



University of  
the Aegean  
**SUMMER  
SCHOOLS**

**INTERNATIONAL DOCUMENT  
IMAGE PROCESSING  
SUMMER SCHOOL 2014  
26 - 30 MAY 2014  
FOURNI GREECE**

*Informations*



## INDEX

ABOUT THE SUMMER SCHOOL.....	4
VENUE .....	5
PEOPLE.....	7
SCHOOL PROGRAM .....	8
MAP OF FOURNI .....	28
USEFUL TELEPHONE NUMBERS.....	29

## ABOUT THE SUMMER SCHOOL

After the successful first Summer School on Document Image Processing (IDIPS 2013), (<http://www.iapr.org/docs/newsletter-2013-03.pdf> p.13-14) the second Summer School on Document Image Processing (IDIPS2014) aims to provide both an objective overview and an in-depth analysis of the state-of-the-art of the Document Analysis Systems (DAS) and their current open issues. The courses will be delivered by world-renowned researchers in the field and relevant companies, covering both theoretical and practical aspects of real problems of Document Analysis Systems, as well as examples of successful applications.

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 their colleagues in a friendly and constructive atmosphere.

The Summer School is hosted on the beautiful island of Fourni which ensures quiet and beautiful environment, with quality food services and housing, which would enjoy the students and lecturers, while the participants will have the opportunity to experience the Greek hospitality.

## VENUE

The Summer School will take place in the City Hall of Fourni. Fourni can be reached by ship by the ports of Piraeus or Karlovassi (Samos). The Organizers will take care of the transportation, from the airport Aristarchos of Samos, of the participants that will arrive there by June 2nd, before 3.00 p.m., and they will be transferred back on the June 8th, about 10.00 a.m. The airport of Samos has direct flights from/to Athens, Germany, Netherlands, England, Belgium, Scandinavian Countries, Switzerland etc.

Fourni Korseon (Greek: Φούρνοι Κορσέων or Greek: Φούρνοι Ικαρίας - Fourni Ikarias), more commonly simply Fourni, is a complex or archipelago of small Greek islands that lie between Ikaria, Samos and Patmos in Ikaria regional unit, North Aegean region. The two largest islands of the complex, the main isle of Fourni (31 km<sup>2</sup>) and the isle of Thymaina (10 km<sup>2</sup>), are inhabited, as is Agios Minas Island (2.3 km<sup>2</sup>) to the east. On the main isle Fourni (town) is the largest settlement and then Chrysomilia in the north the second largest (and third largest overall, after Thymaina). Fourni (town) proper is the main ferry harbor, with ferries also landing on Thymaina.

Fourni, as proven by various archeological sites and other findings on the islands, have been the home of a civilization dating back to the 3rd century B.C. 'till the 2nd to 3rd century A.C.

The islands are uniquely characterized by their extremely complicated coastline, the total length of which is 126 kilometers, while the area of the islands counts merely 48 sq. kilometers.

They're run through by easy to follow asphalt roads, while paths and traditionally built steps lead to lots of beaches with transparent waters, such as Vlichada, Kassidi, Petrokopio, Elidaki, Vitsilia and more.

The highest peaks are Korakas (514m.), Fanas (476m.) and Pappas of Thymaina (470 m.).

Many of the inhabitants are fishermen, although during the summer season the population is also occupied in tourist activities, mostly room rentals and catering.

The climate is arid and hot during summers. Winters are rather mild with average rainfall but constant strong archipelagic winds prevail.

You'll enjoy swimming in exotic waters, staying in hospitable rooms, tasting traditional Greek dishes, fresh fish and lobsters, as well as local honey and cheese of fine quality.

But most of all, we are sure you will discover the beauty of simplicity and serenity, at a place extremely authentic, in a way only Greek islands can be.

# PEOPLE

## *Program Chairs*

Apostolos Antonacopoulos, *PRImA, University of Salford*

Ergina Kavallieratou, *University of the Aegean*

Josep Lladós, *CVC, University Autònoma de Barcelona*

## *Local Organizing Committee*

Ergina Kavallieratou, *University of the Aegean*

Efstathios Stamatatos, *University of the Aegean*

Irini Stathi, *University of the Aegean*

## *Program Committee*

David Doermann, *University of Maryland*

Daniel Lopresti, *Lehigh University*

Jean-Marc Ogier, *Universite de La Rochelle*

Ioannis Pratikakis, *Democritus University of Thrace*

## SCHOOL PROGRAM

Every morning there will two lectures while in the evening either student presentation (Monday) or practice in laboratories (Tuesday, Thursday, Friday) will take place. The last day (Friday), the participants will have the chance to attend the presentation of several case studies.

### **Keywords of the Lectures**

#### System architectures

- end-to-end document analysis systems
- Common issues & implications
- performance evaluation

#### Pre-processing

- Binarization
- Noise removal
- Enhancement
- Show through cancellation
- Geometric correction (page curl removal, Dewarping, Slant&Skew correction)
- Skeletonization



## Layout analysis

- Segmentation and region classification (physical layout)
- Logical layout analysis. Reading order.

## Recognition

- Machine printed text recognition
- Handwriting recognition
- Graphics recognition

## Evaluation

- Datasets
- Metrics
- Examples of performance evaluation methodologies and protocols

## Indexing

- Inverted files
- Hashing
- Latent semantic indexing (LSI)
- information spotting

## Document classification

- Document classification
- Document retrieval

## System Level Issues in Document Image Analysis

*Daniel Lopresti*

Document image analysis techniques are not used in isolation, but rather to solve specific tasks. In this lecture, we shall survey a variety of architectures used when assembling end-to-end document analysis systems, both traditional and novel. We will examine how the nature of the task places demands on the degree of automation that is required, the minimum acceptable accuracy level, and the kinds of errors that can be tolerated. Also discussed will be the implications for performance evaluation and, ultimately, the success of the system.

After attending this lecture, students will understand:

- Representative architectures that are used when building end-to-end document analysis systems.
- Common issues that arise in such systems.
- How to evaluate the performance of such systems.
- Implications for conducting research on a component technology used in such systems.

**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/>

## **Document Image Binarisation**

*Ioannis Pratikakis*

Document image binarisation is the initial step of most document image analysis and understanding systems that converts a grey scale image into a binary image aiming to distinguish text areas from background areas. Binarisation plays a key role in document processing since its performance affects drastically the degree of success in subsequent character segmentation and recog-

nition. Although for modern machine printed documents, binarisation is considered as a problem that has already been solved, when processing degraded document images, binarisation is not an easy task. This advocates the intensive contemporary research efforts for the binarisation of historical document images.

This tutorial aims to survey the main trends on the document image binarisation process combining the theory with hands-on demonstrations. Last but not least, the presentation of established evaluation measures for document image binarisation methods will be addressed as the means to study the algorithmic behaviour by providing qualitative and quantitative performance indication

**Ioannis Pratikakis** is an Assistant Professor at the Department of Electrical and Computer Engineering of Democritus University of Thrace in Xanthi, Greece. He received the Ph.D. degree in 3D Image analysis from the Electronics engineering and Informatics department at VrijeUniversiteit Brussel, Belgium, in January 1999. From March 1999 to March 2000, he joined IRISA/ViSTA group, Rennes, France as an INRIA postdoctoral fellow. From January 2003 to June 2010, he was working as Adjunct Researcher at the Institute of Informatics and Telecommunications in the National Centre for Scientific Research "Demokritos", Athens, Greece. His research interests lie in image processing, pattern recognition, vision and graphics, and more specifically, in document image analysis and recognition, medical image analysis as well

as multimedia content analysis, search and retrieval with a particular focus on visual content. He has served as co-chair of the Eurographics Workshop on 3D object retrieval (3DOR) in 2008 and 2009 as well as Guest Editor for the Special issue on 3D object retrieval at the International Journal of Computer Vision. He will be the co-organiser of the International Conference on Frontiers in Handwriting Recognition (ICFHR 2014). He is Senior Member of the IEEE, member of the Board of the Hellenic Artificial Intelligence Society for the period 2010-2012 and member of the European Association for Computer Graphics (Eurographics).

<http://utopia.duth.gr/~ipratika/>

## **Recognition of Textual and Graphical Patterns**

*Josep Lladós*

Documents contain two main categories of information, namely text and graphics. The recognition of text is solved by Optical Character Recognition (OCR) when the text is machine printed. OCR is a mature research topic, and commercial software exists offering high levels of performance in documents having traditional Manhattan layouts (text blocks are strictly oriented horizontally or vertically). A challenge is the recognition of textual parts is when the text is handwritten. Graphics recognition deals with the interpretation of graphical parts (symbols, logos, lines in tables or forms,

etc.). In this lecture we will review the main techniques addressed to recognize both sources of information. The lecture will be structured in two blocks. First, basic OCR techniques will be reviewed. We will review the traditional shape descriptors and classifiers used in the literature to recognize machine printed text. We will complete the textual analysis block reviewing techniques for handwriting recognition. In the second part we will review the main graphics recognition techniques such as vectorization and symbol recognition. Finally, we will address the problem of documents combining text and graphics, and review the problem of text-graphics separation.

**Josep Lladós** received the degree in Computer Sciences in 1991 from the Universitat Politècnica de Catalunya and the PhD degree in Computer Sciences in 1997 from the Universitat Autònoma de Barcelona (Spain) and the Université Paris 8 (France). Currently he is an Associate Professor at the Computer Sciences Department of the Universitat Autònoma de Barcelona and a staff researcher of the Computer Vision Center, where he is also the director since January 2009. He is the head of the Pattern Recognition and Document Analysis Group (2009SGR-00418). He is chair holder of Knowledge Transfer of the UAB Research Park and Santander Bank. His current research fields are document analysis, graphics recognition and structural and syntactic pattern recognition. He has been the head of a number of Computer Vision R+D projects and published more than 150 papers in

national and international conferences and journals. J. Lladós is an active member of the Image Analysis and Pattern Recognition Spanish Association (AERFAI), a member society of the IAPR. He is currently the chairman of the IAPR-ILC (Industrial Liaison Committee). Formerly he served as chairman of the IAPR TC-10, the Technical Committee on Graphics Recognition, and also he is a member of the IAPR TC-11 (reading Systems) and IAPR TC-15 (Graph based Representations). He is editor in chief of the ELCVIA (Electronic Letters on Computer Vision and Image Analysis) and he serves on the Editorial Board of the IJDAR (International Journal in Document Analysis and Recognition), and also a PC member of a number of international conferences. He was the recipient of the IAPR-ICDAR Young Investigator Award in 2007. He was the general chair of the International Conference on Document Analysis and Recognition (ICDAR'2009) held in Barcelona in July 2009, and co-chair of the IAPR TC-10 Graphics Recognition Workshop of 2003 (Barcelona), 2005 (Hong Kong), 2007 (Curitiba) and 2009 (La Rochelle). Josep Lladós has also experience in technological transfer and in 2002 he created the company ICAR Vision Systems, a spin-off of the Computer Vision Center working on Document Image Analysis, after winning the entrepreneurs award from the Catalonia Government on business projects on Information Society Technologies in 2000.

<http://dag.cvc.uab.es/people/josep-llados>

	Monday 26/5	Tuesday 27/5	Wednesday 28/5
9.30 - 11.00	System Architectures Lopresti	Document Image Binarisation Pratikakis	Recognition Llados
11.00 - 11.30	Coffee		
11.30 - 13.00	Document Image Preprocessing Kavallieratou	Layout analysis Doermann	Document Classification & Retrieval Doermann
13.00 - 13.30	Light Lunch		
13.30 - 16.00	Posters	Organizing Committee meeting	Lab (preprocessing)
16.00 - 18.00		EXCURSION	
18.00 - ...	Dance course/football		Fishing Contest



/5	Thursday 29/5	Friday 30/5	
	<b>Indexing</b> Jean-Marc OGIER	<b>Case study 1</b> <b>(Archives/Researchers'</b> <b>point of view)</b> Irini Stathi	<b>9.30 - 11.00</b>
<b>Break</b>			<b>10.45 – 11.45</b>
	<b>Document Analysis</b> <b>System Evaluation</b> Stamatatos	<b>Case study 2</b> <b>(Design of commercial</b> <b>DIA systems)</b>	<b>11.30 - 13.00</b>
<b>lunch Break</b>			<b>13.00 – 13.30</b>
ng)	Lab( recognition)	Summer School Evaluation	<b>13.30 - 16.00</b>
			<b>16.00 - 18.00</b>
t	<b>Dance</b> <b>course/beach</b> <b>volley</b>	<b>Goodbye Party</b>	<b>18.00 - ...</b>

## Indexing Techniques

*Jean-Marc Ogier*

Thus 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 :

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

**Jean-Marc Ogier** (07/04/1967) a obtenu son diplôme de doctorat en informatique de l'Université de Rouen en 1994, dans le cadre d'un partenariat avec la société Matra Systèmes et Informations. Depuis 2001, il est professeur titulaire à l'Université de La Rochelle. Les intérêts de ses recherches actuelles portent sur la reconnaissance, l'interprétation, et l'indexation de documents numériques.

Jean-Marc Ogier est directeur du laboratoire L3I de l'Université de la Rochelle, qui comporte 90 chercheurs. Il a été le leader national ou international plusieurs projets nationaux financés par l'Agence Nationale pour la recherche nationale (MADONNE, NAVIDOMASS, ...) et par la communauté européenne (Eureka !). Il est directeur adjoint du GDR I3 du CNRS depuis 2005 (structure regroupant environ 800 personnes). Il est le président général du comité technique numéro 10 (reconnaissance graphique - 300 chercheurs dans le monde) de l'Association internationale pour la Reconnaissance des Formes (IAPR), et a été membre du Governing Board de l'IAPR (association internationale scientifique regroupant plus de 10 000 chercheurs dans le monde), représentant la France. Il est également vice-président de l'Université de La Rochelle. Il est par ailleurs président de l'association VALCONUM, consortium académo-industriel visant la compétitivité française en termes de recherche et d'innovation. Il est auteur de plus de 200 publications

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

## Archives/Researchers' point of view

*Irini Stathi*

This lecture will focus on critical issues raising from archival policies regarding collection development, access, exhibition, cataloging, preservation, and restoration of audiovisuals. Introduction to principle models and methodologies of moving image archive practices such as collection development of classical, national, regional (small gauge formats, independent and amateur productions, new media) materials; changing role of technology in preservation and restoration; cultural impact of public programming and elaboration of possible metadata and useful information mined from the audiovisual text.

Finally, this lecture will examine and evaluate the current preservation standards for storage and duplication. Discussion of critical preservation problems and management of the data produced by digitization practices will be focused as well.

**Irini Stathi** Received a Laurea from DAMS, University of Bologna, in 1990 and the Ph.D in Cinema Theory and Communication from the Department of Mass Media and Communication, Panteion University of Athens, in 1996. She is an Associate Professor in Theory of Cinema and Audiovisuals at the Department of Cultural

Technology and Communication, at the University of the Aegean. Her research interest include: film theory and film history, audiovisuals and new technologies, relationship between cinema and the other arts, Film preservation. She has collaborated as well with the Greek director Theo Angelopoulos for more than 10 years and she is the responsible for the digitalization of his Archive.

[http://aviac.ct.aegean.gr/?page\\_id=172](http://aviac.ct.aegean.gr/?page_id=172)

## **Document Image Preprocessing**

*Ergina Kavallieratou*

By document image preprocessing, we refer to all the procedures before the main task of a Document Image Analysis (DIA) 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 DIA 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. 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 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 2012. Her research interests include Optical Character Recognition, Document Image Analysis, Computer Vision and Pattern Recognition.

<http://www.icsd.aegean.gr/lecturers/kavallieratou/cv.htm>

## Page Layout Segmentation and Analysis

*David Doermann*

Historically, documents in printed form have been organized into pages in a way that facilitates sequential browsing, reading and/or search. The structure of these documents varies by genre, there is no doubt that there

is an implicit organization that helps readers navigate the content in order to enable the transfer of information from the author to the reader.

As authors, humans have a canny ability to tweak the basic rules of organization to enhance the experience of the reader, and as readers we are able use this information implicitly to comprehend what we are reading. While ideally we should be teaching our document analysis systems to be able to parse structures and use them to their advantage, traditionally page analysis has simply been a way to divide up content in a way that enables content analysis algorithms to perform more efficiently. Pages are divided into zones and these zones labeled and analysis by content specific routines. Zones of text are divided into lines, lines of text into words, etc. As part of interpretation, the logical relationships between zones or regions on the page can be analyzed for reading order or as functional units of organization.

In this session we cover two fundamental problems of Page Segmentation and Logical Layout Analysis. While traditional these two tasks have been addressed in sequence, it is more effective to fully understand the challenges that face the community including dealing with handwritten and degraded material

## Document Image Classification and Retrieval

*David Doermann*

Traditional approaches to document retrieval focus on conversion to electronic text followed by indexing of the text content. Recently some work in the community has focused on indexing document image content directly. In this talk, we will overview work at Maryland on Classification and Indexing that scales to millions of documents. First we present a learning based approach for computing structural similarities among document images for unsupervised exploration in large document collections. The approach is based on multiple levels of content and structure. At a local level, a bag-of-visual words based on SURF features provides an effective way of computing content similarity. The document is then recursively partitioned and a histogram of codewords is computed for each partition. Structural similarity is computed using a random forest classifier trained with these histogram features. We experiment with three diverse datasets of document images varying in size, degree of structural similarity, and types of document images. Second, we present a scalable algorithm for segmentation free content retrieval in document images. The contributions of this paper include the use of the SURF feature for image passage retrieval, a novel indexing algorithm for efficient retrieval of SURF features and a method to filter results using the orientation of local features and



geometric constraints. Results demonstrate that logo, signature block and stamp retrieval can be performed with high accuracy and efficiently scaled to a large datasets.

**Dr. David Doermann** is a senior research scientist in UMIACS. He received a B.Sc. degree in Computer Science and Mathematics from Bloomsburg University in 1987, and a M.Sc. degree in 1989 in the Department of Computer Science at the University of Maryland, College Park. He continued his studies in the Computer Vision Laboratory, where he earned a Ph.D. in 1993. Since 1993, he has served as co-director of the Laboratory for Language and Media Processing in the University of Maryland's Institute for Advanced Computer Studies and as an adjunct member of the graduate faculty.

His team of researchers focuses on topics related to document image analysis and multimedia information processing. Recent intelligent document image analysis projects include page decomposition, structural analysis and classification, page segmentation, logo recognition, document image compression, duplicate document image detection, image based retrieval, character recognition, generation of synthetic OCR data, and signature verification. In video processing, projects have centered on the segmentation of compressed domain video sequences, structural representation and classification of video, detection of reformatted video sequences and the performance evaluation of automated video analysis algorithms.

In 2002 he received an Honorary Doctorate of Technology Sciences from the University of Oulu for his contributions to digital media processing and document analysis research. He is a founding co-editor of the International Journal on Document Analysis and Recognition, has the General Chair or Co-Chair of over a half dozen international conferences and workshops and is the General Chair of the International Conference on Document Analysis and Recognition (ICDAR) to be held in Washington DC in 2013. He has over 30 journal publications and over 125 refereed conference papers.

<http://www.umiacs.umd.edu/people/doermann>

## **Document Analysis System Evaluation**

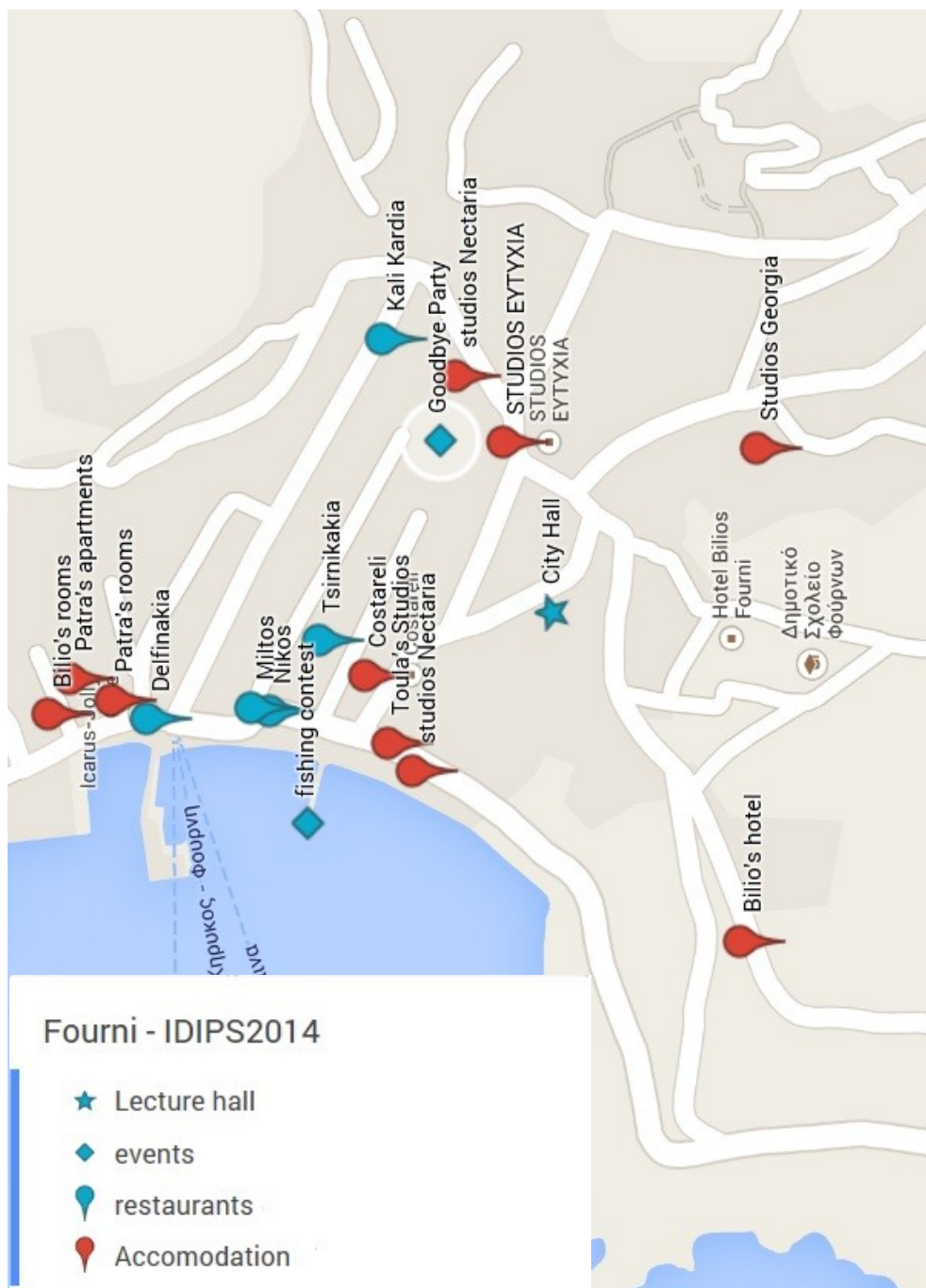
*Efstathios Stamatatos*

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/>



## USEFUL TELEPHONE NUMBERS

Municipality of Fourni	22753 50600
K.E.P (Citizen Service Center)	22750 50601
Police Department of Fourni	22750 51222
Port Authorities of Fourni	22750 51207
Medical Center	22750 51202
Ticket Agents	22750 51481-51546





**ΠΑΡΟΣ -  
ΑΝΤΙΠΑΡΟΣ Ν.Ε.**





**TC10**  
Technical Committee on  
Graphics Recognition

TC-11 **IAPR**  
Reading  
*Systems* 