Laboratory of Searching and Dialogue

/ Department of Computer Systems and Communications
/ Faculty of Informatics / Masaryk University

RESEARCH GROUP CONTACT »»	Botanická 554/68a, 602 00 Brno, Czech Republic http://lsd.fi.muni.cz/tiki-index.php	HEAD Prof. Pavel Zezula PHONE + 420 549 497 992 E-MAIL zezula@fi.muni.cz

THEMATIC RESEARCH FOCUS ↘

RESEARCH AREA

» Digital data processing, large-scale data collections, parallel processing, performance

EXCELLENCE

» New methods for extensible and scalable similarity search in digital data

MISSION

We want to be a workplace of research at international level, educating excellent IT specialists and successfully cooperating with companies.

DEVELOPED TECHNOLOGIES ↘

CONTENT OF RESEARCH

Similarity has been a central notion throughout our lives – it is a main concept in human perception, speech or face recognition, object classification, memory, and many others. As almost everything we see, read, hear, write or measure can now be in digital form, the group is trying to develop computer tools to manage similarity. To this objective, we assume a very universal concept of similarity that is based on the mathematical notion of metric space. In this concept, data collection is seen as objects together with a method to measure similarity between pairs of objects.

Our software tools are based on the following three development pillars:

» Extensibility – they can be used practically on data of any type – we only have to define how to measure "similarity between pairs of objects". Examples of similarity concepts that can be managed include various aspects of image visual descriptors (shapes, colours, combination of shapes & colours), similarity of video sequences, face recognition, similarity of biometric data such as fingerprints or iris scan, etc.

» Scalability – they are efficient even for very large databases. Designs offer various search structures including distributed mechanisms that are suitable for fast processing of large datasets and are mostly based on peer-to-peer principles

» Infrastructure independence – the search technologies can run on various HW infrastructures including large-scale distributed computer clouds. Migrating the search engine to different hardware can be used for performance tuning (response time and query throughput)

MAIN CAPABILITIES

Basic research
We have developed several generic search methods and published papers in the top scientific conferences and most prestigious scientific journals. We have also build prototypes and tested their performance on very large collections of data. Basic software tools are publicly available.

Application research + protection forms
Our open-source software for non-text similarity searching has been applied by Pixmac and Profimedia photo selling sites. The fastest developing application area concerns security. In general, the application areas include: - Information retrieval

» Multimedia indexing and searching (images, audio, video, etc.)
» Classification and recommender systems
» Large-scale biometric identification problems
» Biological sequence processing, etc.

FIELDS OF RESEARCH RESULTS APPLICATION

» ICT (Information and Communication Technologies) - software, internet, IT security, ecommerce
» Security (Biometric identification).

ALUMNI PROFILE

The students gain a comprehensive overview in the field of acquisition and processing of image information starting with simple image modifications using point transforms or linear filters and ending up with sophisticated tools such as mathematical morphology of deformable models. Graduates are equipped with knowledge and skills that enable them to solve challenging research and development problems in computer science and lead research teams engaged in the development and subsequent implementation of new technologies in the field of modern computer science. Graduates can work in academic and also industrial sectors, which occupy higher positions that require nontrivial management, analytical and technical skills.

NUMBER OF RESEARCH POSITIONS ↘

SENIOR RESEARCH STAFF

5

JUNIOR RESEARCH POSITIONS (INCL. PH.D. STUDENTS)

8

BUDGET ↘

TOTAL (MIL. CZK/ MIL. EUR)

7.5 / 0.3

PART OF THE TOTAL BUDGET FROM PRIVATE RESOURCES (%)

10

PART OF THE TOTAL BUDGET FROM FOREIGN RESOURCES (%)

0

MAIN PROJECTS ↘

2009-2011: Searching in Large Multimedia Databases (project GA201/09/0683 financed by the Czech Science Foundation)

2006-2008: Search on Audio-visual content using Peer-to-peer Information Retrieval (project IST 045128 financed by the 6th Specific RTD Programme, European Union)

2004-2006: Network of Excellence on Digital Libraries (project IST 507618 financed by the 6th Specific RTD Programme, European Union)

ACHIEVEMENTS ↘

» Metric similarity search (M-tree, VLDB 1997), the most cited article about similarity searching;

» Springer US, 2006, Similarity Search: the Metric Space Approach - the first book on similarity searching;

» IBM Shared University Research Award 2008 on Web-scale Similarity Search in Multimedia Data

MAIN COLLABORATING PARTNERS ↘

COLLABORATION WITH ACADEMIC PARTNERS

» Faculty of Mathematics and Physics, Charles University in Prague (Prague, CZ)

» Max-Planck-Institut für Informatik (Saarbrücken, DE)

» Institute of Information Science and Technologies, Italian National Research Council (Pisa, IT)

» École polytechnique fédérale de Lausanne (Lausanne, CH)

COLLABORATION WITH COMPANIES

» IBM Haifa Research Lab (Haifa, IL)

» Picsearch AB (Stockholm, SE)

» Telenor ASA (Oslo, NO)

» Javlin (Prague, CZ)

» Bull.cz (Prague, CZ)

EXPECTATIONS ↘

REQUIREMENTS

From our potential partners, we mainly expect cooperation on definition of similarity search problems and specification of effectiveness of search - which is domain (application) specific. We are interested in partners with credible business models that would lead to the development of profitable products. We are also interested in research partners to define common projects and submit proposals for grants.

OFFERS

» Software and software support for similarity search execution, as needed for numerous contemporary digital data collections, such as multimedia, biometric, biological, chemical, statistical, or other scientific data

» Cooperation on common projects which need a similarity search

» Share of know-how through consultancy

centrope_tt
TECHNOLOGY TRANSFER COMMUNITY POWERED & CONNECTED BY CENTROPE

CENTRAL EUROPE
COOPERATING FOR SUCCESS.

EUROPEAN UNION
EUROPEAN REGIONAL DEVELOPMENT FUND

B | R | N | O

South Moravian Region

RIS>3

JIC SOUTH MORAVIAN INNOVATION CENTRE