

I-SEARCH

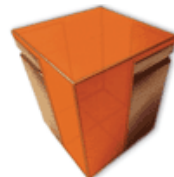
A Unified Framework for Multimodal Content Search

Dimitrios Tzovaras

CERTH-ITI

I-SEARCH Project Coordinator

Thomas Steiner
Google, Germany





What is I-SEARCH?

The mission

- To provide a novel unified framework for multimodal content indexing, sharing, search and retrieval
- To handle specific types of multimedia and multimodal content (text, 2D image, video, 3D objects and audio) along with real world and emotional information
- To develop a user-centric search-engine (only the content of interest will be delivered to the end-users)
- To optimally present the retrieved results to the user via novel visualization schemes
- To dynamically adapt the search engine to end-user's device (varying from a simple mobile phone to a high-performance PC)





What is STREP I-Search?

Objectives

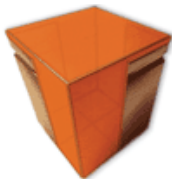
- Research and development of an innovative Rich Unified Content Description (RUCoD)
- Development of intelligent content interaction mechanisms so that only the content of interest will be delivered to the users
- Provide a novel way for the presentation of the multimodal data retrieved by the search engine, by utilizing visual analytics technologies





What is STREP I-Search?

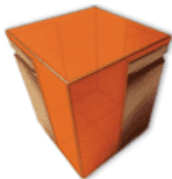
The aim of the I-SEARCH project is the **development of a search engine** able to handle **multimedia** and **multimodal content** (text, 2D image, video, 3D objects, audio and combination of the above), **which can be used as queries** and retrieve any available relevant content of any of the aforementioned types



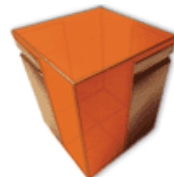
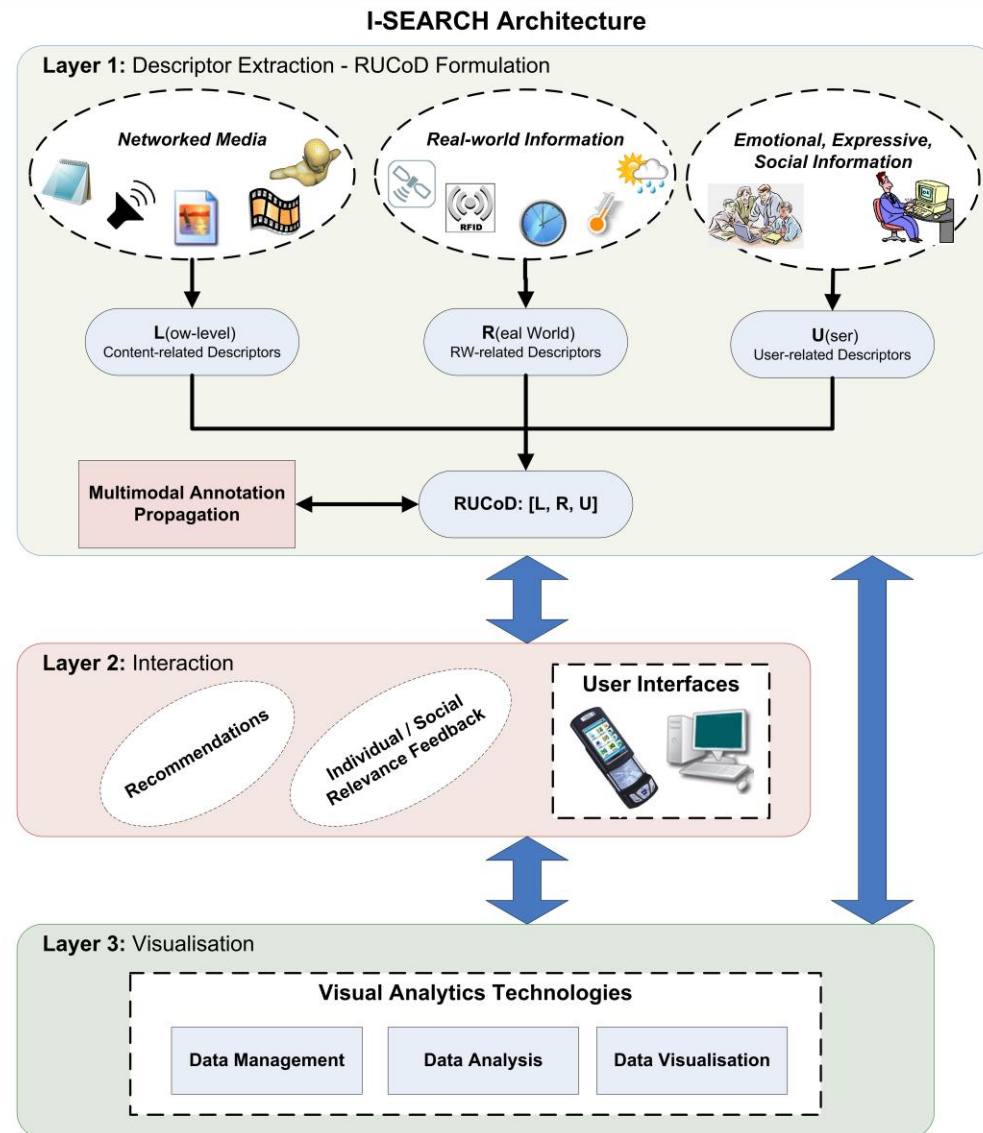


I-Search Use Cases

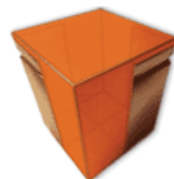
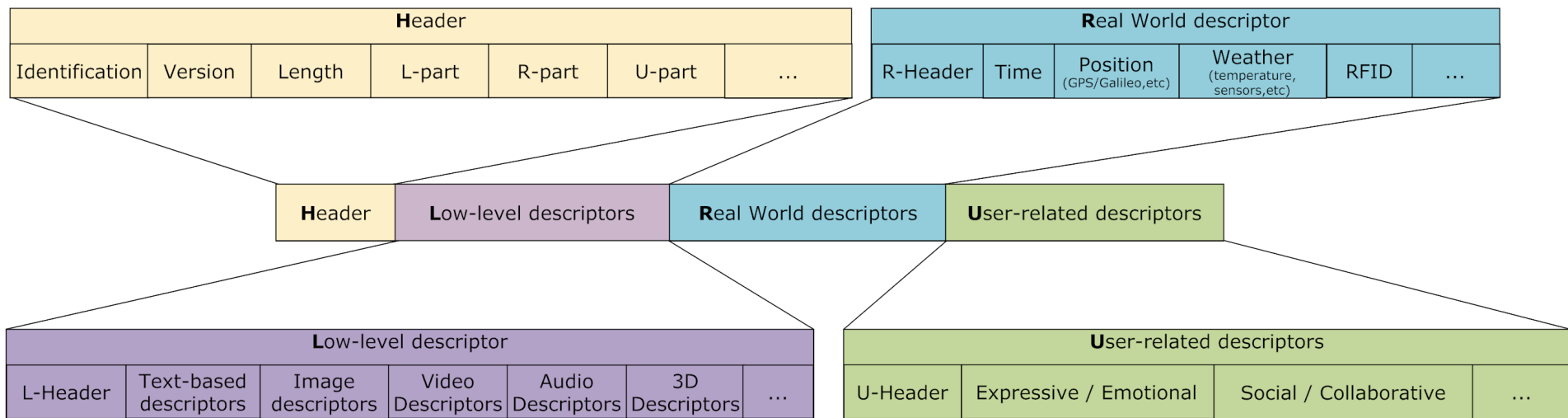
- UC1 - Music retrieval through expressive embodied queries
- UC2 - Social music retrieval through expressive embodied queries
- UC3 - Search and retrieval of a piece of furniture
- UC4 - Search and retrieval of multimedia content using the end-user's smartphone
- UC5 - Personalized approach to customer needs
- UC6 - 3D game component retrieval
- UC7 - Game avatar retrieval



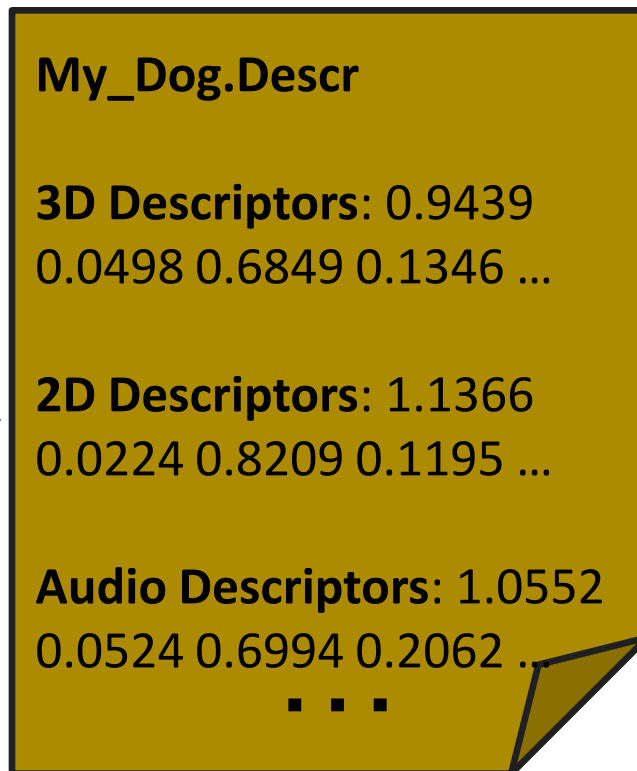
I-Search Architecture



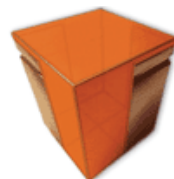
RuCoD Specifications



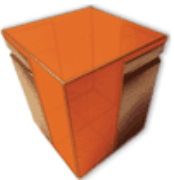
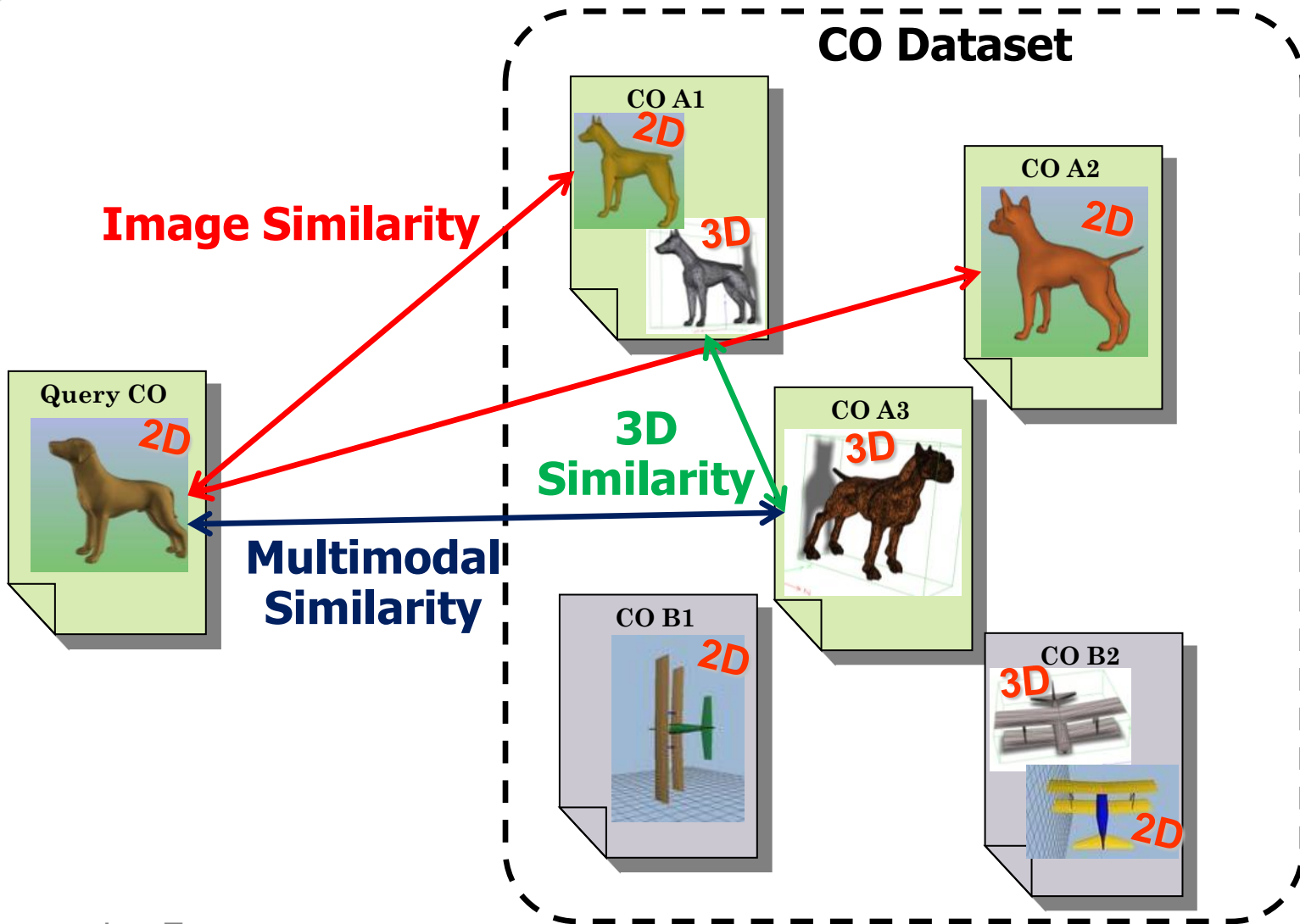
The Concept of Content Object (CO)



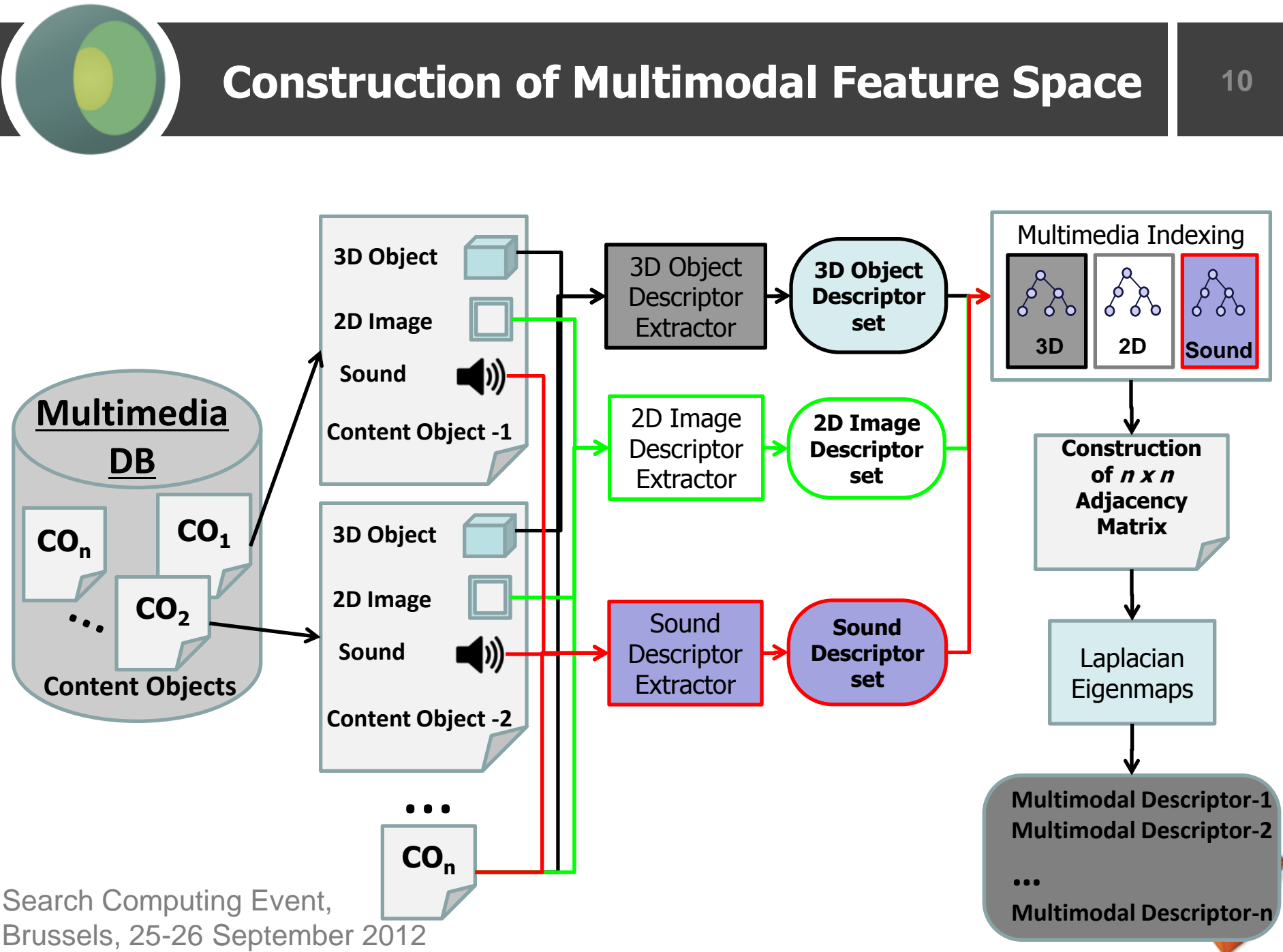
Content Object: a container of multimedia objects (Images, 3D objects, audio) that carry the same semantics information



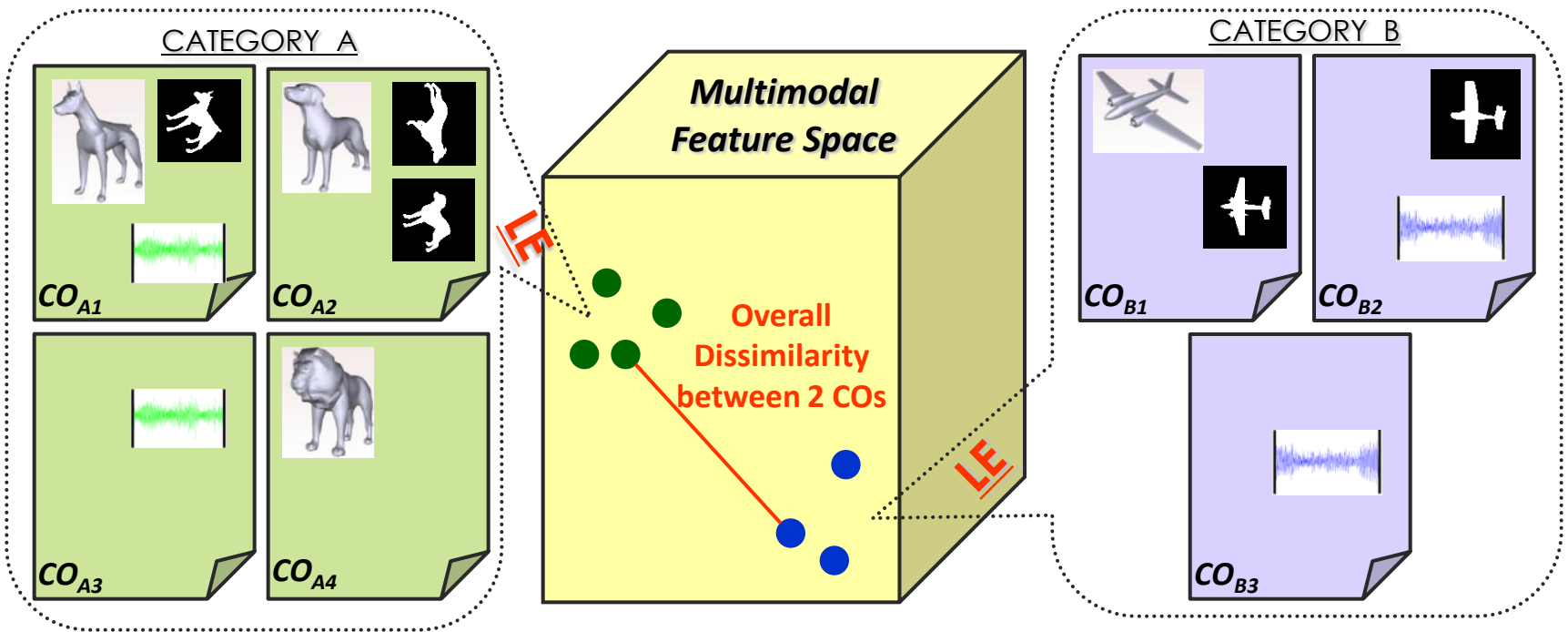
Multimodal Retrieval Example



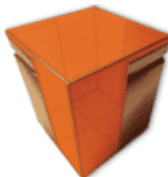
Construction of Multimodal Feature Space



Multimodal Feature Space

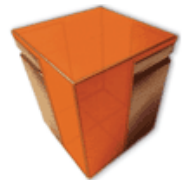
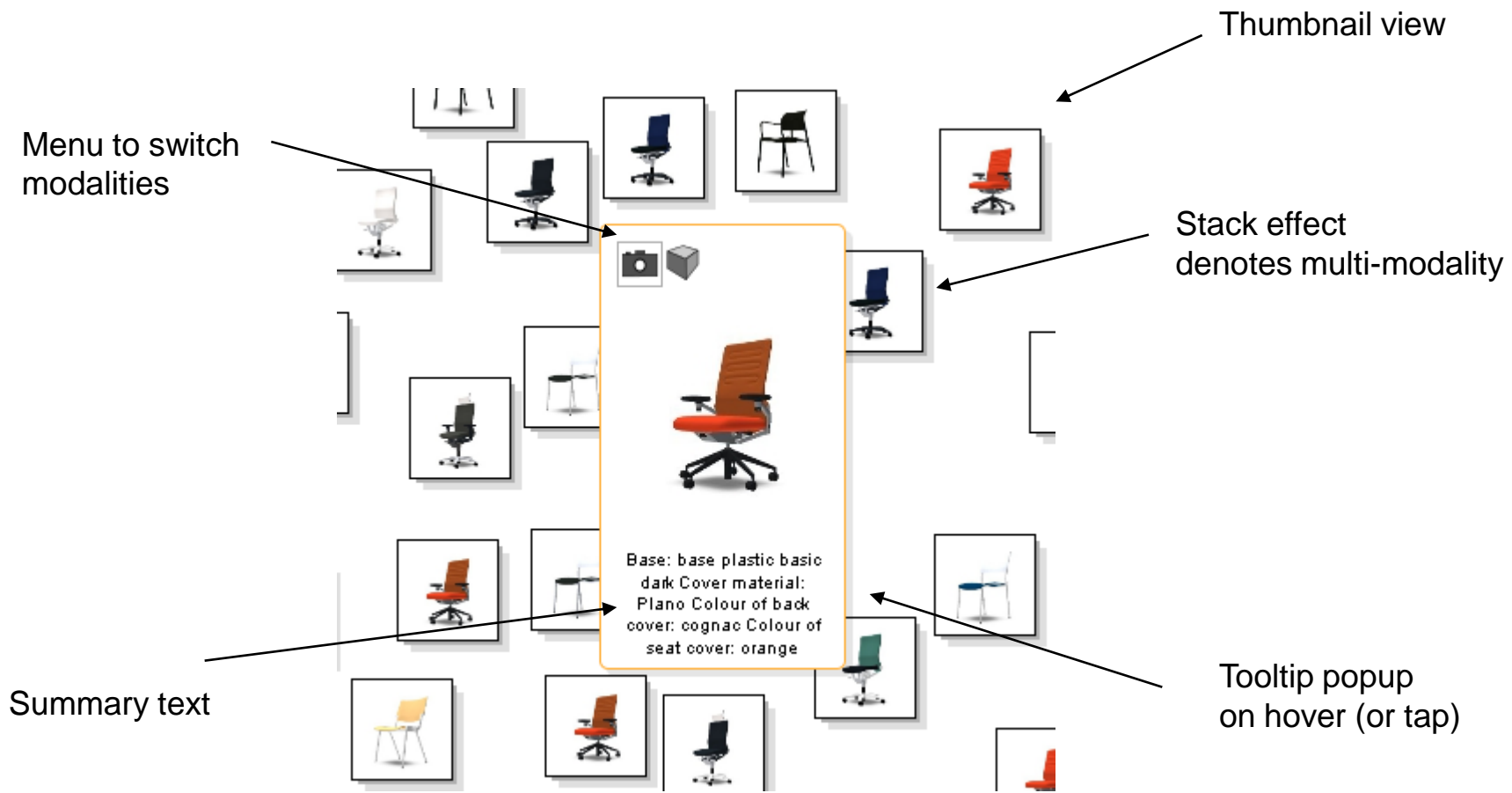


LE: Laplacian Eigenmaps





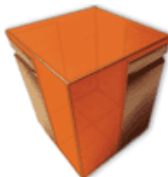
Results Presentation





Results Presentation

- Images: thumbnails
- 3D: rendered image thumbnails (look-around effect)
- Video: slideshow of keyframes
- Audio: generation of synthetic visual previews from aural features.



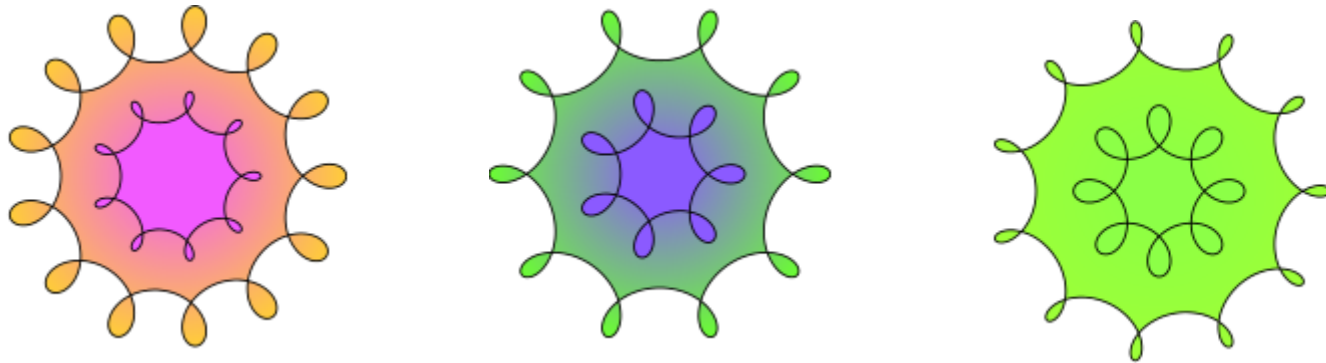


Audio Visualisation

MFCC spectral features

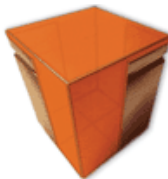
↳ Dimensionality Reduction

↳ Parametric shape generation

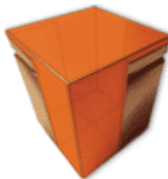
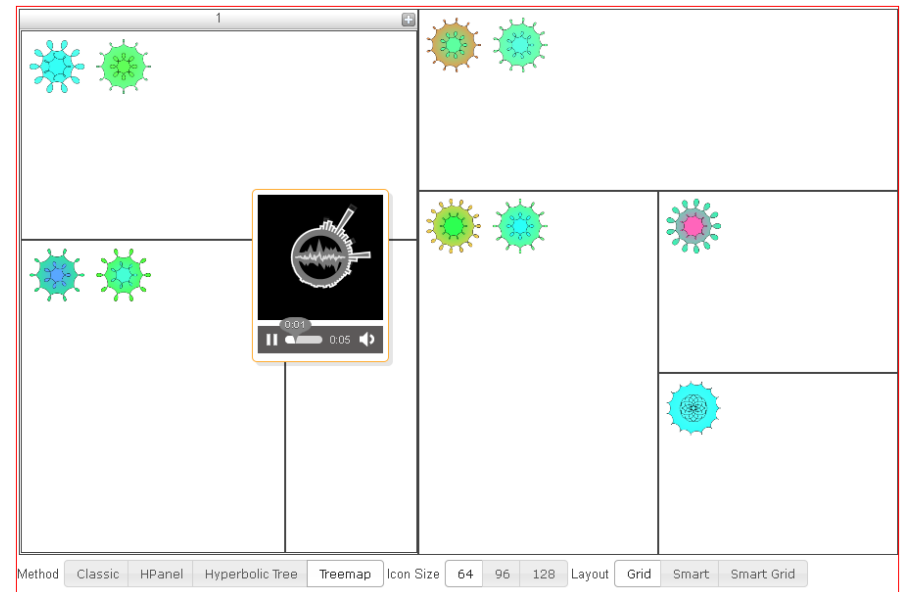
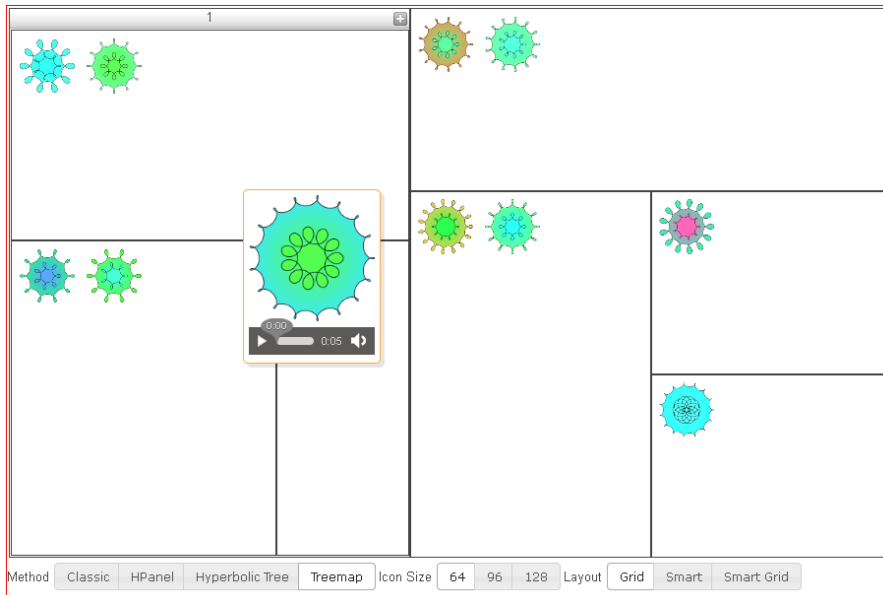


Parameters:

shape and number of internal/external petals +
internal/external colors



Audio Visualisation



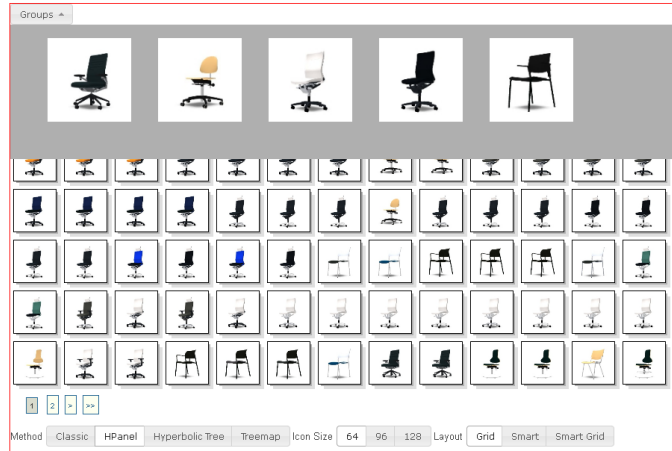


Cluster Visualisation

Classic

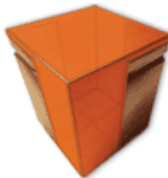
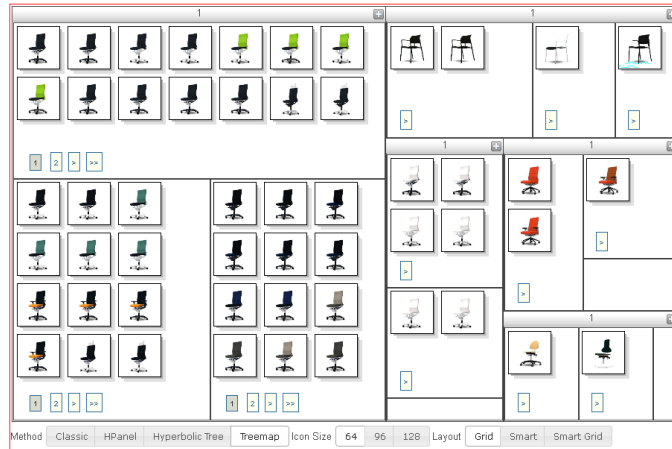
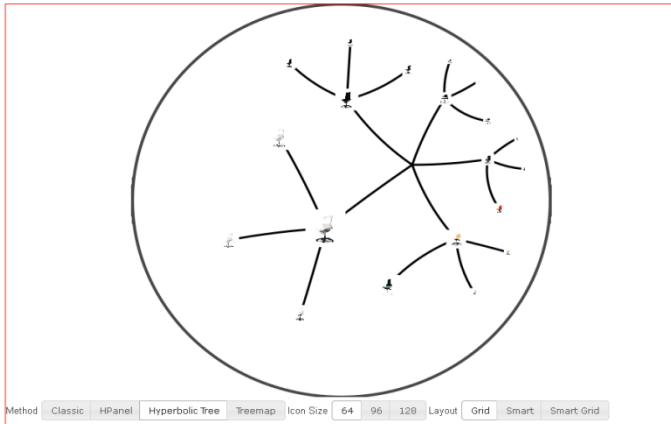


Panel



Treemap

Hyperbolic Tree



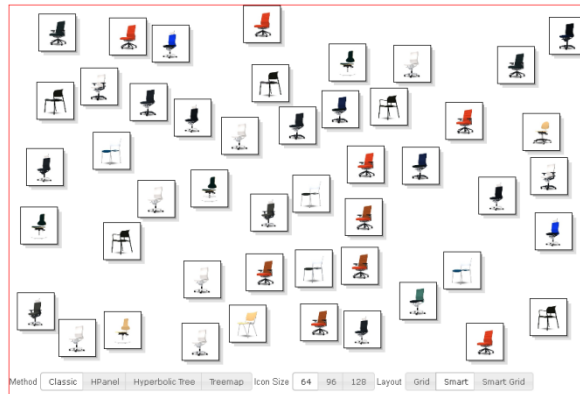


Thumbnail arrangement



Grid

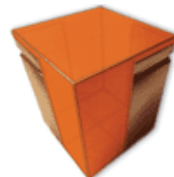
Multiple pages



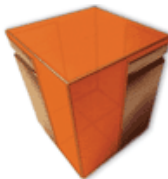
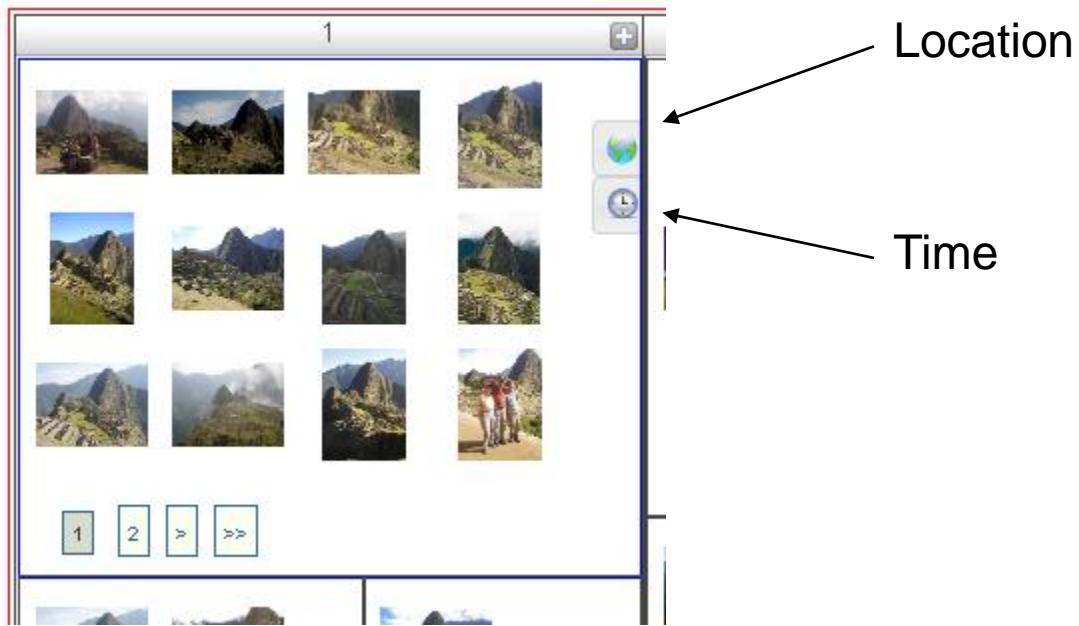
Smart



Smart Grid

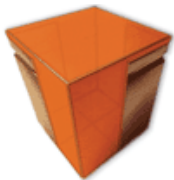


Real-world info



Location

The screenshot shows a web application interface for mapping document locations. The main window is titled "Geographic location of documents" and contains a Google Map of Europe. A popup window is open over the United Kingdom, displaying a photograph of a Lincoln Mark LT pickup truck. The popup text reads: "The Mark LT is a luxury pickup truck that was manufactured by the Lincoln division of American Location: A275, Uckfield, East Sussex TN22, UK". The map includes various city labels and a "Map" dropdown menu. On the left side, there is a "Media:" section with a vertical stack of image thumbnails. On the right side, there is a "Start from scratch" button and another vertical stack of image thumbnails. The bottom of the map area shows copyright information: "Map data ©2012 GIS Innovatsia, GeoBasis-DE/BKG (©2009), Google, Tele Atlas, and DeLorme".





Time

Timeline of documents

Media: a

Start from scratch

Timeline © SIMILE

2006 2007 2009

Apr May Jun Aug Sep Oct

The Fender Precision Bass (often shortened to "P Bass") is an Electric bass. Designed by Leo Fender and brought to Location: 12 Fitzherbert Spur, Portsmouth PO6, UK



Adaptation

Filtering (modality, tags, date, time, relevance)

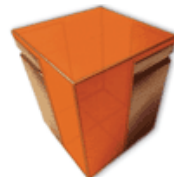
The screenshot shows a search interface with a search bar containing 'red car'. Below the search bar, there are controls for 'Media' (all, image, video, audio, document), 'Sort by' (list, clock, globe), and 'Tags'. A 'Query settings' panel is open on the right, showing 'Number of results' set to 100, 'Number of clusters' set to 5, and 'Transformation method' set to 'Minimum Spanning Tr'. The main area displays a grid of image thumbnails, including a white sports car, a yellow car, a blue car, a green car, a red car, a blue car, a yellow car, a blue car, a red motorcycle, a deer, and a guitar. At the bottom, there are navigation controls including 'Method' (grid, list, zoom, refresh), 'Icon Size' (small, medium, large), 'Layout' (grid, list, zoom), and 'Navigation Mode' (Browse, Feedback). A 'HistorySave' button and copyright notice '© 2011 I-SEARCH Project - All rights reserved' are also visible.

Query settings

Clustering

Thumb size

Thumb layout





Feedback

Relevance

User tagging

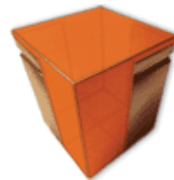
Add/Edit Tags

- Bird x
- Swallow x
- Hirundinidae x
- Aves x
- Birds x
- Oklahoma x
- Tillmancounty x
- Hackberryflatwma x
- Documentation x
- Seagull x
- Forest x
- Cuckoo x
- Bees x
- wire



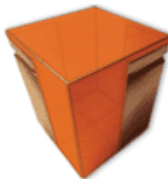
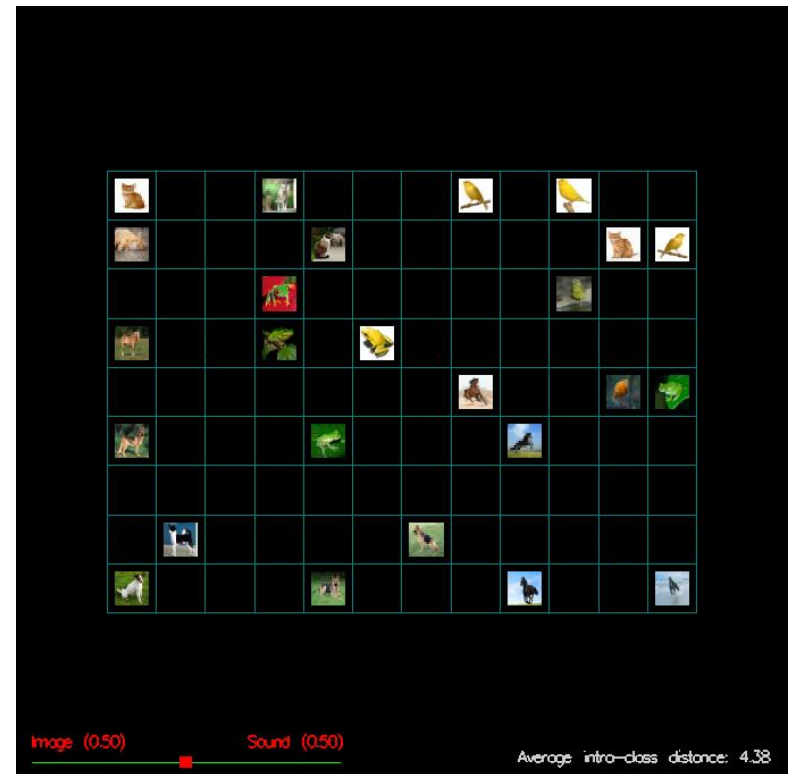
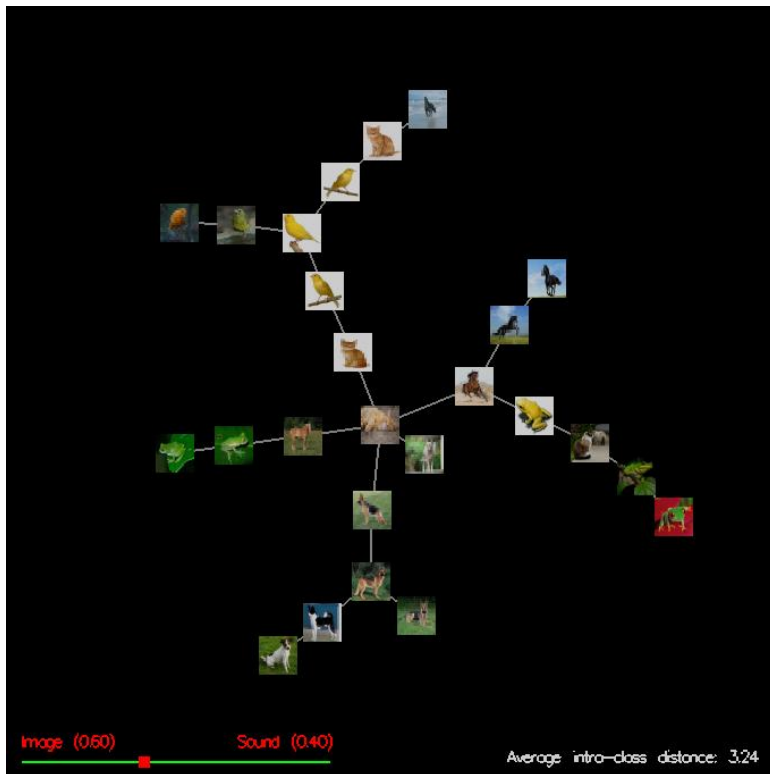
Method Icon Size Layout

Navigation Mode **Browse** **Feedback**



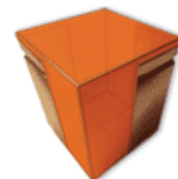
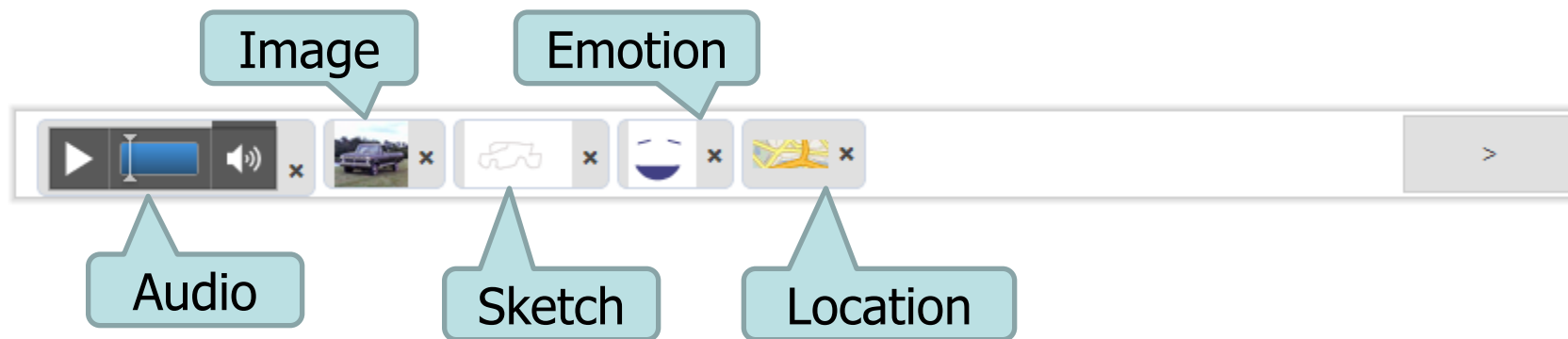
Advanced interactive feedback

Exploit user feedback to provide more informative visualisations



MuSeBag - Multimodal Search Interface

- To support : Audio, Video, 3D, Sketch , Image, Real world data, texts
- *Tokens for different modalities representation*
- *HTML5 based*
- *Adaptiveness and Device independence*





I-Search Demonstrators

Demonstrators

- I-SEARCH User Interface

<http://isearch.ai.fh-erfurt.de>

- CoFetch tool for semi-automatic creation of Content Objects (HSF)

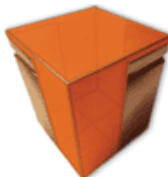
<http://isearch.ai.fh-erfurt.de/cofetch/>

- Multimodal Search (CERTH)

<http://vcl.itι.gr/is/isearch/client/musebag/>

- Demo presentation of UC3 - Furniture retrieval (CERTH)

<http://195.251.117.110/isearchgui2/client/musebag/>





I-Search: Overview

Thank you for your attention!

