

Yannis Kalantidis

National Technical University of Athens

Email: ykalant@image.ntua.gr

PhD thesis: *Nearest Neighbor Search and Clustering for Large scale Visual Search*

Estimated graduation date: November 2014

Research Statement

I am a PhD student at the National Technical University of Athens since November 2009 under the supervision of Prof. Stefanos Kollias and Dr. Yannis Avrithis, working also closely with Dr. Giorgos Tolias. During the last four and a half years I have worked on different aspects of large scale visual search, clustering and approximate nearest neighbor search. My publication record includes papers in CVPR, ECCV, CVIU, ACM Multimedia, ICMR and ICIP. Figure 1 presents a **graphical overview** of main research areas and publications.

In 2010 I worked on geometry indexing and clustering of large image databases, research that resulted in two oral **ACM Multimedia** papers and two journal publications. One of the works was an effort to include global image geometry in the indexing stage of visual search. We proposed ways to cope with large scale geometry indexing using hashing and feature selection (**CVIU 2014**). The second publication presented an unsupervised approach for clustering large geo-tagged urban photo collections in coherent visual clusters for landmark and non-landmark recognition (**MTAP 2011**). Our *Scene Maps* are able to both compress the index size and boost recall at the same time. I was also a principal developer for *VIRaL*¹, *Visual Image Retrieval and Localization*, an online demonstration of the aforementioned research.

In the summer of 2010 I worked as a research intern at *Yahoo! Research Barcelona*, where collaborated with Roelof Van Zwol, Lluís Garcia Pueyo and Michele Trevisiol on scalable logo recognition. This research was accepted at **ICMR 2011**, where we presented an *multi-scale Delaunay Triangulation* approach for local features that was able to index and recognize thousands of logos given a query image in milliseconds.

In 2011 I worked on feature selection for unique images in the context of visual search. We exploited self-similarities, symmetries and repeating patterns to select features within a single image and retained similar search performance using only a small fraction of the index size. This research was accepted as an oral presentation in **ACM Multimedia 2012**.

In 2012 my main focus was on visual vocabulary construction for large-scale visual search. This requires clustering 10^7 high-dimensional points into 10^6 clusters – a scale for which only very few and limited clustering options exist. We proposed *Approximate Gaussian Mixtures*, a modification of the EM algorithm that is not only able to scale well via approximate and incremental assignments but is also able to automatically select the final number of clusters needed, exceeding the current state-of-the-art in terms of visual search performance. This work was accepted at **ECCV 2012**.

I spent the summer of 2012 in California, working at *Yahoo! Research Silicon Valley*, where I collaborated with Lyndon Kennedy and Li-Jia Li. The project given to me was *visual clothing recognition, segmentation and product search*. My work resulted in a fully automated approach that is able to recognize clothing classes and suggest relevant clothing products given only a single real-world image. This cross-scenario approach performs as good as the state-of-the-art while being more than an order of magnitude faster. Results are published in the proceedings of **ICMR 2013** and a relevant **US Patent application** has been submitted.

In 2013 I worked on large scale approximate nearest neighbor search. Our method, *Locally Optimized Product Quantization*, extends product quantization by locally optimizing rotations and sub-quantizers in a multi-index framework. It achieved impressive results, exceeding the current state-of-the-art by more than 10% in terms of precision on a billion-scale dataset. It was accepted as a poster presentation in **CVPR 2014**.

For the summer of 2014 I am an intern at *Yahoo! Labs* in San Francisco, working with Lyndon Kennedy and the Flickr Computer Vision team. We are working towards the visual and semantic indexing of the Flickr corpus and zero-shot learning.

¹<http://viral.image.ntua.gr/>

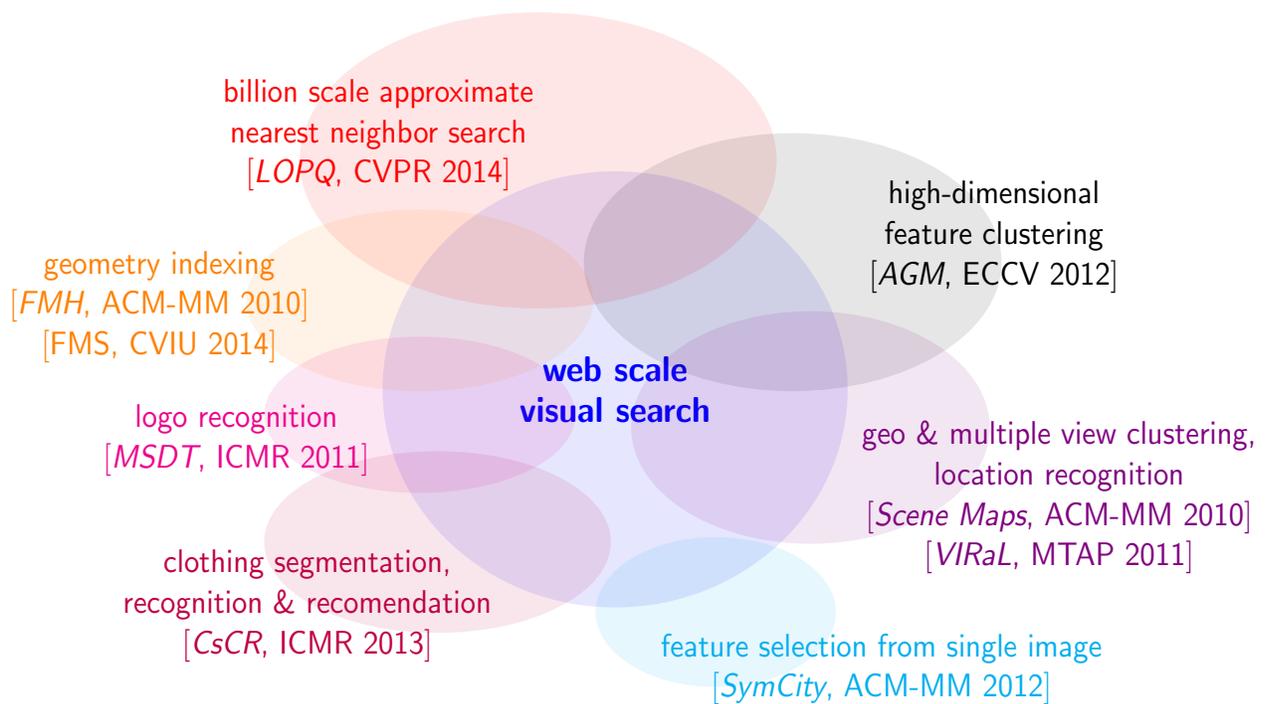


Figure 1: A graphical overview of selected research areas and related publications.