

# WT2.17: The Informatics & Telematics Institute, Thessaloniki, Greece

## Report

Activities performed during the visit

in           Wrocław, POLAND

period:   29<sup>th</sup> September 2013 -2<sup>nd</sup> October 2013

author:   Stefanos Vorochidis, Jerzy Sas

## 1 Personal Information

Mr. Stefanos Vrochidis, faculty member of Information Technologies Institute, Centre for Research and Technology Hellas (ITI-CERTH), Greece visited

Wroclaw University of Technology, Poland. in the period from 29<sup>th</sup> September 2013 to 2<sup>nd</sup> October 2013 in order to carry out research and training activities in the field of data mining and machine vision, multimedia analysis and retrieval

## 2 Information about Seminars

The seminar presentation was organized on Tuesday, October 1, 2013

It was entitled: “Collaboration in the areas of multimedia analysis and retrieval”

## 3 Description of scientific activities

The purpose of this seminar/meeting was a presentation of the research activities of the Multimedia Group of ITI-CERTH and of Wroclaw University of Technology in the areas of multimedia analysis and retrieval with a view to establish close cooperation, know how exchange, as well as submission of common publications and proposals.

ITI-CERTH presented their work in multimedia analysis focusing in two axes:

- The presentation of a new fp7 project MULTISENSOR (starting date November 2013), which is coordinated by ITI-CERTH. MULTISENSOR will mine heterogeneous data from the aforementioned resources and apply multidimensional content integration. To achieve multidimensional integration of heterogeneous resources, MULTISENSOR proposes a content integration framework that builds upon multimedia mining, knowledge extraction, analysis of computer-mediated interaction, topic detection, semantic and multimodal representation as well as hybrid reasoning. The developed technologies will be validated with the aid of 2 main use cases: a) International mass media news monitoring and b) SME internationalisation.
- The presentation of works relevant to patent image search including image extraction from patents, feature generation, visual similarity search and concept extraction. Feature extraction is based on Adaptive Hierarchical Density Histograms, while concept extraction builds upon supervised machine learning using Support Vector Machines.

The publications of Multimedia Group are available at: <http://mklab.iti.gr/publications>.

On the other hand the Wroclaw University of Technology presented their main works in the area of multimedia analysis and retrieval focusing on:

- Image retrieval approaches based on SIFT features. These include development of search engines to support visual similarity search by example and automatic annotation.

- Map processing to extract geographical and topological information. Optical character recognition is applied to identify geographical locations, which are compared to geographical resources such as GeoNames. Map scale is automatically calculated based on the geographical names distance.
- Automatic processing of diagrams, graphs and flowcharts based on text to graphics separation, OCR and shape recognition finding out connectivities between recognized shapes and determining additional attributes of links like arrows, link line styles and widths. This technique seems to operate very well for circuits and flowcharts.

Based on the aforementioned presentations and discussions the following areas are defined for cooperation in the context of the ENGINE project including specific targets.

- MULTISENSOR-ENGINE cooperation: ITI-CERTH will provide constant updates and direct information on the results produced by MULTISENSOR project. These could include deliverables, tools and resources that will be developed during the project. Additional interactions between MULTISENSOR and ENGINE will be also investigated.
- Image and video retrieval. Both teams are interested in developing search engines for video and image retrieval. We can investigate common participation in evaluation workshops such as TRECVID and ImageCLEF 2014. The deadlines for the instance search task of TRECVID are usually end of August 2014, while the deadline for ImageCLEF is around end of April 2014.
- Patent Image Search. Based on the ITI-CERTH works on patent image retrieval and the works of Wroclaw University on document image analysis, figure segmentation and flowchart processing we could create a common paper for patent image search to be published in conferences such as IRFConference, which is very relevant to patent search (deadline ~ May of 2014) and Journals such as the World Patent Information. CLEF-IP would have been an additional option for participation, however CLEF-IP was not organized in 2013 it is possible that it will not be organized for 2014 as well.
- Maps processing. By combining common works we can perform experiments for automatic extraction of information from maps that include geographical names and focus on environmental information such as air quality and pollen forecasts. Such works could be published in special sessions of information retrieval conference relevant to environment, workshops, as well as in Journals such as the Ecology Informatics.

The aforementioned targets will enhance the collaboration of the two groups and facilitate the know-how and knowledge exchange. The results will be also reflected by common participation in evaluation workshops and publications, while participation in common proposals will be investigated.

In order to define the collaboration directions more precisely, the participating parties should exchange papers and articles published so far to better understand used and developed methods. The locally gathered benchmark datasets can be also exchanged in order to test elaborated methods and algorithms using the second party data. Next,

necessary improvements can be proposed by combining approaches used by both parties. It can be a basis for common publications.

Initial tests on exchanged data should be completed in the beginning of February 2014. Then the next short incoming visit can be organized near March 2014. The aim of it will be to work together on particular improvements of methods and approaches elaborated as a result of collaboration.