

# Preface

Volume 11:

**Workshop on Applications of Pattern Analysis**

*Tom Diethe, Nello Cristianini, John Shawe-Taylor*

## Preface

Pattern Analysis and Statistical Learning cover a wide range of technologies and theoretical frameworks, and significant activity in the past years has resulted in a remarkable convergence and many advances in the theory and principles underlying the field.

Bringing these technologies to real world demanding applications is however often treated as a separate problem, one that does not directly affect the field as a whole. It is instead important to consider the field of Pattern Analysis as fully including all issues involved with the applications of this technology, and hence all issues that arise when deploying, scaling, implementing and using the technology.

The workshop called for contributions in the form of Demos, Case Studies, Working Systems, Real World Applications and Usage Scenarios. Challenges may stem from the violation of common theoretical assumptions, from the specific types of patterns and noise arising in certain scenarios, or from the problem of scaling up the implementation of state of the art algorithms to real world sizes, or from the creation of integrated software systems that contain multiple pattern-analysis components.

We were also interested in new application areas, where Pattern Analysis has been deployed with success, and in issues involving the visualisation and delivery and exploitation of the patterns discovered by PA technologies. Systems working in noisy and unstructured environments and situations are particularly interesting.

The goal was to discuss and reward work aimed at making theory useful and relevant, without requesting the researchers to propose new theoretical methods, but rather requesting to show how they solved the many challenges related to applying these methods to real world scenarios, or how they benefited other fields of research. Getting ideas to work in real scenarios is what this is about.

We believe that the papers are of a high standard and meet the goals of the workshop.

## Invited Talks

1. Facial Image Analysis using Directional Statistics and Shape-from Shading  
Edwin Hancock, University of York

*Abstract: Although the recovery of facial shape using shape-from-shading is an appealing idea, it is frustrated by problems such as concave-convex ambiguities, variable albedo, self shadowing and non-Lambertian reflectance. As such the devil resides in the detail. In this talk I will show how these problems can be overcome by incorporating a statistical model for surface normal direction within the shape-from-shading process. The main contribution of the talk is to develop a representation of the distribution of surface normals using the equidistant azimuthal projection from cartography, which transforms a distribution of surface normal direction on a unit sphere to a distribution of points on a tangent plane. I will show how this model can be adapted*

to deal with shadowing by fitting the statistical model to image brightness data using robust statistics. I will also show to adapt the process to deal with non-Lambertian reflectance, through fitting a reflectance model that can capture the behaviour of both shiny and rough surfaces. Finally, I will show how the shape information delivered by the process can be used to perform face recognition and gender determination. This talk will provide a synopsis of recent work by Smith and Hancock (PAMI 07, IJCV09, IJCV 2010) and Wu, Smith and Hancock (IVC 2010).

2. TrueSkill and AdPredictor: Large Scale Machine Learning in the Wild  
Thore Graepel, Microsoft Research

Abstract: *Probabilistic Graphical Models play a crucial role in Microsoft's online services. In this talk, I will describe two powerful applications of machine learning in practice. TrueSkill is Xbox Live's Ranking and Matchmaking system and ensures that gamers online have balanced and exciting matches with equally skilled opponents. AdPredictor is the system that estimates click-through rates (CTR) for ad selection and pricing within Microsoft's search engine Bing. The two systems have in common that they are based on factor graph models and approximate Bayesian inference. They operate at a very large scale involving millions of gamers and billions of ad impressions, respectively. However, in this talk, I will put particular emphasis on those aspects of these applications that are not part of the generic machine learning setting: a) The difficulties that arise because these are closed-loop systems in which the predictions determine the future composition of the training sample. b) The consequences of the fact that these systems make decisions that have an impact on more or less rational agents (advertisers, users, gamers) with the ability to influence the training sample. Time permitting, I will show the two systems in action. This is based on joint work with Ralf Herbrich, Thomas Borchert, Tom Minka, and Joaquin Quionero Candela.*

3. Europe Media Monitor (EMM) System  
Erik van der Goot, EC Joint Research Centre

Abstract: *The Europe Media Monitor (EMM) is the text gathering and analysis engine underlying a number of European media monitoring and other information analysis applications (e.g. EMM (<http://emm.newsbrief.eu>), MediSys (<http://medisys.newsbrief.eu>) ) that are serving EU policies especially those concerned with crisis management. The EMM engine consists of a growing number of text analysis and information processing modules currently performing the following tasks: language detection, known entity extraction, geo-tagging, sentiment analysis, categorization, duplicate detection, clustering, event detection and indexing. The system has a number of information aggregation modules to present the analysis results per category, per country, as a story etc. In the case of the EMM NewsBrief, the system harvests and analyses around 100.000 news articles per day in 40 languages from around 5500 RSS feeds and HTML pages, and categorizes these in approximately 1000 different categories defined by 35000 different keywords and keyword combinations. The system is developed and operated by the European Commission's Joint Research Centre (JRC). The presentation will focus on the history, development and architecture of the system.*

4. Invited Talk: Visual Pattern Recognition  
Giovanni Maria Farinella, University of Catania

*Abstract: Computer vision researchers are increasingly using algorithms from pattern recognition and machine learning to help build robust and reusable vision systems that act taking into account the visual content of images and videos. Just as learning is an essential component of biological visual systems, the design of machine vision systems that learn and adapt represent an important challenge in modern computer vision research. In this seminar I will focus on the task of recognition in computer vision by explaining how, why and where the techniques of pattern recognition and machine learning are used as essential ingredients to represent and recognize visual patterns in different application contexts.*

*September 2010*

The Editorial Team:

Tom Diethe  
University College London  
[t.diethe@cs.ucl.ac.uk](mailto:t.diethe@cs.ucl.ac.uk)

Nello Cristianini  
University of Bristol  
[Nello.Cristianini@bristol.ac.uk](mailto:Nello.Cristianini@bristol.ac.uk)

John Shawe-Taylor  
University College London  
[J.Shawe-Taylor@cs.ucl.ac.uk](mailto:J.Shawe-Taylor@cs.ucl.ac.uk)