lopresti joined the Department of Computer Science and Engineering at Lehigh where he leads a research group examining fundamental algorithmic and systems-related questions in pattern recognition, bioinformatics, and security. Dr. Lopresti is director of the Lehigh Pattern Recognition Research (PatRec) Lab. On July 1, 2009, he became Chair of the Department of Computer Science and Engineering.

<http://www.cse.lehigh.edu/~lopresti/>

Historical Handwritten Documents

Enrique Vidal

Huge amounts of historical data are currently "buried" behind thousands of terabytes of handwritten text images. Automated methods are needed to make such important textual content accessible to scholars and users in general. OCR transcription technology is generally useless for this kind of documents because it is practically impossible to reliably segment the images of interest into individual characters or even into individual words. This is why holistic Handwritten Text Recognition (HTR) approaches have emerged in the last two decades.

This lecture will present state-of-the-art HTR technologies for the transcription of handwritten text images. Both fully automatic and semi-automatic, computer-assisted

HTR will be presented. Finally, HTR applications where hand-held devices can be used for (camera-based) image acquisition will be outlined, along with multimodal interactive techniques which can be used to take advantage of the multiple input modalities of these devices.

Empirical results with real, large image collections will be reported and live demonstrations with these collections will be presented. Moreover, the attendants will have the opportunity of a hands-on experience of using these advanced handwritten image transcription systems.

Enrique Vidal is a full professor of Computer Science in the Universitat Politècnica de València (Spain) and a co-founder and former co-leader of the PRHLT Research Center of this university. He has published more than two hundred research papers in the fields of Pattern Recognition and Multimodal Interaction and applications to Language, Speech, Handwriting and Image Processing, and has led many important projects in these fields. Dr. Vidal is a member of the IEEE and a fellow of the International Association for Pattern Recognition (IAPR).

<https://www.prhlt.upv.es/~evidal/>

Segmentation

Jean-Marc Ogier

This talk will provide an overview concerning the main approaches of the literature concerning segmentation techniques for document analysis.

These methods will be discussed according to their context of utilization.

The outline of the presentation will be the following

1. Introduction : challenge related to document segmentation ; particularities in the context of mobile capture
2. Top down segmentation : XY cut, Split & Merge, Run Length Smoothing algorithm, texture based approaches (probabilistic, geometric, frequential), Voronoi based approaches, autocorrelation based approaches
3. Bottom-up segmentation strategies : Connected components based approaches
4. Hybrid and alternatives approaches : gradient based approaches, multi-resolution approaches;
5. Text - Graphic segmentation using alternatives approaches : keypoints based approaches
6. Context based approaches

7. Conclusion - discussion

Jean-Marc Ogier received his PhD degree in computer science from the University of Rouen, France, in 1994.

During this period (1991-1994), he worked on graphic recognition for Matra Ms&I Company. From 1994 to 2000, he was an associate professor at the University of Rennes 1 during a first period (1994-1998) and at the University of Rouen from 1998 to 2001. Now full professor at the university of la Rochelle, Pr Ogier is the head of URL laboratory which gathers more than 120 members and works mainly of Document Analysis and Content Management.

Author of more than 200 publications / communications, he managed several French and European projects dealing with historic document analysis, either with public institutions, or with private companies. Pr Ogier is a Deputy Director of the GDR I3 of the French National Research Centre (CNRS). He is also Chair of the Technical Committee 10 (Graphic Recognition) of the International Association for Pattern Recognition (IAPR), and is the representative member of France at the governing board of the IAPR. At last he is also Vice rector of the university of La Rochelle.

<http://pageperso.univ-lr.fr/jmogier>

Business Documents

Franck Lebourgeois

The recognition of Business documents like forms, invoices, postal mail, is one of the most important application in Document Image Analysis (DIA) which concern many private companies and public organizations in the world. Business documents processing is also a challenging problem, rarely studied in the scientific literature. These documents are complex and have heterogeneous contents. Business documents contain degraded printings and handwritings which overlapped complex coloured backgrounds or pre-printings forms. Moreover, there is no unique model of business document, each private company or public organization, design their own template. Consequently, the layout and the logical structure vary from a company to an another. Numerous commercial systems exist, but they simplify the recognition process by using a manual model which designs the informative zones for each type of document. Most of existing software use a large database of predefined models which describes the template of each existing business document from the main companies. This solution is suited for large companies which process the same model of documents from other large companies. For small companies which have to process daily a large quantity of documents from different origins, the manual modelisation is impossible to achieve. The only solution consists to develop a system which recognizes the logical structure without any predefined

model. A data-driven recognition must replace the model-driven classical approach. We will present the main works from the state of the art in image pre-processing and segmentation, layout analysis and logical structure recognition applied to business documents.

Frank LeBourgeois is a senior researcher from LIRIS laboratory from INSA Lyon (France) since 1992. He is specialized in the domain of Document Images Analysis (DIA) and its works cover a wide range of applications in the domain of document imaging. With more than 25 years of experience in this field, he has published more than 70 international papers in international conferences and journals and has served in numerous program committees of international and national conferences. He is reviewer of international journals and international conferences of the domain. Since 2000, he has participated to several international projects : DEBORA(2001) Digital AccEss to BOoks of the RenAissance, European project for a better access to early printed documents of the Renaissance, VECMAS (2011) Valorisation of Arabic Manuscripts from Timbuktu, PHC VOLUBILIS (2014) Mathematical Models applied to robust recognition of texts in natural images and digital videos, GUWENSHIBIE (2015) Recognition of Historical Chinese Manuscripts with Tsinghua university of Beijing, several public national projects, MADONNE (2006) MAssE de DONnees issues de la Numerisation du patrimoiNE, GRAPHHEM (2011) Grapheme based Retrieval and Analysis for PaleographHic Expertise of Middle Age manuscripts, for digital Palaeography assistance, DIGIDOC (2014) Document Image diGitisation with Interactive DescriptiOn Capability, ORIFLAMMS (2015) Analysis on Multilingual Medieval manuscripts, DREAM (2015) : Digital Restoration of mEdievAl Manuscripts CNRS. He also participates directly to several industrial projects since 2000 : BookRestorer (2001) with I2S for post processing solution to restore scanned documents, RASADE (2003) with EVER TEAM for an Automatic Recognition of structures of documents, CESA (2008) with VINCI group for the Improvement of postal mail sorting system, MEDIABOX (2011) with SPIGRAPH for advertisement detection in digitized press, DOD (2015) Documents On Demand with ITESOFT for Business documents recognition, OZALID (2015) with ORANGE for Crowdsourcing to correct errors of the OCR with the French National Library. He has co-supervised 12 PhD during the past years in the field of documents imaging until today. He actually employs 3 Post Doc on different Research projects. He is a co-founder of the start-up CORENUM in 2009 <http://www.corenum.fr/> and he participates to the implementation of LibCrn <http://sourceforge.net/projects/libcrn/> and PLEIAD <http://liris.cnrs.fr/pleiad/> platforms for researchers.

<http://liris.cnrs.fr/membres/?id=834>

Multi-Script Document Processing

Umapada Pal

A single camera/video image may contain two or more scripts and the presence of such multiple scripts in a single image makes the text recognition task of the image challenging. Moreover video frames suffer from low resolution, blur, complex background etc. and these factors add extra difficulty in recognition task. Although, there are techniques available for multi-script recognition of scanned document images but only a few pieces of work have been done towards the text recognition of multi-script camera/video images. In this talk we will discuss about text recognition techniques from multi-script camera/video images. There are two approaches for multi-script text recognition: (i) develop a generalized recognition system which can recognize characters of all scripts (ii) first identify portions of the different scripts present in an image and then use script-specific recognition system on the identified portion. First approach is very difficult and it is, in fact, not possible to develop such a system when number of scripts to be considered is large. In this talk we will discuss different issues towards these approaches. For better understanding, demonstrations of some systems will be given to the audience. Finally, some open and new challenging problems will be discussed.

Umapada Pal received his Ph.D. in 1997 from Indian Statistical Institute. His PhD work was on the development of printed Bangla OCR system. He did his Post Doctoral research at INRIA (Institut National de Recherche en Informatique et en Automatique), France. From January 1997, he is a Faculty member of Computer Vision and Pattern Recognition Unit of the Indian Statistical Institute, Kolkata and at present he is a Professor. His fields of research interest include Digital Document Processing, Optical Character Recognition, Camera/video text processing, Biometrics, etc. He has published more than 260 research papers in various international journals, conference proceedings and edited volumes. Because of his significant impact in the Document Analysis research, in 2003 he received "ICDAR Outstanding Young Researcher Award" from International Association for Pattern Recognition (IAPR). In 2005-2006 Dr. Pal has received JSPS fellowship from Japan government. In 2008, 2011 and 2012, Dr. Pal received visiting fellowship from Spain, France and Australia government, respectively. Dr. Pal has been serving as General/Program/Organizing Chair of many conferences including International Conference on Document Analysis and Recognition (ICDAR), International Conference on Frontiers of Handwritten Recognition (ICFHR), International Workshop on Document Analysis and Systems (DAS), Asian Conference on Pattern recognition (ACPR) etc. Also he has served as a program committee member of more than 50 international events. He has many international research collaborations and supervising Ph.D. students of 5 foreign universities. He is serving as associate Editor of the journal of ACM Transactions of Asian Language Information Processing (ACM-TALIP), Pattern recognition Letters (PRL), Electronic Letters on Computer Vision and Image Analysis (ELCVIA) etc. Also he

has served as a guest editor of several special issues. He is a Fellow of IAPR (International Association of Pattern Recognition).

<http://www.isical.ac.in/~umapada/>

Indexing

Jean-Marc Ogier

This talk give provide an overview concerning the main approaches of the literature concerning indexing techniques. These methods will be discussed according to their context of utilization.

The outline of the presentation will be the following

1. Introduction : challenge related to document indexing
2. Space-partitioning-based methods : KD Trees, NKD and PKD Trees, PA Trees, R Trees, R* Trees, SR Trees, X Trees
3. Clustering-based methods : K-Means Clustering Trees, Agglomerative clustering Trees, K Medoids clustering trees,
4. Hashing-based methods : Locality Sensitive Hashing,(LSH), Kernelized LSH, Entropy based LSH, Multi-probe LSH
5. Latent Semantic Indexing (LSI)
6. Main schema related to Information spotting problems : from information characterization to indexing
7. Conclusion - discussion

Jean-Marc Ogier received his PhD degree in computer science from the University of Rouen, France, in 1994.

During this period (1991-1994), he worked on graphic recognition for Matra Ms&I Company. From 1994 to 2000, he was an associate professor at the University of Rennes 1 during a first period (1994-1998) and at the University of Rouen from 1998 to 2001. Now full professor at the university of la Rochelle, Pr Ogier is the head of URL laboratory which gathers more than 120 members and works mainly of Document Analysis and Content Management.

Author of more than 200 publications / communications, he managed several French and European projects dealing with historic document analysis, either with public institutions, or with private companies. Pr Ogier is a Deputy Director of the GDR I3 of the French National Research Centre (CNRS). He is also Chair of the Technical Com-

mittee 10 (Graphic Recognition) of the International Association for Pattern Recognition (IAPR), and is the representative member of France at the governing board of the IAPR. At last he is also Vice rector of the university of La Rochelle.

<http://pageperso.univ-lr.fr/jmogier/>

Performance Evaluation

Efstathios Stamatatos

Common evaluation frameworks for Document Image Analysis will be presented. Evaluation techniques used in DIA competitions and systems will be discussed focusing on their pros and cons. Moreover, current evaluation techniques from other related research fields and evaluation campaigns will be presented. Through this presentation the students will get familiar with several evaluation techniques and methodologies and will understand their strengths and limitations.

Efstathios Stamatatos received the diploma degree in electrical engineering (1994) and the doctoral degree in electrical and computer engineering (2000), both from the University of Patras, Greece. He was a research associate of the Wire Communications Lab. of the University of Patras from 1995 to 2003. He also joined the Polytechnic University of Madrid (1998) as a visiting researcher, the Austrian Research Institute for Artificial Intelligence as a post-doc researcher (2001-2002), and the TEI of Ionian Islands as an adjunct professor (2002-2004). Since 2004 he is a faculty member of the Dept. of Information and Communication Systems Engineering, University of the Aegean.

<http://www.icsd.aegean.gr/lecturers/stamatatos/>

Robust Reading

Dimosthenis Karatzas

Robust reading, or else the ability of machines to read text in unconstrained settings such as scene images and videos, has been an important application field of document image analysis for the past 20 years. This lecture will give an introduction on robust reading systems, with a special focus on the detection and segmentation of text in real scene images.

We will review classical pipelines for text detection based on sliding window and patch based text/non-text classification, and will study the evolution of these pipelines towards selective search approaches. The nature of text will be explored and we will

demonstrate the importance of context in text extraction, and the tendencies stemming from perceptual organisation. The current OpenCV implementation of text extraction algorithms will be reviewed as an open, state of the art methodology that can be used as the basis for further research.

Finally, the lecture will give an overview of the last five years of Robust Reading Competition activity, reviewing open tools and methodologies for performance evaluation. Lessons to be learnt after receiving over 2500 submissions to the Robust Reading Competition over the past year will be outlined and an outlook to the future of robust reading will be offered.

CONTENTS (preliminary):

- Sliding window approaches for text detection
- Patch based text/ no text classification, and limitations
- Text segmentation / segmentation for recognition
- Selective search approaches for combined text detection and segmentation
- Open source state of the art resources (OpenCV implementation)
- The Robust Reading Competition – overview and lessons learnt
- Performance evaluation protocols for text detection, segmentation, word recognition
- Next steps in robust reading: End to end systems (full recognition or word spotting), from focused text to accidental text.

Dimosthenis Karatzas is a “Ramon y Cajal” Research Fellow at the Computer Vision Centre, Universitat Autònoma de Barcelona, within the Document Analysis Group (DAG). He received his degree in Physics from the Aristotle University of Thessaloniki, Greece in 1998 and his PhD in Computer Science titled “Text Segmentation in Web Images Using Colour Perception and Topological Features”, from the University of Liverpool, UK in 2003.

From 2002 to 2004 he worked as a Research Associate within the Pattern Recognition and Image Analysis group (PRImA) at the University of Liverpool. From 2005 to 2007 he worked as a Research Fellow in the School of Electronics and Computer Science, University of Southampton within the Electronics Systems Design (ESD) and Image, Speech and Intelligent Systems (ISIS) research groups.

His main research interests are computer vision and colour science and in particular colour and historical document image analysis, perceptual methodologies for DIA, and colour calibration. He has produced over 40 peer reviewed publications in journals, book chapters and international conferences.

Dr Karatzas is the director of the spin-off company TruColour, which specialises on perceptually based colour calibration solutions. The company was setup in 2007 and has since received venture capital funding and separate funding for prototype development. He has also provided consultancy and technical support to AIMES, UK and Perpetuum

Ltd, UK.

Dr Karatzas has been the primary investigator in 4 research and knowledge transfer projects (two of them currently active). He serves as a Publicity Chair of ACPR 2013 and Publications Chair of ICDAR 2009 and is serving on the program committees of ICDAR, DAS, PRIS and CVCRD. Dr Karatzas has been a founding member and a member of the executive committee of the UK Chapter of the SPIE, while he is currently a member of the leadership committee of TC-11 (Reading Systems), a member of the IAPR -Industrial Liaison Committee and a member of IEEE, the SPIE and the IAPR.

<http://dag.cvc.uab.es/people/dimosthenis-karatzas>

Camera-based OCR

Seiichi Uchida

Now OCR is extending its target from scanned documents to scene images. Please look around and find texts around you --- there may be various texts printed in various fonts. Thinking about scene texts will provide us many interesting problems, which never happened on the conventional OCR for scanned documents.

First, how about recognition of scene texts? It is a difficult and challenging problem because the font shapes are very different from each other even though they come from the same category. Reviewing the recent scene text/character recognition technologies will introduce us how we can tackle with this huge shape variations.

Second, how about "non-verbal" information given by a specific font? For example, please recall the logo of a famous gorgeous brand. Is the font used in the log the same as that in a toyshop? The answer is, maybe, "no". This suggests that each font has its own "atmosphere", which is typical "non-verbal" information.

Third, how about "visual perception of scene texts?" Most scene texts are put in the scene for providing some message to people. They, therefore, should be "eye-catching", or, visually salient. Recently, this visual saliency is a hot topic of computer vision and thus it is not bad to examine that scene texts are really salient or not.

Fourth, how about concealing texts? Some scene texts, such as a car license plate, are privacy sensitive and we should conceal them before publishing the image containing them. How can we "damage" them to make them unreadable with less side-effect?

Like the above four problems, we can find many other "new" problems around scene texts. I hope that the participants will enjoy discussion on some practical issues and less practical issues on scene texts for a wider view.

Seiichi Uchida received B.E., M.E., and Dr. Eng. degrees from Kyushu University in 1990, 1992 and 1999, respectively. From 1992 to 1996, he joined SECOM Co., Ltd., Ja-

pan. Currently, he is a professor at Kyushu University, Fukuoka, Japan. His research interests include sequential pattern recognition, character recognition, image processing, optimization, and bioimage-informatics. He received 2007 IAPR/ICDAR Best Paper Award and 2010 ICFHR Best Paper Award, in addition to many domestic awards from Japanese academic societies. He is a workshop chair of ICDAR2015 and ACPR2015 and has been a general co-chair of CBDAR2009, a program chair of DAS2013, an area chair of ICPR2012, and a regular program committee member of ICPR, ACCV, ICDAR, S+SSPR, DAS, ICFHR and CBDAR. Recently, he wrote a book chapter titled "Text Localization and Recognition in Images and Video" for "Handbook of Document Image Processing and Recognition", Springer-Verlag.

<http://human.ait.kyushu-u.ac.jp/~uchida/index-e.html>

COMPANY PRESENTATIONS

The Historical Archives of Cephalonia: problems and limitations in the digitized collections research

Dr. Dora Zafeiratou

Director in General State Archive

The various categories of historical documents kept in General State Archives- Archives of Cephalonia are briefly presented with emphasis on those that are already digitized. Problems and limitations faced by the historian in his research of the digitized archival material are analyzed.

Case Study of a Commercial DIA System

Anagnostopoulos Tassos

Processing of hand-written prescriptions for the Greek Social Security Agency.

We will have a detailed look into the implementation of a system for the automated processing of hand-written documents for the Greek Social Security Agency, which is one of the largest organizations in Greece. The system needed to be able to process over 200,000 hand-written prescriptions on a daily basis, with minimum user input. We will explain the key challenges that had to be addressed during the implementation of the project, along with the approach that was used to solve each one of them. The presentation will be an interactive one, and we will be eliciting feedback from the audience while describing the issues that we encountered

From laboratory to market :

Natural Intelligent Technologies' experience

Adolfo Santoro

POSTER ABSTRACTS

Automated Detection of Neuropsychological Impairments from Image-Based Visuo-constructive Screening Tests

Momina Moetesum

Drawing tests have been long used by practitioners and researchers for early detection of psychological and neurological impairments. These tests allow subjects to naturally express themselves as opposed to an interview or a written assessment. Some neuropsychological testing procedures are lengthy and comprise of comprehensive batteries including a broad array of subtests like the Halstead-Reitan Neuropsychological Test Battery (HRB) and the Luria-Nebraska Neuropsychological Battery (LNNB) while others are much briefer and are typically used as screens for neuropsychological impairment rather than full-fledged assessment tools like the Bender Visual-Motor Gestalt Test (BGT) and the Rey-Osterrieth Complex Figure Test (ROCF). Nevertheless a common feature amongst these tests is the use of figure or shape drawing. Poor performance on these tests is indicated by a variety of errors that subjects may make while drawing the figures, including missing details, collision or overlapping, inability to accurately complete shapes like circles or squares, disproportionate size and angles or orientation. The patterns of scores on these subtests can go a long way towards pinpointing specific cognitive weaknesses. The manual scoring of these tests, however, is a time consuming and lengthy procedure especially when a large number of subjects is to be analyzed. This limits their applicability in the clinical psychology despite their numerous advantages. The purpose of this study is provide an automated solution which applies image analysis techniques to automatically score most of such tests in order to facilitate psychologists.

Equivalence between two document images at hierarchical levels through structural attributes

Sahana Deve Gowda

A lot of memory is consumed most of the time because multiple copies of the document images are saved. The reason for this could stretch from more copies for safety to downloading the same content from different origins. Even all copies of slightly edited versions involving simple modifications such as reshuffling of blocks/paragraphs, splitting/merging of paragraph, change in layout formats are retained. Detection of these multiple copies using Optical Character Recognition (OCR) could be possible for automating the process provided the document images are entirely composed of text in one language and the language of the text is a-priori known. If the document image contains tables, images and graphs and it is multilingual for which an OCR may not exist. Hence

in this research work different components composing a document are extracted and are recursively decomposed into finer possible levels. At every level suitable structural attributes have been derived.

Depending upon the possible types of equivalence between two document images, the two documents under test are said to be Identically equivalent document images –if and only if both the documents exhibit equivalence in terms of both layout structure and content; Content equivalent document images – if and only if both the document images exhibit equivalence in terms of content but not in terms of layout; Layout equivalent document images – if and only if both the document images exhibit equivalence in terms of layout but not in terms of content; else the two document images are said to be dissimilar.

In this research work a new tight segmentation procedure is evolved, which automatically accomplishes segmentation of components at all hierarchical levels in a top down sequence, starting with the entire page as a component. An intelligent analysis of the volume of blank space existing between the components induces the automation in the process of tight segmentation.

Every text component at different hierarchical levels is characterized in terms of two types of attributes-geometrical structural attributes and content structural attributes. The geometrical structural attributes are used to analyze the position and span content of the component. The content structural attributes are used to analyze the text content of the component. Content of a component is interpreted by its structure as the foreground overlay on the null background of the document image. When there is a superimposition of the component structure, the intensity value of the corresponding null pixel changes, thereby causing a sudden change in intensity value which can be encountered during the run length smearing either in horizontal or vertical direction. The sudden change in intensity values are regarded as the points of transition. The transition points may be regarded as points, exhibiting changes in the level of energy. Therefore such locations of transitions may be identified as packets of energy or energy chunks. Analyzing these energy patterns would enable the understanding of the content. Hence energy values act as content structural attributes. The energy values are quantified using conventional entropy quantifier (CEQ) and spatial entropy quantifier (SEQ) based on the frequency of occurrence and position of occurrence respectively. However the large mass of entropy values generated by CEQ and SEQ cause computational concern. The problem is alleviated by assimilating the entropy values into frequency histograms and spatial histograms. The regression based distance function measures the match between the histograms of these text components for understanding the equivalence or otherwise.

The frequency histograms and spatial histograms of conventional-entropy values and frequency histograms and spatial histograms for spatial entropy values are compared with their respective counterparts, to establish the equivalence between the text content of the components in the document images.

Apart from establishing the equivalence between the text components the problem of establishing the equivalence between the corresponding non-text components- tables, graphs and images has been achieved. Further, suitable extension algorithms are devised

to withstand the scale up/scale down of the components and to establish equivalence between the text components having different font sizes, however having same font style. In sequel, the newly developed models are extended to solve some applicative problems. Two such problems termed as Advertiser's problem and Information seeker's problem are formulated and customized algorithms are devised.

Image Document Analysis of Ancient Palm Leaf Manuscripts

Made Windu Antara Kesiman

One very valuable cultural heritages that are found in Southeast Asia is the collection of palm leaf manuscripts. Ancient palm leaf manuscripts store various forms of knowledge and historical records of social life in Southeast Asia. But unfortunately, the physical condition of natural materials from palm leaves certainly cannot last long and certainly cannot fight time. Usually, palm leaf manuscripts are of poor quality since the documents have degraded over time due to storage conditions. The ancient palm leaf manuscript is written on a dried palm leaf by using some sort of sharp pen or small knife, which is then scrubbed with natural dyes. Due to these specific characteristics, ancient palm leaf manuscripts are providing new challenges in document analysis. The automatic analysis of these documents, in order to extract relevant information, is a real challenge. This research develops a document image analysis system for document images of palm leaf manuscripts, that includes several image processing tasks, beginning with digitization of the document, ground truth construction and ending with character recognition and document retrieval.

Keywords— document analysis; palm leaf manuscript; low quality document image; ground truth construction;

AWARDS

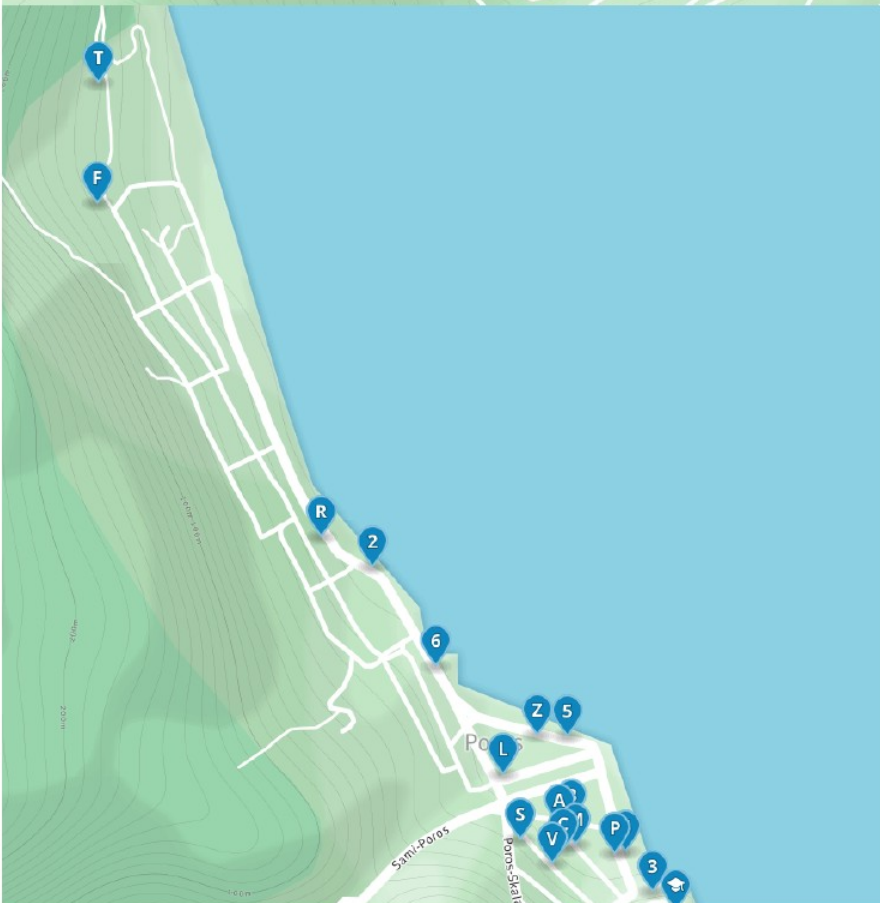
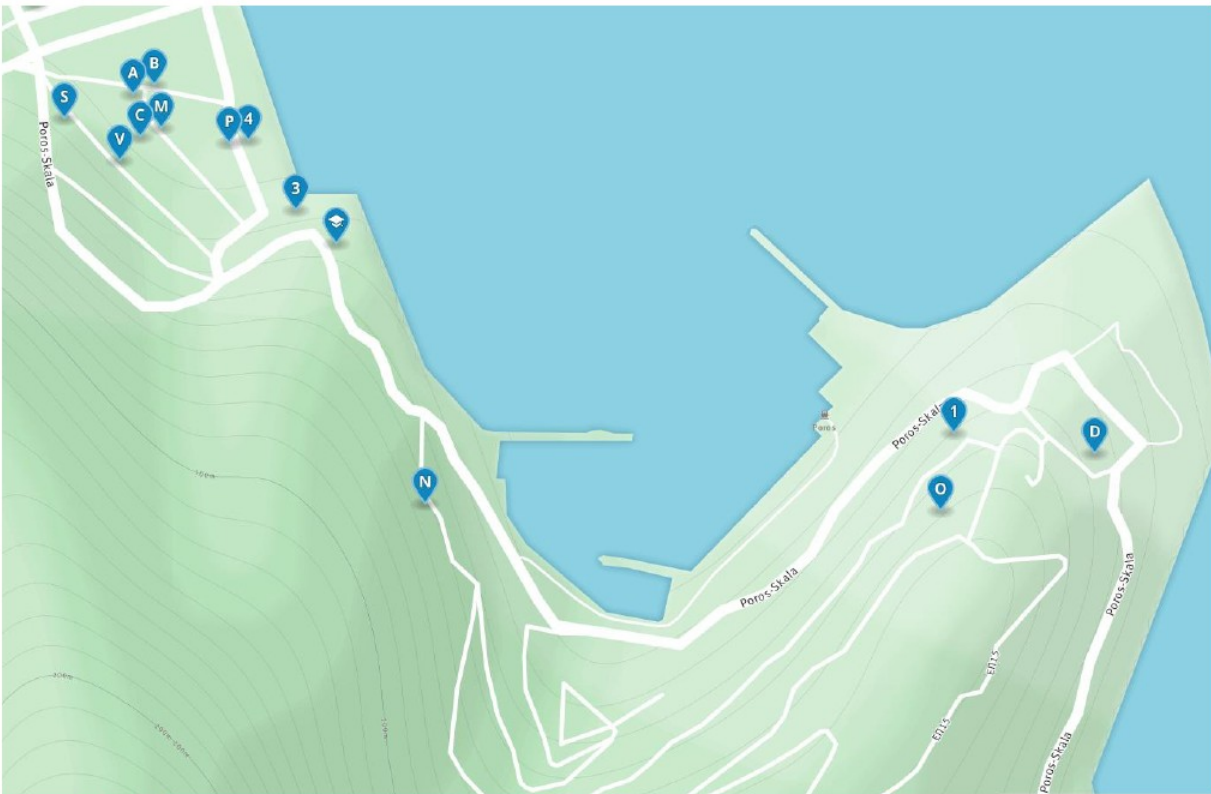
There will be 2 awards of 500 euros:

The **best poster award**, offered by Ailab, will be selected by a scientific committee during the poster session.

The **excellence award**, offered by Natural Intelligent Technologies s.r.l.: the winner will be selected among the participants for his knowledge, collaboration ability and his work in the practice team.



MAP OF POROS



-  Lecture Hall
- 1 Restaurant Agrapidos
- 2 Restaurant Memas
- 3 Restaurant Fotis Family
- 4 Restaurant Pantelis
- 5 Restaurant Steki
- 6 Restaurant M'anavis
- A. Alecos Studios
- B. Anemos Studios
- M. Ariston studios
- C. Captain Georgio
- D. Dioni
- F. Filoxenia
- L. Louisa
- N. Oceanis Apartments
- O. Oceanis Hotel
- P. Pantelis Rooms
- R. Riviera
- S. Santa Irena
- T. Travlos
- V. Vavasi Margarita
- Z. Zanza Apartments

USEFUL TELEPHONE NUMBERS

Poros Harbor-Police Station	2674 072460
Poros Police Station	2674 072210
Taxi Station in Poros	2674 072340
	2674 072329
Bus station in Poros	2674 072483
Municipality of Argostoli	2671 022847
Argostoli Police Station	2671 022200
General Hospital in Argostoli	2671 024641-6
Ticket Agents - Ainos Travel	2671 022333

