



I-SEARCH

Newsletter

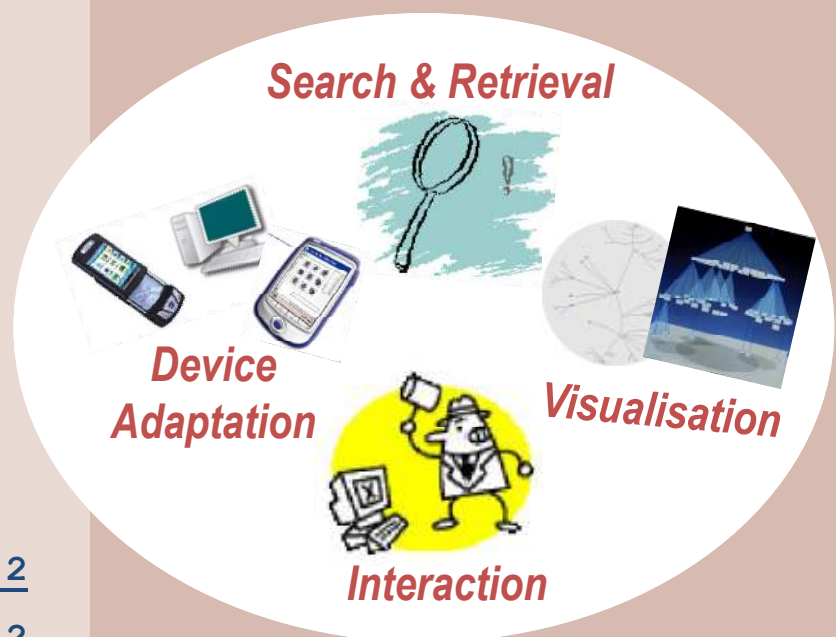
Issue 3

www.isearch-project.eu

June, 2011

CONTENTS :

<u>Welcome Note</u>	2
<u>Publications</u>	2
<u>Events</u>	2
<u>Multimodal Search</u>	3
<u>I-SEARCH Use Cases</u>	4
<u>I-SEARCH Architecture</u>	5
<u>3D Object Retrieval</u>	7
<u>Consortium</u>	8
<u>Contact</u>	8
<u>Project Data sheet</u>	8



A unified framework for
multimodal content SEARCH



Newsletter Outline:

Publications, Events	2
Multimodal Search	3
I-SEARCH Use Cases	4
I-SEARCH Architecture	5
3D Object Retrieval	7
Consortium	8
Contact	8

Welcome Note

Welcome to the third newsletter of the European project I-SEARCH. In this issue you will find description of latest status of the project, including more precise description of use cases, I-SEARCH architecture, novel 3D and multimodal retrieval, as well as low level descriptors extraction methods. Further information can be found on the project

web site: <http://www.isearch-project.eu>

Publications

1. Thomas Steiner, "Making Video A First Class Semantic Web Citizen", SemWebPro Conference, 17-18 January 2011, Paris, France
2. P. Daras, A. Axenopoulos, V. Darlagiannis, D. Tzovaras, X. Le Bourdon, L. Joyeux, A. Verroust-Blondet, V. Croce, T. Steiner, A. Massari, A. Camurri, S. Morin, A.D. Mezaour, L. F. Sutton, S. Spiller, "Introducing a Unified Framework for Content Object", IJMIS special issue on "Challenges in Scalable Context Aware Multimedia Computing"
3. Thomas Steiner, Jan Algermissen, "Fulfilling the Hypermedia Constraint Via HTTP OPTIONS, The HTTP Vocabulary In RDF, And Link Headers", WS-REST 2011, 28 March 2011, Hyderabad, India
4. H. Dutagaci, A. Godil, P. Daras, A. Axenopoulos, G. Litos, S. Manolopoulou, K. Goto, T. Yanagimachi, Y. Kurita, S. Kawamura, T. Furuya, R. Ohbuchi, B. Gong, J. Liu, "Generic Shape Retrieval", 3DOR 2011, 10 April 2010, Llandudno, UK
5. A. Axenopoulos, G. Litos, P. Daras, "3D Model Retrieval using Accurate Pose Estimation and View-based Similarity", ICMR2011, 17-20 April 2011, Trento, Italy
6. Thomas Steiner, "SemWebVid - An Automatable Approach Towards RDF Video Description Generation", SEMTECH 2011, 5-9 June 2011, San-Francisco, USA
7. A. Axenopoulos, S. Manolopoulou, P. Daras, "Multimodal Search and Retrieval using Manifold Learning and Query Formulation", WEB3D Conference, 20-22 June 2011, Paris, France.

Events

1. 7th Concertation meeting, 14-15 April 2011, Trento, Italy
2. Press Conference for the presentation of the Accademia Nazionale di Santa Cecilia 2011-2012 season, 4 May 2011, Rome, Italy
3. "Tecnologie per la Musica" workshop, University La Sapienza of Rome, 22 June 2011, Rome, Italy



4. CERTH/ITI Open Day 2011, 6 May 2011, Thessaloniki, Greece

Newsletter Outline:

Publications, Events	2
Multimodal Search	3
I-SEARCH Use Cases	4
I-SEARCH Architecture	5
3D Object Retrieval	7
Consortium	8
Contact	8

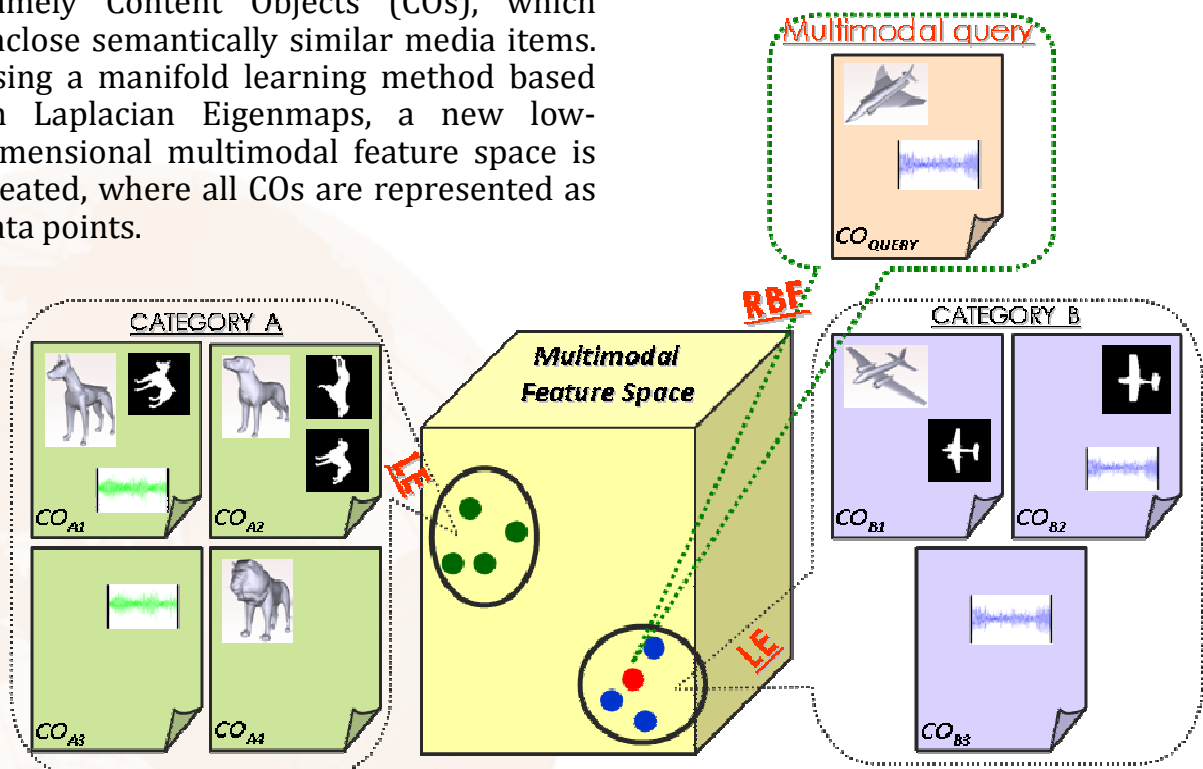
Multimodal Search and Retrieval

Due to the widespread availability of multimedia content over the Internet, the need for methods for effective search and retrieval of this content emerged. Towards this direction, several approaches for content-based multimedia retrieval have been introduced. Most of these methods are specialized to deal with only one single modality, namely 3D objects, images, video or audio. However, the emerging demands of Internet users asking for multiple media types simultaneously brought up a new generation of content-based methods, known as Cross-media Retrieval or Multimodal Retrieval.

Within the I-SEARCH project, a novel method for multimodal search and retrieval was developed. Multimedia data are organised in rich media containers, namely Content Objects (COs), which enclose semantically similar media items. Using a manifold learning method based on Laplacian Eigenmaps, a new low-dimensional multimodal feature space is created, where all COs are represented as data points.

During multimodal search and retrieval, multimodal queries are matched with the database COs with respect to their multimodal descriptor vectors. In case the query does not belong to the database, a RBF network predicts its multimodal descriptor with high accuracy.

The multimodal search method accepts queries of multiple modalities simultaneously, which is not supported by any of the cross-modal retrieval methods reported so far.



I-SEARCH Use Cases

After thorough analysis, contacts with end-user and business modelling, I-SEARCH Consortium decided to enlarge initial choice of use cases by modifying existing use cases and including additional scenarios. New scenarios even more highlight the need for enhanced semantic content analysis and representation techniques and describe relevant tasks that should be followed by the end-users in the multiple domains, aiming to efficiently perform search and retrieval activities. Moreover, updated use cases are developed taking into account possible future extensions of each use case.

Updated use cases reflect the needs of the variety of groups, such as professional and amateur musicians, teachers, furniture architects, designers and dealers, game developers and gamers, tourists and even motorcycle dealers and bikers! All these groups' needs are analyzed for exploiting a multimedia search engine for search and retrieval of multimodal content in the variety of domains:

- ❑ Music or dance **teachers, composers** or just **amateurs**:
 - Could retrieve appropriate music content using gestures, similar sound by tapping or singing, as well as real-world information.

- ❑ Big or small **party**, several friends or **disco** at night club:
 - Could realise the collective DJ experience by enabling the selection of a piece of social music through gestures and user expressions from a group of people

Newsletter Outline:

Publications, Events	2
Multimodal Search	3
I-SEARCH Use Cases	4
I-SEARCH Architecture	5
3D Object Retrieval	7
Consortium	8
Contact	8

- ❑ Living or office space **architects**, furniture **designers** and **dealers**:
 - Could access the most appropriate piece of furniture using multimodal (including GPS data for Customer proximity) content for queries

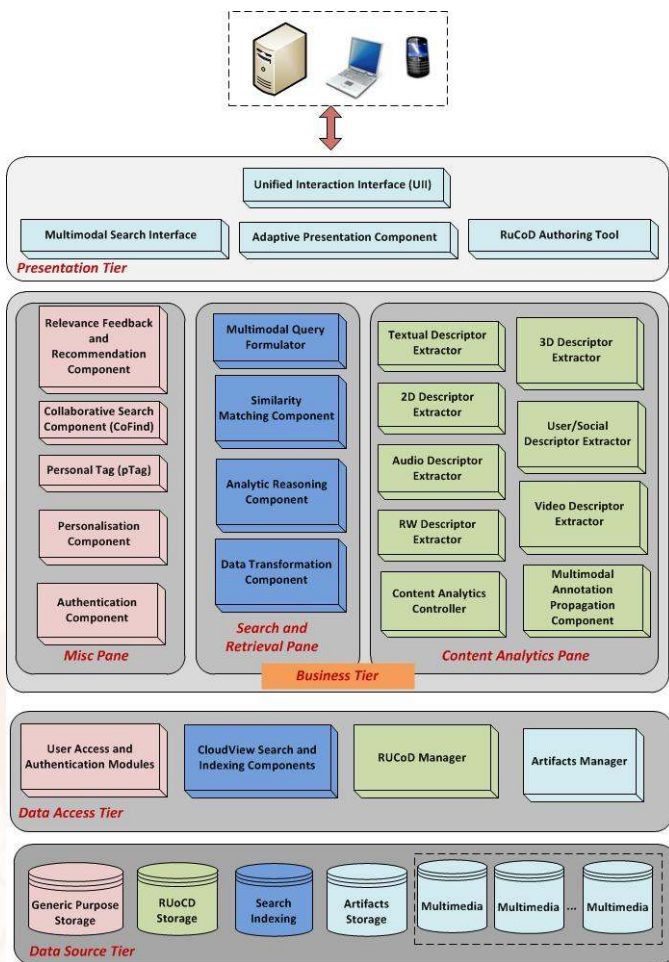
- ❑ **General applications** for multimodal search:
 - Such as retrieving multiple media types simultaneously by e.g. tourists (**cultural heritage**), usual people (**consumer goods, footwear and clothes**), teachers, etc.

- ❑ **Motorcycle dealers and bikers**
 - Could use system based on I-SEARCH concept for retrieval of motorcycles using custom details, such as sketch of shape of fuel tank or exhaust pipes, color, or even engine sound

- ❑ In the **3D games development domain**, I-SEARCH will:
 - Let a professional **3D games developer** to identify 3D content for enhancing games' visualisation
 - Enable a **gamer** to replace a virtual game avatar in order to individuate his/her an online game experience

The I-SEARCH Architecture

Despite of recent changes on the use cases, the I-SEARCH Architecture hasn't been modified and stays mainly unchanged, with the exception of adding some more components. These minor changes to the architectural structure have to do with the support of additional functionalities, especially towards the collaborative search performed by two different I-SEARCH users and the attachment of personal tags to the received results, which are now applied to the architectural design.

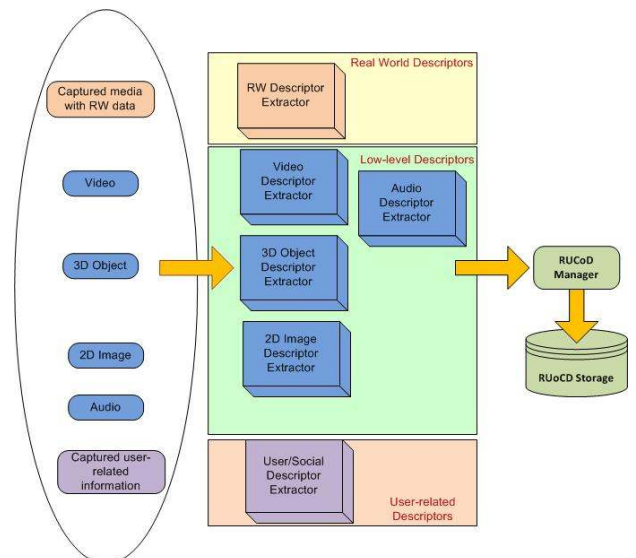


Newsletter Outline:

Publications, Events	2
Multimodal Search	3
I-SEARCH Use Cases	4
I-SEARCH Architecture	5
3D Object Retrieval	7
Consortium	8
Contact	8

Low-level multimedia descriptors

With respect to the overall architecture of I-SEARCH, extraction of low-level multimedia descriptors takes place at the Content Analytics Phase, where the RUCoD descriptors are created. More specifically, the Content Analytics Controller (CAC) is responsible for invoking the Text/ Audio/ Image/ Video/3D Object Descriptor Extractor component in order to extract low-level descriptors from audios, images, videos or 3D objects, respectively. Then, the RUCoD Manager undertakes the task to convert the extracted low-level descriptors into RUCoD files and to store them.



The extraction components in the Content Analytics Pane

Low-level multimedia descriptors (continued)

An analysis of the various low-level descriptors for each of the aforementioned media types (text, audio, image, video and 3D) has been accomplished. More specifically, the various intrinsic features for each type of content have been investigated and evaluated. The descriptor selection took seriously into account the I-SEARCH use cases. Therefore, for each use case, a mapping of the optimal descriptors has been performed. Finally, the best features have been selected and the generic low-level descriptors formats have been determined. Several issues have been addressed, such as descriptor dimension, the percentage of each feature to the descriptor vector, the vocabulary of the descriptors (i.e. numerical values, words) etc.

The final outcome of this task led to the specification of the descriptor formats for all types of media. We specified both global and local low level features for images, a specific video descriptor for video browsing, three 3D descriptors, five audio and five textual descriptors.

Generic social and real world descriptors

Generic social, emotional and real-world descriptors are extra descriptors for every RUCOD object. According to the I-SEARCH architecture, Content Analytics Controller (CAC) also invokes the components that are responsible for the emotional and real-world information.

Newsletter Outline:

Publications, Events	2
Multimodal Search	3
I-SEARCH Use Cases	4
I-SEARCH Architecture	5
3D Object Retrieval	7
Consortium	8
Contact	8

The social descriptors will be captured during the query formulation.

An analysis of the Use Cases has been performed and various descriptors have been considered. Two well known emotional descriptor alternatives has been studied and a variety of social descriptors has been considered. The real-world descriptors have been divided into four broad categories (location, environment, time and identification) and for every type a variety of representations has been considered. For each type of descriptors (social, emotional and real-world) the appropriate descriptors has been selected in order to fulfil all the needs that have been extracted during the user requirement extraction procedure and have been mapped to the use cases.

The final outcome is the formal specification of social, emotional and real-world descriptors, their representations in the RUCoD and the relative distance metrics. The defined features will be included in the RUCoD specification and, consequently, in the I-SEARCH platform.

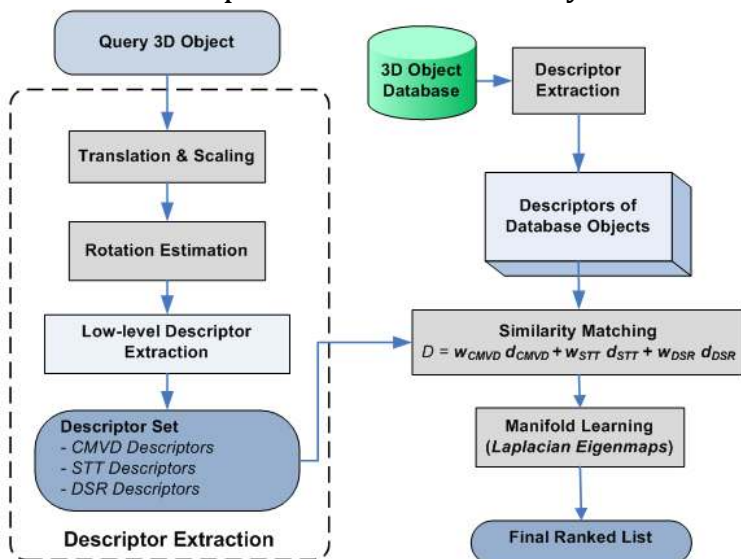
Newsletter Outline:

Publications, Events	2
Multimodal Search	3
I-SEARCH Use Cases	4
I-SEARCH Architecture	5
3D Object Retrieval	7
Consortium	8
Contact	8

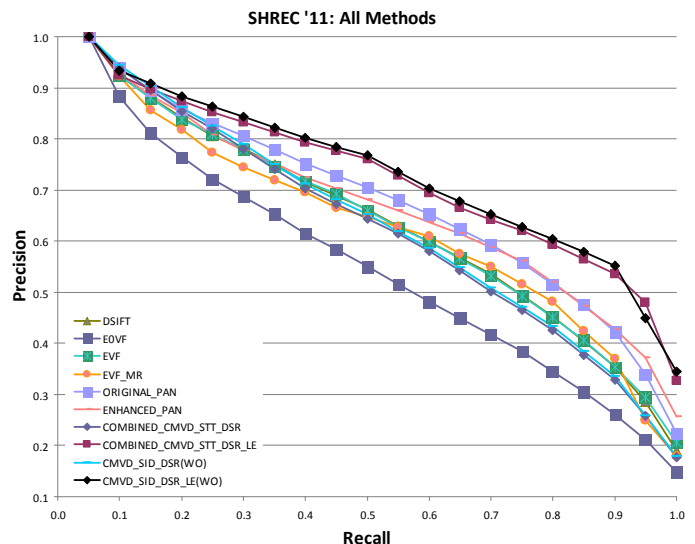
Towards More Accurate 3D Object Retrieval

A unified framework for 3D object retrieval was developed within I-SEARCH. The framework is a combination of three 3D object retrieval approaches: the Compact Multi-View Descriptor (CMVD) the Spherical Trace Transform (STT) and the Depth - Silhouette - Radial - EXTent descriptor (DSR). Moreover, two novel features are introduced in order to improve the retrieval performance: The first is a new method for rotation estimation and the second is a manifold learning approach based on Laplacian Eigenmaps.

Rotation estimation consists of the following steps: as a first step, Continuous PCA (CPCA) is applied to the input 3D object to produce a first pose estimation. Then, the reflection symmetry for the three CPCA-coordinate planes is computed. If symmetry is observed in two or three coordinate planes, the transformation is kept as it is and the process terminates. In case symmetry is observed in only one or zero coordinate planes, then, the algorithm proceeds to a correction step based on rectilinearity.



The overall dissimilarity between two 3D objects is a weighted sum of the dissimilarities of each descriptor separately. In order to improve the retrieval accuracy, an additional manifold learning based on Laplacian Eigenmaps is used.



The participation in the worldwide Shape Retrieval Contest 2011 (SHREC '11) was a great opportunity to evaluate the retrieval accuracy of the proposed framework among the best presented shape retrieval methods. The method developed within I-SEARCH achieved the best retrieval performance. More information on SHREC 2011, as well as a more detailed view of the results is available at:

<http://www.aimatshape.net/event/SHREC/>

Consortium



INFORMATICS & TELEMATICS INSTITUTE
Centre for research and technology

Centre for Research & Technology Hellas – Informatics & Telematics Institute



JCP – Consult



INRIA

INRIA



Athens Technology Center



ENGINEERING INGEGNERIA INFORMATICA

Engineering Ingegneria Informatica S.p.A.



Google



University of Genoa



exalead®
connect the dots

Exalead



FACHHOCHSCHULE ERFURT UNIVERSITY OF APPLIED SCIENCES
Angewandte Informatik

Erfurt University of Applied Sciences



Accademia Nazionale di Santa Cecilia



EasternGraphics
visualize your business

EasternGraphics

Newsletter Outline:

Publications, Events	2
Multimodal Search	3
I-SEARCH Use Cases	4
I-SEARCH Architecture	5
3D Object Retrieval	7
Consortium	8
Contact	8

Contact

Project Coordinator

Dr. Dimitrios Tzovaras
Centre for Research & Technology Hellas
Informatics & Telematics Institute
e-mail: Dimitrios.Tzovaras@iti.gr

Technical Manager

Dr. Petros Daras
Centre for Research & Technology Hellas
Informatics & Telematics Institute
e-mail: daras@iti.gr

Project Data Sheet

Acronym	: I-SEARCH
Full Name	: A unified framework for multimodal content SEARCH
URL	: http://www.isearch-project.eu
Programme	: FP7-ICT-2009-4
Strategic Objective	: ICT-2009.1.5: Networked Media and 3D Internet
Start Date	: 1 January 2010
Duration	: 36 Months

The 4th issue of the I-SEARCH Newsletter will be released on January 2012 presenting the progress of the project for the period since the 3rd issue.