Status Report of the Project "Advanced Querying and Coordination of Multimedia Information in ETH World"

Can Türker
Institute of Information Systems
ETH Zürich, Switzerland
tuerker@inf.ethz.ch

3. November 2000

Project Members

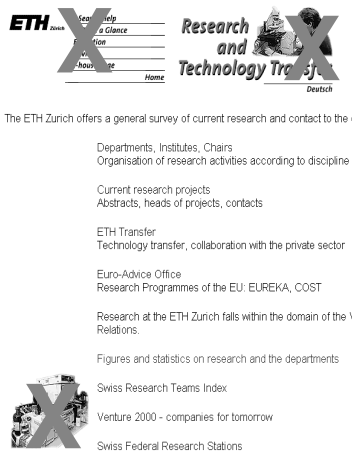
- **Prof. Dr. H.-J. Schek**
- **Dr. Klemens Böhm**
- **Dr. Can Türker**
- **Klaus Haller**
- **Michael Mlivoncic**
- **Heiko Schuldt**
- **Christoph Schuler**
- **Roger Weber**

Problem and Project Focus

- **Todays Information Space**
  - **Huge, heterogeneous, multimedia, rapidly increasing**
  - **Problem: How to find relevant information?**
  - **Advanced tool assistance needed!**
- **Querying & Coordinating Information Sources**
  - **Combined data, text, and image retrieval**
    - Context-aware and feedback-aware
  - **Providing consistent information sources**
    - Increasing retrieval quality

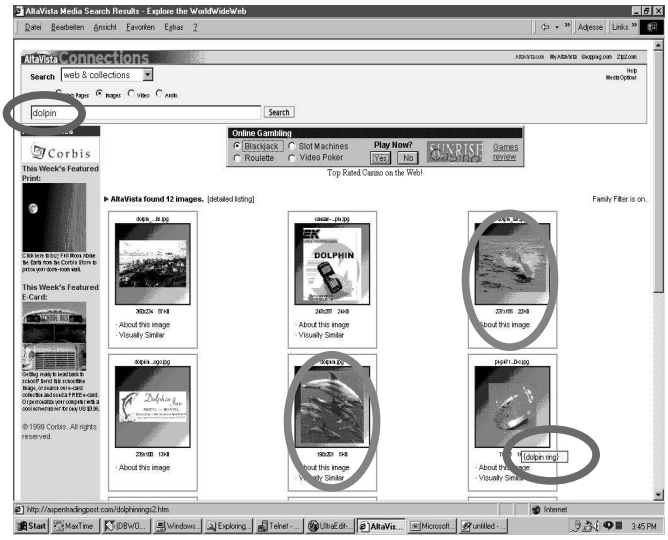
Searching for Information

- **Indexing**

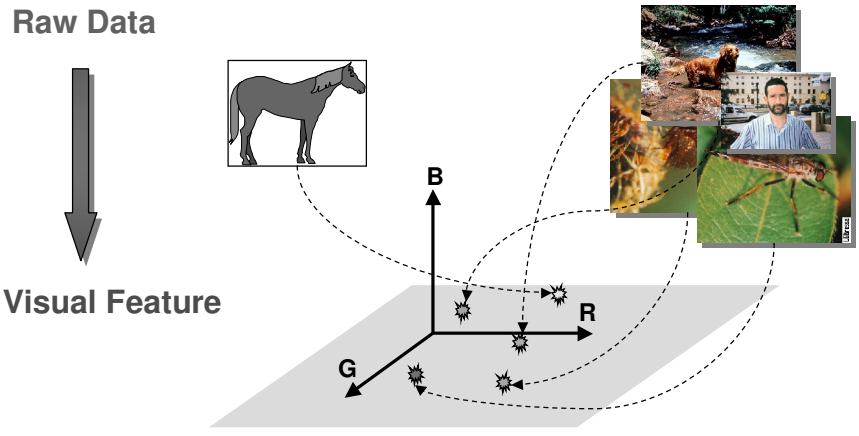


The ETH Zurich offers a general survey of current research and contact to the departments. Departments, Institutes, Chairs Organisation of research activities according to discipline Current research projects Abstracts, heads of projects, contacts ETH Transfer Technology transfer, collaboration with the private sector Euro-Advice Office Research Programmes of the EU: EUREKA, COST Research at the ETH Zurich falls within the domain of the Vice-president for Research who is assisted by Staff Research and Private Sector Relations. Figures and statistics on research and the departments Swiss Research Teams Index Venture 2000 - companies for tomorrow Swiss Federal Research Stations

Image Search with Keywords



Feature-based Image Search



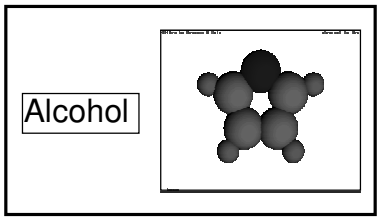
Similarity Search with Relevance Feedback

- **It is hard to formulate an image query!**

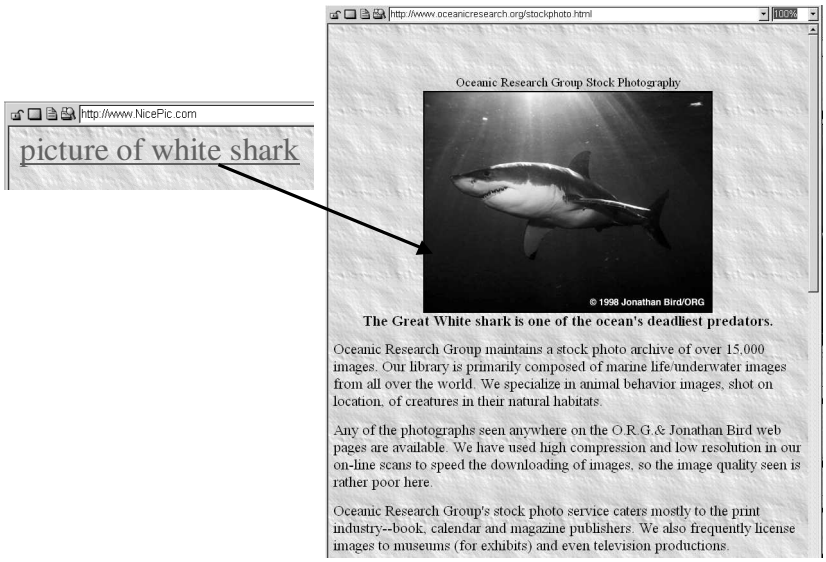
Combining Text and Image Search

- **Example: Searching for alcohol molecules**

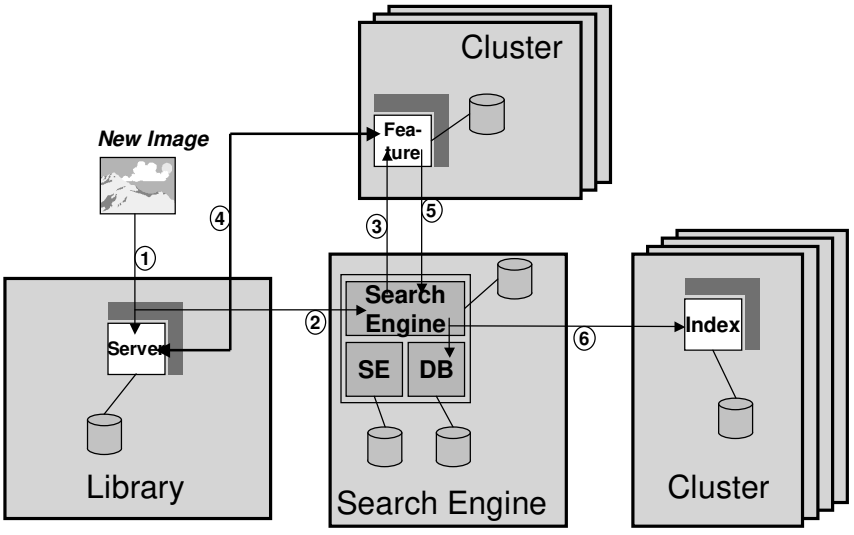
Query=



Extraction of Text Information



Coordination of Information Sources



Finished and Ongoing Work

- **Image search engine with similarity search**
- **Integration of a query planner**
  - **Dynamic selection of best execution plan based on user-specified conditions**
- **Combining Text and Image Retrieval**
  - **Considering structured predicates**
  - **Extraction of image annotations**
- **Coupling Application Systems, e.g.,**
  - **ImageFinder (image database of the ETH Library)**
  - **Monitoring and triggering relevant system events**

Future Steps

- **Improving existing search mechanisms**
  - **Interface**
  - **Relevance feedback**
  - **Feature selection**
  - **Query plan selection**
  - **Context-awareness**
- **Coordination agents for other important ETH information sources**
  - **Publish&Subscribe**

Talks

- **H.-J. Schek:** Hyperdatabases, ***Int. Conf. on Web Information Systems Engineering, Hong Kong, June 2000***
- **R. Weber:** Parallel Search, ***Int. Conf. on Data Engineering, April 2000***
- **K. Böhm:** Approximative Search, ***Int. Conf. on Extending Database Technology", Konstanz, April 2000***
- **H.-J. Schek, R. Weber:** Hyperdatabases & Similarity Search, ***Microsoft Research, Cambridge, May 2000***
- **R. Weber:** Similarity Search, ***Kolloquium, University of Halle, June 2000***
- **R. Weber:** Complex Queries for ETH World, ***Internal Workshop of the Database Group, July 2000***
- **R. Weber:** Verwalten von und Suche nach Informationen in ETH World, ***Kolloquium, Institute of Information Systems, October 2000***

Talks and Other Activities

- **H.-J. Schek:** Visuell- und kontextbasierte Informationssuche in ETH World, ***Forum für Informatik-AnwenderInnen, ID-Forum, 24. November 2000***
- **K. Böhm, H.-J. Schek:** ***Organisation of the DELOS Network of Excellence Workshop on*** Information Seeking, Searching and Querying in Digital Libraries, ***December 2000***

Publications

- **R. Weber, K. Böhm:** Trading Quality for Time with Nearest-Neighbor Search. ***In: Proc. 7th Conf. on Extending Database Technology (EDBT 2000), Konstanz, 2000.***
- **K. Böhm, T. Grabs, U. Röhm, H.-J. Schek:** Evaluating the Coordination Overhead of Synchronous Replica Maintenance in a Cluster of Databases. ***In: Euro-Par 2000 Parallel Processing, Proc. 6th Int Euro-Par Conference, Munich, 2000.***
- **H.-J. Schek, R. Weber:** Higher Order Databases and Multimedia Information. ***In: Proc. of the Swiss/Japan Seminar Advances in Database and Multimedia, Kyoto, 2000.***
- **H.-J. Schek, K. Böhm, T. Grabs, U. Röhm, H. Schuldt, and R. Weber:** Hyperdatabases. ***In: Proc. 1st Int. Conf. on Web Information Systems Engineering (WISE'00), Hong Kong, 2000.***
- **R. Weber, K. Böhm, and H.-J. Schek:** Interactive-Time Similarity Search for Large Image Collections Using Parallel VA-Files. ***In Proc. 4th European Conf. on Research and Advanced Technology for Digital Libraries (ECDL2000), San Diego, 2000.***

Publications (Cont.)

- **A. Lazcano, G. Alonso, H. Schuldt, and C. Schuler:** The Wise Approach to Electronic Commerce. ***In: Int. Journal of Computer Systems Science & Engineering. Special Issue on Flexible Workflow Technology Driving the Networked Economy, 2000.***
- **C. Schuler, H. Schuldt, and H.-J. Schek:** Transactional Execution Guarantees for Data-Intensive Processes in Medical Information Systems. ***To appear in: Proc. 1st European Workshop on Computer-based Support for Clinical Guidelines and Protocols. Leipzig, 2000.***
- **R. Weber, M. Mlivoncic, K. Böhm:** Quality-Aware and Query Planning Local Sensitive Image Search. ***To appear in: Proc. 17th Int. Conf. on Data Engineering (ICDE2001), 2001.***
- **T. Grabs, K. Böhm, H.-J. Schek:** High-level Parallelisation in a Database Cluster: a Feasibility Study Using Document Services. ***To appear in: Proc. 17th Int. Conf. on Data Engineering (ICDE2001), 2001.***

Dissertations

- **Roger Weber:** Similarity Search in High-Dimensional Vector Spaces**, December 2000**
- **Heiko Schult:** Transactional Process Management**, December 2000**

Finished Diploma Theses and Semester Work

- **J. Bosshard: Relevanz Feedback for Content-based Image Retrieval**
- **R. Grob: Query Planning Component for Similarity Search**
- **K. Stöckli: Cost-based Execution of Transactional Processes**
- **A. Naef: Monitoring Module for Nested Process Management Systems**
- **R. Chisholm: Implementing/Evaluating the Vector-Space Model**
- **L. Calgari/H. Gröller: Benchmarks for Similarity Search**
- **Y. Langisch: Efficient Web-Spider**
- **R. Locher: Graphical Conflict Specification in Transactional Workflow Systems**
- **C. Stolz: Communication Server for Transactional Process Management**
- **N. Atanasoski: Transactional Data Transfer in Medical Information Systems**
- **Internship:**
  - **R. Kumar: Providing XML Document Collections**
  - **P. Manadhata: Testing a Text Retrieval System for Databases**

Ongoing Diploma Theses and Semester Work

- **E. Burgherr: Combined Text and Image Retrieval**
- **H.P. Purtschert: Publish&Subscribe Techniques and their Integration into ETH World**
- **K. Jauslin: Web-Portal for ETH World**

- **K. Diouf: Coupling Application Systems to the Publish&Subscribe System of ETH World**
- **P. Miotti: Agents for Monitoring Directory Structures**
- **P. Sieber: Extraction of Image Annotations from HTML-Pages**

- **Hiwis:**
  - **P. Wallmann: Latent Semantic Indexing for Text and Image Retrieval**
  - **R. Widmer: Extended Search Algorithms for Similarity Search**

Homepage

- **http://www-dbs.inf.ethz.ch/projects/ethworld**
- **Online demo of the search engine**
- **More detailed information will be available soon**