

Data Science Research

Özgür Ulusoy

<http://www.cs.bilkent.edu.tr/~oulusoy/>

- **Text-to-SQL Systems**
- **Web Search Engines**
- **Multimedia Database Systems**
- **Social Network Analysis**



Text-to-SQL Systems

- **Keyword Mapping**
- **Explainable Text-to-SQL**
- **Query Recommendation**

Text-to-SQL Systems

Keyword Mapping - mapping between tokens in the query and relational database elements

- ❖ Keyword mapping formulated as sequence tagging problem in NLP
- ❖ An end-to-end keyword mapper (DBTagger)

Text-to-SQL Systems

Explainable Text-to-SQL - based on Explainable AI

- ❖ Explaining the decisions made by the keyword mapping system using LIME
- ❖ A novel wrapper around LIME
- ❖ Using a schema graph to explain join-path inference

Text-to-SQL Systems

Query Recommendation

- ❖ A ranking method to create a list of suggested queries
- ❖ Distributed representations for database tuples, trained on a relational database
- ❖ Graph convolutional networks (GCNs) to learn distributed representations



Web Search Engines

- **Learning to Rank for Educational Web Search**
- **Diversification of Search Results**
- **Efficiency and Scalability Issues for Web-IR**

<http://www.cs.bilkent.edu.tr/~bilweb>

Web Search Engines

Learning to Rank for Educational Web Search - machine-learned ranking models

- ❖ A rich set of features employed in educational search
- ❖ Domain knowledge utilized to build query-dependent ranking models



Web Search Engines

Diversification of Search Results

- ❖ **Multidimensional result diversification**
- ❖ **Supervised learning methods for search result diversification**
- ❖ **Impact of index pruning on diversification performance**



Web Search Engines

Efficiency and Scalability Issues

- ❖ **Methods and algorithms for caching, indexing, and query processing in search engines**



Multimedia Databases

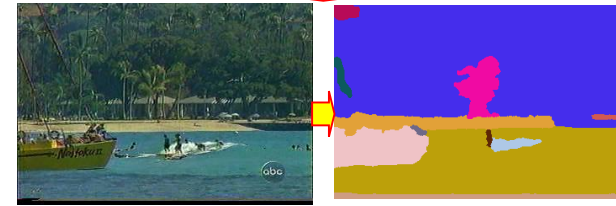
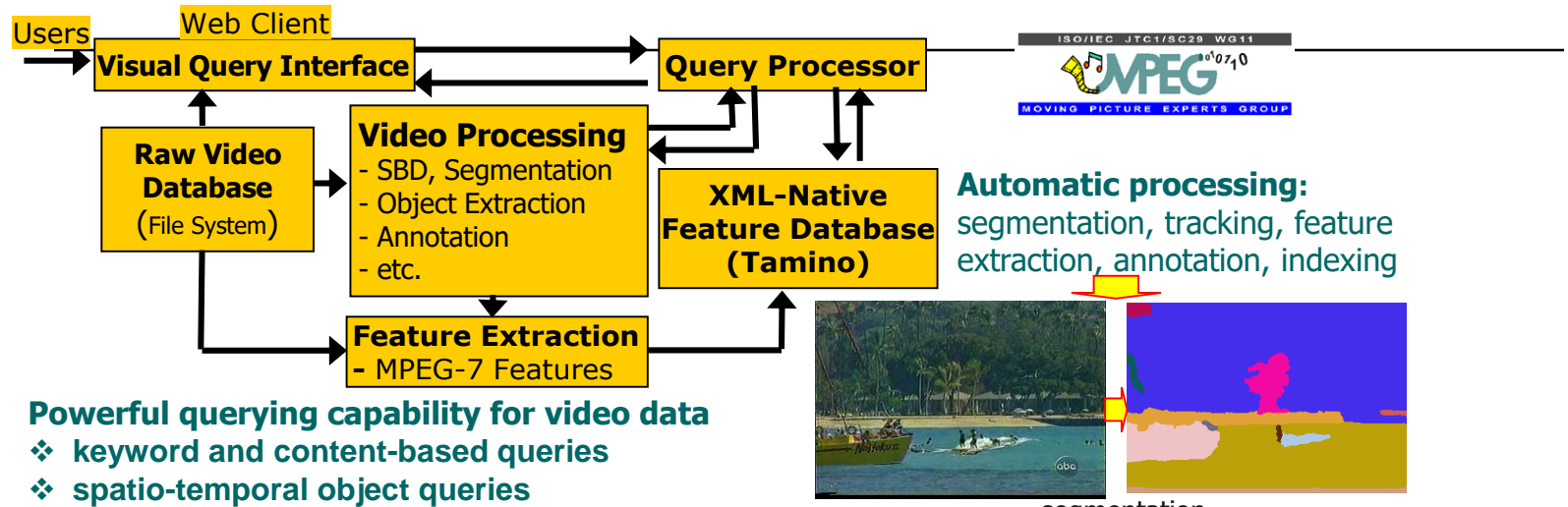
(joint work with Prof. Uğur Gündükbay)

- **Video Retrieval Systems**
- **Mobile Visual Search**
- **Learning Visual Similarity for Image Retrieval**

<http://www.cs.bilkent.edu.tr/~bilmdg>

Multimedia Databases

BilVideo-7: An MPEG-7 Compatible Video Retrieval System



segmentation

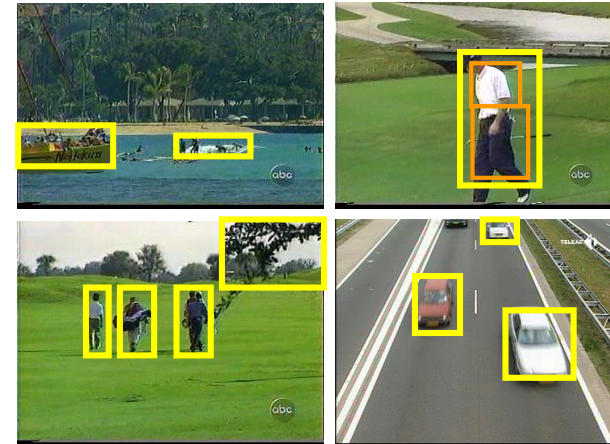
Powerful querying capability for video data

- ❖ keyword and content-based queries
- ❖ spatio-temporal object queries



keywords: trees, greenery, sky – bush, putin, dog

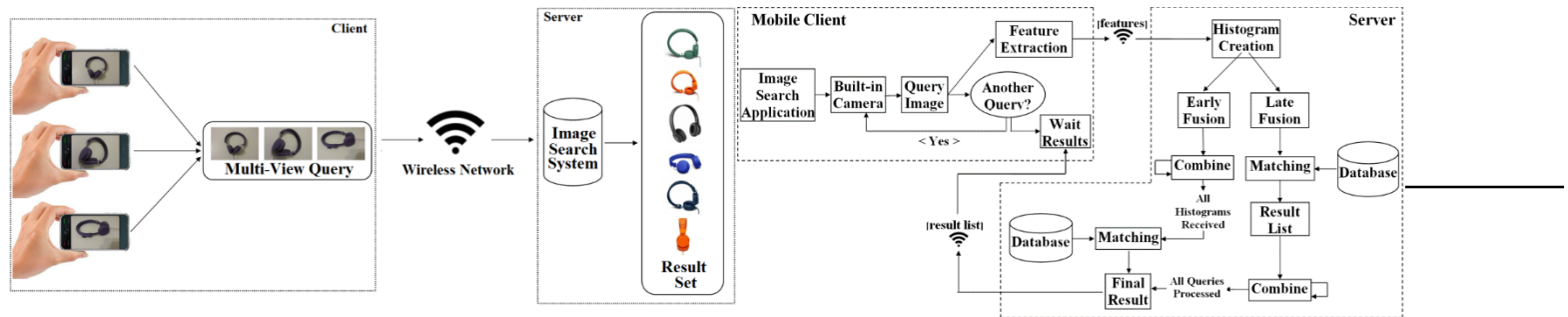
Example query formulation



Salient video object extraction

Multimedia Databases

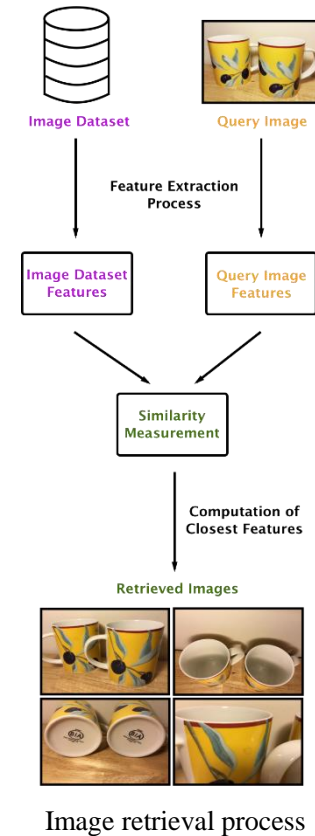
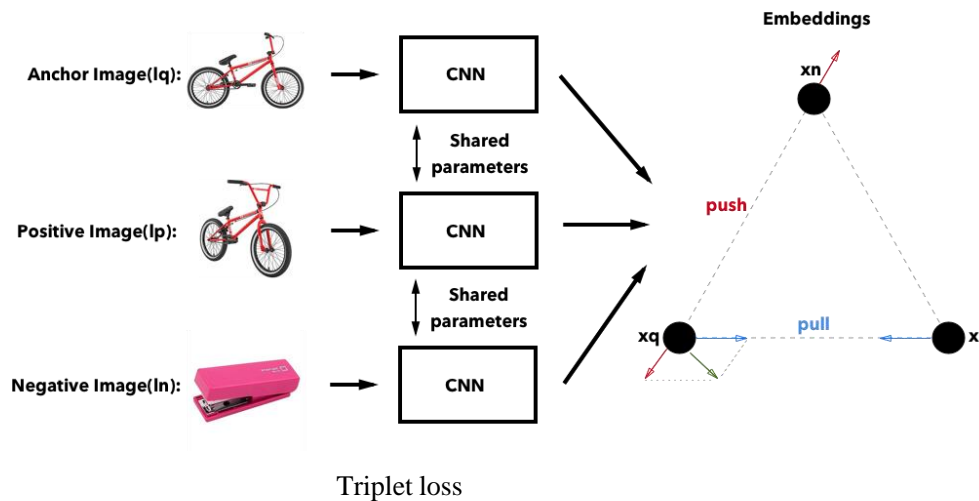
Mobile Image Search Using Multi-View Object Image Queries



Multimedia Databases

Learning Visual Similarity for Image Retrieval with CNNs

- ❖ Learning efficient visual similarity for image retrieval by revealing resemblances and differences between product images





Social Network Analysis

- **Misinformation Propagation in Social Networks**
- **Social Network Data Analysis on Big Data Processing Platforms**



Social Network Analysis

Misinformation Propagation in Social Networks

- ❖ Misinformation problem modelled as a game based on the notion of cooperative games
- ❖ Agents defined to maximize the total reward
- ❖ A blockchain - machine learning hybrid approach for addressing misinformation in a crowdsourced environment



Social Network Analysis

Social Network Data Analysis on Big Data Processing Platforms

(joint work with Prof. İbrahim Körpeoğlu)

- ❖ **Community detection formulated as a multi-k-core problem**
- ❖ **Distributed multi-k-core construction and maintenance algorithms running on a big data platform**