

AVVISO DI SEMINARI

Dr. Eric Paquet (Ph.D)

National Research Council, Ottawa, Canada

Giovedì 22 luglio 2010 ore 11.00, Aula Gödel, II piano, Dipartimento di Informatica
Multimedia Data Mining for Images and Three-Dimensional Objects:
Applications to Cultural Heritage and Structural Proteomics

This talk discusses our research regarding content-based retrieval of 2D images and 3D objects and details the application of our work in the cultural heritage and proteomics domains. We also present a theoretical perspective on index spaces and the concept of distance.

Part I: Cultural Heritage

Digital cultural heritage data covers a wide range of techniques from acquisition, modelling, visualisation, navigation, archiving, indexing and, restoration, just to mention a few. There is an urgent need for better integration of, and synergy between, these components, both from an efficiency and utility point of view. An approach based on content-based indexing and retrieval of 2-D images and 3-D objects is proposed which creates a thread, and provides a common methodology to better integrate and manage these otherwise disconnected processes. To this end, we describe our content-based indexing and retrieval system and show its applicability within the cultural heritage domain.

Part II: Structural Proteomics

Research suggests that the complex geometric shapes of amino-acid sequence folds often determine their functions. In order to aid domain experts to classify new protein structures, and to be able to identify the functions of such new discoveries, accurate shape-related algorithms for locating similar protein structures are thus needed. We present our Content-based Analysis of Protein Structure for Retrieval and Indexing system, which locates protein families, and identifies similarities between families, based on the 2D and 3D signatures of protein structures. Our approach is novel in that we utilize five different representations, using a query by prototype approach.

Part III: Index Spaces: Theoretical Considerations

Distance is a fundamental concept when considering the information retrieval and cluster analysis of 3D information. That is, a large number of information retrieval descriptor comparison and cluster analysis algorithms are built around the very concept of the distance, such as the Mahalanobis or Manhattan distances, between points. Although not always explicitly stated, a significant proportion of these distances are, by nature, Euclidian. This implies that it is assumed that the data distribution, from a geometrical point of view, may be associated with a Euclidian flat space. In this talk, we draw attention to the fact that this association is, in many situations, not appropriate. Rather, the data should often be characterised by a Riemannian curved space.



Eric Paquet is a Senior Research officer at the Visual Information Technology Group of the National Research Council of Canada (NRC) and Adjunct Professor at the School of Information Technology and Engineering of the University of Ottawa. He received his Ph.D. in Computer Vision from Laval University (Quebec City, Canada) and the National Research Council in 1994. After finishing his Ph.D., he worked on optical information processing at the University of Valencia (Spain), on laser microscopy at the Technion-Israel Institute of Technology and on 3D hand held scanners at Tricorder plc (London, England). He is currently pursuing research on content-based indexing and retrieval of 2D images and 3D objects, as well as on visualisation, at the NRC. His current research interests include content-based description of multimedia and multi-dimensional objects, anthropometric databases, proteomics and cultural heritage applications.