

Multimedia Information Retrieval on Very Large Scale ^{*}

Giuseppe Amato, Fabrizio Falchi, Claudio Gennaro

ISTI-CNR, Pisa, Italy

`firstname.lastname@isti.cnr.it`

1 Duration: full-day tutorial

2 Abstract

This tutorial will discuss the issues related to Multimedia Information Retrieval in very large dataset of visual documents. Content based retrieval typically is not performed using the visual content itself, rather visual features are extracted and retrieval is performed searching by similarity on the extracted features. Similarity search is a difficult task because efficient techniques to process database or text queries cannot be applied here. Therefore in the last decades researcher have investigated techniques for executing similarity search efficiently and in a scalable way.

3 Motivation, Target audience, and Interest to the SAC community

Information on the web and on the social networks is becoming more and more multimedia driven rather than text driven. People more often publish and post pictures and videos, however, the associated textual information is generally poor and does not help in searching and analyzing effectively this type of content. Techniques able to effectively and efficiently deal with multimedia data is becoming of paramount importance. All major search engines are nowadays providing visual search functionalities and the research effort in this direction is rapidly increasing.

Target audience includes, researcher willing to know more about the issues and solutions on this topic, and student doing research in topics related to Multimedia Information retrieval.

4 Outline

This tutorial will discuss the issues related to Multimedia Information Retrieval in very large dataset of visual documents.

The tutorial is composed of three distinct parts:

- Feature extraction techniques for multimedia information retrieval (2h:30m)
- Efficient indexing strategies for similarity searching on a large scale (2h:30m)

^{*} This work was partially supported by the EAGLE project, co-funded by the co-funded by the European Commission within the CIP-Best Practice Network Programme.

- Multimedia information retrieval on a large scale using open source solutions (1h)

The tutorial will first give an overview of the most popular visual features for content based retrieval. It will introduce the basics related to *global visual descriptors*, then will introduce the *local visual descriptors*, which are more suited for object recognition, and finally it will also introduce the most recent results related to *deep features*, obtained using Convolutional Neural Networks, which provide us with a leap forward in dealing with the semantic gap.

When visual documents are described using visual features, they are generally compared using distance (or dis-similarity) measures. Various indexing strategies and search algorithms based on distance function were defined during the last decade. Among these methods the tutorial will introduce the LSH methods (*Locality-Sensitive Hashing*), based on similarity hashing techniques, and the *permutation-based methods*, where documents are represented as permutations of a set of reference objects, and similarity between documents is approximated by comparing permutations. These indexes allow similarity retrieval to be executed very efficiently, in datasets containing hundred millions images, with limited computing and storage resources.

Finally, the tutorial will discuss how the above concepts can be put in practice using Open Source solutions. It will introduce the OpenCV library, that offers various tools for image analysis and feature extraction. The tutorial will also discuss how Lucene, the text retrieval engine, can be easily used to index and search visual documents.

4.1 Relevant Bibliography

- **Image Features:**
- "MPEG-7". ISO/IEC JTC1/SC29/WG11N6828. 2004
<http://mpeg.chiariglione.org/standards/mpeg-7/mpeg-7.htm>.
- Distinctive Image Features from Scale-Invariant Keypoints. Lowe, David: Kluwer Academic Publishers Hingham, MA, USA , 2004, International Journal of Computer Vision, Vol. 60.
- "SURF: Speeded Up Robust Features. Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool,. 3, 2008, Computer Vision and Image Understanding (CVIU), Vol. 110, pp. 346–359.
- "Aggregating local descriptors into a compact image representation", Jegou H., Douze M., Schmid C., Perez, P., IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010, 13-18 June 2010, San Francisco, CA
- CNN Features Off-the-Shelf: An Astounding Baseline for Recognition, Razavian A.S., Azizpour H., Sullivan J., Carlsson S., IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 23-28 June 2014, Columbus, OH
- **Image retrieval/recognition libraries:**
- OpenCV, [Online] <http://opencv.org/>
- Apache Lucene, [Online]<http://lucene.apache.org/>
- **Efficient similarity Search:**
- Similarity Search The Metric Space Approach, Pavel Zezula, Giuseppe Amato, Vlastislav Dohnal, Michal Batko, 2005 Springer

- Effective proximity retrieval by ordering permutations,, E. Chavez Gonzalez, K. Figueroa, and G. Navarro, IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 9, pp. 16471658, 2008
- Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions, A. Andoni and P. Indyk,, Commun. ACM, vol. 51, no. 1, Jan. 2008
- Use of Permutation Prefixes for Efficient and Scalable Approximate Similarity Search. Esuli A. Information Processing and Management, Vol. 48, N. 5, 2012
- MI-File: Using Inverted Files for Scalable Approximate Similarity Search, Multimedia Tools and Applications Journal, Springer, Volume 71, N. 3, August 2014
- **Image search recognition classification**
- Tour the World: building a web-scale landmark recognition engine. Yan-Tao Zheng, Ming Zhao, Yang Song, Hartwig Adam, Ulrich Buddemeie, Alessandro Bissacco, Fernando Brucher, Tat-Seng Chua, Hartmut Neven. 2009. CVPR.
- Geometric consistency checks for kNN based image classification relying on local features. Giuseppe Amato, Fabrizio Falchi, Claudio Gennaro. 2011. SISAP '11 Proceedings of the Fourth International Conference on Similarity Search and Applications.
- Fast Image Classification for Monument Recognition, Giuseppe Amato, Fabrizio Falchi, Claudio Gennaro, ACM Journal on Computing and Cultural Heritage (JOCCH), Volume 8 Issue 4, August 2015

5 Specific goals and objectives

Multimedia information retrieval has to face two issues. On one hand visual information extracted from images should be good enough to allow effective retrieval of documents to be executed. On the other hand search algorithms should be able to scale to huge (web scale) datasets. The tutorial will address both topics and will stress how solutions on one side affect the other side. Finally, the tutorial will show how already existing open source solutions can be used to address these problems.

6 Expected background of the audience

Basic computer science background. Information retrieval might be a plus.

7 Bibliographic sketch of the proposers

Giuseppe Amato graduated in Computer Science at the University of Pisa, Italy, in 1992 and was awarded a PhD in Computer Science at the University of Dortmund, Germany, in 2002. Since 1994 he has been a member of the research staff of ISTI-CNR (previously IEI-CNR) in Pisa, working in the area of Multimedia Information Systems. His main research interests are content-based retrieval of multimedia documents, access methods for similarity search of multimedia documents, wireless sensor networks. He has published in international journals and conferences in the areas of information

systems and multimedia information retrieval. He has participated in several EC and national funded research actions in the area of multimedia information retrieval.

Fabrizio Falchi is researcher at the Networked Multimedia Information System Laboratory (NeMIS) of the Information Science and Technologies Institute (ISTI) of the Italian CNR in Pisa. He has a Ph.D. in Information Engineering from University of Pisa (Italy), and a Ph.D. in Informatics from Faculty of Informatics of Masaryk University of Brno (Czech Republic). He also received an M.B.A. from Scuola Superiore Sant'Anna in Pisa. His research interests include similarity search, distributed indexes, multimedia information retrieval, computer vision, peer-to-peer systems. He has participated in several EC and National research projects. He has published in international journals and conferences in the areas of information systems and multimedia information retrieval.

Claudio Gennaro is a researcher at CNR-ISTI. He received the "Laurea" degree in Electronic Engineering from the University of Pisa, Italy, in 1994 and Master Degree in Information Technology at CEFRIEL of Milan. He received PhD Degree in Computer Engineering and Automatica in 1999 from Politecnico di Milano. His main interests are: Wireless Sensor Networks, Access Structures for Multimedia Retrieval, Peer-to-Peer Systems, Digital Libraries, Model of Metadata for Audio/Video, Performance Evaluation, and Parallel Computing. Claudio Gennaro has had considerable previous experience in participation of European projects and has published several articles in the area of Multimedia Information Retrieval

8 Needed equipment: Projector

9 Teaching materials

Other tutorials given by the proposers are:

- "Information retrieval and Web Search Engines", G. Amato, F. Sebastiani, Eighth International World Wide Web Conference, May 11-14 1999, Toronto Canada.
- "Audio/Video Digital Libraries: designing, searching for documents, and generating Metadata, G. Amato, C. Gennaro, P. Savino, at the 2003 Joint Conference on Digital Libraries, JCDL 2003, Houston on May 27-31.
(slides: <http://www.nmis.isti.cnr.it/amato/JCDL03/index.htm>)
- Indexing and searching Audio/Video Documents in Multimedia Digital Libraries, G. Amato, C. Gennaro, P. Savino, at the 2003 European Conference on Digital Libraries, ECDL 2003, Trondheim, Norway, on August 17-22.
(slides: <http://www.nmis.isti.cnr.it/amato/ECDL03/index.htm>)
- "Similarity Search: The Metric Space Approach", G. Amato, P. Zezula, Vlatislav Dohnal at the 22nd Annual ACM Symposium on Applied computing, Seoul, Korea, March 11-15, 2007
(slides: <http://www.nmis.isti.cnr.it/amato/similarity-search-book/tutorial-SAC-2007.htm>)
- "Content Based Retrieval on Very Large Visual Document Archives" G. Amato, Russian Conference on Digital Libraries, RCDL 2012, Pereslavl, Russia, 15-18 Ottobre 2012
(slides: <http://www.nmis.isti.cnr.it/amato/RCDL2012/RCDLTutorial.pdf>)