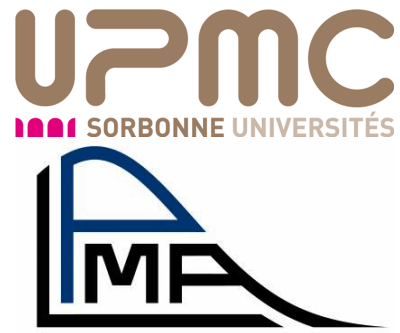




+33 (0)6 09 09 43 54
victor@fotonower.com
www.fotonower.com



Gilles Pagès
gilles.pages@upmc.fr
LPMA

Beyond deep learning : Massive Media Indexing with
Parallel Computing.

- Indexing database** Unsupervised and supervised learning algorithm are used to index media. Fotonower proposes to enhance several algorithms and work on their parallelization :
- Kohonen self organizing map, Quantization (Lloyds) [5].
 - Building and enhancing a hierarchical tree [2].
 - Cuda implementation to index media [4].

Practical work and requirements

Student Team The student should have a practical experience of Monte-Carlo simulation and a theoretical knowledge of Lloyd's algorithm. An understanding of some classical machine learning algorithm like Random Forest, Hierarchical tree or others would be of high interest. The student team will need to be able to have good knowledge in programming like python or c++ and should be keen to learn to develop in a parallel language like CUDA.

Fotonower Fotonower will provide access to computing power, GPU in the cloud as well as data of photos described in deep neural network.

State of the art

Datas : Massive sharing of photos and videos from Social Media Nowadays massive amount of photos and videos are shared publicly on social media, Twitter, Instagram and some part of Facebook for the most well known in US and Europa. These photos can be described as vector of coefficients describing each pixel in three different colors for each frame in case of a video. These datas can be considered in an euclidean space but the euclidean distance is not linked to any particular interesting property. At this point a media is stored in around 100,000 to a few millions of dimensions.

Theoretical background : Supervised Convolutional Neural Network for euclidean distance sensible with similarity of photos (inverse or opposite of the distance). Neural network, developed since a few decades have proven to be able to identify automatically written text and more recently more complex scenes (mountain, city, indoor, demonstration), objects (shoes, animal, fashion accessories) or distinguish different faces. They uses manually labeled datas and stochastic algorithm descent in order to calibrate themselves. The convolutional property of the first layer lower the amount of parameters to optimize and enable the neural network to get a description of low features of images. Therefore it can describe in a sensible way any images although it was not in the manually labelled datas. After this step, a media is stored in around 10,000 dimensions. We will call them features. We therefore uses it in a non supervised methodology known as k-mean or Lloyd's algorithm in order to index these datas.

Theoretical background : Quantization. This methodology is known as quantization. It is used for numerical integration, clustering of datas or historically transmission of datas. Building optimal quantizer uses the Lloyd's algorithm and has the property of minimizing the distortion, known as inertia when using the quadratic norm.

Preliminary study : First tests and implementation The issue in our case is to be able to store in our index tens of millions to billions of media. We have currently some implementation prototype with brute technic storing 5,000,000 photos media. We have studied, inspired by [1], the statistic of the features. We have proposed an enhancement of Lloyd's algorithm know as Split and Stick. Thanks functional quantization we got some hint on the optimal dimension of quantizing or indexing these datas. The idea is to build a hierarchical tree for high dimension datas [2]. The brute methodology we have already implemented is a tree with one cell. The querying with a GPU takes around one second. Our aim is to attains the same efficiency with 20 to 200 times more media. Here is some bibliography attached that uses non deep learning feature for a similar task : [4]. [6] presents general knowledge on nearest neighbour search technic.

Enhancement of parallel indexing media technology We aim to implement in parallel advanced technic for quicker indexing and retrieval, like partial distance search or split and stick enhancement of Lloyd Algorithm. Gilles Pagès has initiated in the 90's using quantization for numerical integration with application to option pricing and he is today a specialist of this field which is developed in [3]. Intuitively the curse of the dimension makes this inefficient for dimension above 12. Moreover historic classification methodology like Kohonen (an extension of quantization) apply to a few hundred of datas. We wish to break both these limits by applying quantization clustering in dimension around hundred to thousand and with million of items to cluster. The goal is to develop knowledges in related domains.

Context - Start-up

Fotonower is a search engine of photos crawled on social media, www.fotonower.com. Fotonower develops technology to search and sort photos. One of its main source of data is social media. It delivers services to media and tourism industry. Fotonower photo recognition technologies were built in collaboration with members of IMAGINE at ENPC and LPMA at UPMC.

Bibliography

Références

- [1] Mathieu Aubry and Bryan C Russell. *Understanding deep features with computer-generated imagery*. In Proceedings of the IEEE International Conference on Computer Vision, pages 2875–2883, 2015.
- [2] Marius Muja David G. Lowe. *Fast matching of binary features*. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 36(11), November 2014.
- [3] Pagès. G. *Introduction to Numerical Probability applied to Finance*. Springer, to appear, 2017.
- [4] Vincent Garcia, Eric Debreuve, and Michel Barlaud. *Fast k nearest neighbor search using gpu*. In Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on, pages 1–6. IEEE, 2008.
- [5] Harald Luschgy and Gilles Pagès. *Sharp asymptotics of the Kolmogorov entropy for Gaussian measures*. J. Funct. Anal., 212(1) :89–120, 2004.
- [6] David M Mount. *Ann programming manual*. [http link : http://www.cs.umd.edu/~mount/ANN/Files/1.1](http://www.cs.umd.edu/~mount/ANN/Files/1.1), 1, 2006.