Toothpic: Who took this picture?

# TOOTHPIC: WHO TOOK THIS PICTURE?

*D. Valsesia, G. Coluccia, T. Bianchi, E. Magli*

Politecnico di Torino - DET, Italy
{name.surname}@polito.it

## 1. INTRODUCTION

Managing photos over the Internet is an increasingly bigger problem with sizeable social and financial implications. More than 1 billion Facebook users upload in excess of 350 millions of new pictures per day and 250 billions overall, not to mention other social media sites such as Flickr, Instagram, Google+, Tumblr and Pinterest, and these figures are growing steadily. It is hence not surprising that a huge problem has arisen for the users and social media sites alike: it is very difficult to track down a wide range of improper uses of the photos, such as exploiting them for commercial purposes, re-posting others' photos without consent or infringing copyright, posting photos containing unethical or illegal contents, and so forth. To quantify the magnitude of these issues, consider that social media company The Content Factory reportedly got sued $8,000 for using an image on a blog post that got less than 100 visitors. Pinterest's users are held responsible for the pictures they post or re-post, plus they automatically grant licence for reuse on Pinterest. Since many of Pinterest's 70+ million users arguably do not own these rights, they may greatly suffer from legal actions, and this may in turn disrupt the service's popularity.

The solution to this problem requires the ability to effectively and properly manage large-scale databases of photos; this is an extremely desirable feature for several types of users and has plenty of potential applications with huge economical and legal impact. At the same time, it is a very challenging problem with few feasible technical solutions. Camera identification is a key technology to solve this problem, allowing to link a photo to the device that has shot it. Specifically, any digital imaging sensor leaves its own unique fingerprint, called Photo Response Non-Uniformity (PRNU) [1], in all pictures, and the fingerprint can be detected and compared with a database of known fingerprints. Despite having numerous appealing applications, however, camera identification has been so far confined to a few forensic use cases. Indeed, state-of-the-art techniques are limited to a small scale because of the complexity of fingerprint matching and storage required by the fingerprint database. As a consequence, only few individuals have been able to benefit from this technology, and the applications that are more promising from the marketing and social standpoints still remain infeasible.

This work is supported by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement n.279848 and under the European Community's Horizon 2020 Programme / ERC Grant agreement n.665421.

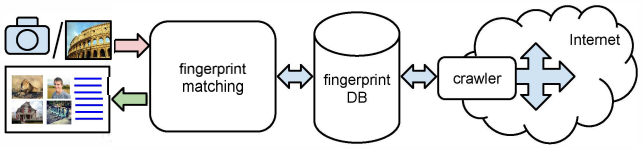


**Fig. 1**. High-level block diagram of the proposed system.

This project aims at validating a breakthrough technology to solve the problems described above. The proposed technology applies some research results of an ongoing ERC starting grant project, namely a new compressed fingerprint format based on random projections, which advances the state-of-the-art of camera identification, boosting its performance by orders of magnitude in terms of camera matching speed, energy consumption, bandwidth usage and complexity. Toothpic will go beyond a simple lab demo, providing a sound demonstration of the feasibility of this technology, and proving that it can achieve sufficiently high throughput so as to (a) enable large-scale camera identification and hence detection of improper image use over large image databases, e.g. social media sites, (b) do this in real-time and at a low cost, allowing to reach a huge number of potential users.

## 2. THE TECHNOLOGY

A large collection of pictures is obtained by scanning portions of the web, using a crawler software that downloads photos from publicly available repositories. From this collection, a large fingerprint database is automatically generated by extracting the PRNU pattern of each individual photo. A query is presented to the system in the form of a camera fingerprint, and the goal is to retrieve all the photos acquired by the same device. The proposed system, visualized in Fig. 1, is a realistic example of the envisioned search engine, and can be used to demonstrate that the above task is indeed feasible. Notice that this image retrieval problem is significantly different from the problems addressed by the well-researched area of content-based image retrieval where the *content* of an image is the query and the user searches for images with similar con-

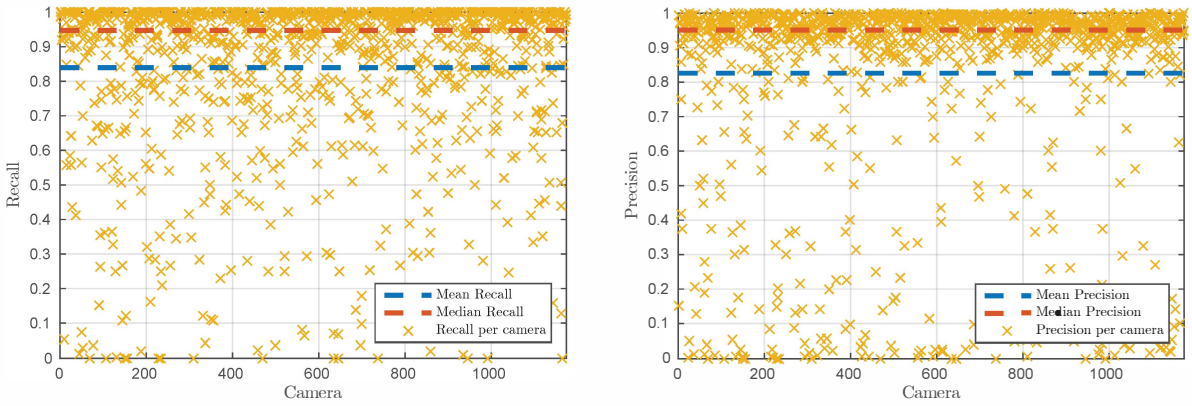


**Fig. 2**. Perfomance of Toothpic. Database: 500,000 photos. Disk usage: 25 GiB. Total RAM usage: 3.6 GiB

tent.

Due to the characteristics of PRNU, a technique for obtaining a compact representation of PRNU fingerprints is essential to deal with such a big data scenario. For example, assuming a collection of 6 billion photos, which is the number of photos hosted by Flickr in 2011, and an average sensor resolution of 12 megapixels, the whole uncompressed fingerprint database, represented in single precision, would require about 250 petabytes of storage. Toothpic solves the above problem by using compressed fingerprints obtained via proper quantization of random projections. A recent work [2] has shown that random projections permit to obtain very compact representations with limited performance loss in terms of matching accuracy. Moreover, this scheme is flexible, meaning that the number of projections can be tuned according to the desired trade-off between compression and accuracy. Hence, this technique appears to be one of the best candidates for performing camera-based image retrieval in very large scale scenarios [3]. In order to address search problems on huge scales, we propose a solution based on a two-step approach, in which a first search is performed over the entire database using a coarse version of the compressed fingerprint returning a subset of the database; a second search is then performed on this subset using a refined fingerprint. An improved version of the above technique is then considered, in which the coarse version of the compressed fingerprint is obtained by adaptively choosing the random projections with the largest magnitude. Toothpic is also robust to image rescaling.

Preliminary results have been presented in [3] (see Fig.2) showing that the system can achieve a precision and recall around 95% on a dataset composed of about 500,000 pictures collected in an unsupervised way from photo sharing website Flickr.

## 3. ONGOING WORK

A prototype camera search engine is under development with a target database size of 50+ million photos. The system is expected to reply to a query in near real time by using the proposed compression and retrieval algorithms.

First, the database is assembled by scanning a portion of Flickr, downloading user photos and extracting an estimate of the fingerprint from each photo. This fingerprint is stored in our compressed format along with some metadata. The metadata include username and Exif data of the photo. This information is used to generate a ground truth used in the testing phase. Moreover, the URL of the web page containing the photo is stored to be returned upon a user query. The system also stores the aspect ratio of the images, as it will be used to speed up the search by filtering only the images matching the aspect ratio of the query.

The database is implemented with Aerospike, a popular high-performance NoSQL database. Aerospike has several desirable features to accommodate our algorithms. First, it allows to keep tables in RAM which is the preferred solution for the first stage of the search algorithm (prefiltering with a secondary index on the aspect ratio and coarse search using adaptively embedded fingerprints). Second, it follows the Map-Reduce computational paradigm, which allows distributed computations and easily scales to multiple machines.

The tests are being conducted on two servers equipped with 384 GB of RAM and 1.6 TB of fast SSD storage each. This is a high-performance platform but with a limited degree of parallelism. In order to show how effectively the system scales with the number of machines, tests will be conducted on large number of slower machines.

## 4. REFERENCES

[1] J. Fridrich, "Digital image forensics," *Signal Processing Magazine, IEEE*, vol. 26, no. 2, pp. 26–37, 2009.

[2] D. Valsesia, G. Coluccia, T. Bianchi, and E. Magli, "Compressed fingerprint matching and camera identification via random projections," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 7, pp. 1472–1485, July 2015.

[3] D. Valsesia, G. Coluccia, T. Bianchi, and E. Magli, "Large-scale image retrieval based on compressed camera identification," *Multimedia, IEEE Transactions on*, vol. 17, no. 9, pp. 1439–1449, Sept 2015.