Face Match for Family Reunification: Real-World Face Image Retrieval

Eugene Borovikov, Communications Engineering Branch, U.S. National Library of Medicine, Bethesda, MD, USA
Szilard Vajda, Central Washington University, Ellensburg, WA, USA
Michael Gill, U.S. National Library of Medicine, Bethesda, MD, USA

ABSTRACT

Despite the many advances in face recognition technology, practical face detection and matching for unconstrained images remain challenging. A real-world Face Image Retrieval (FIR) system is described in this paper. It is based on optimally weighted image descriptor ensemble utilized in single-image-per-person (SIPP) approach that works with large unconstrained digital photo collections. The described visual search can be deployed in many applications, e.g. person location in post-disaster scenarios, helping families reunite quicker. It provides efficient means for face detection, matching and annotation, working with images of variable quality, requiring no time-consuming training, yet showing commercial performance levels.

KEYWORDS

Face Detection, Face Recognition, Image Retrieval, Post-Disaster Family Reunification

INTRODUCTION

The Content Based Image Retrieval (CBIR) technology has seen significant advances recently resulting in many useful web-scale image search techniques (Dharani & Aroquiaraj, 2013). Several web search engines (e.g. bing.com/images, images.google.com, yandex.com/images) employ those techniques to provide visual search capabilities. The face recognition (FR) technology has also seen a considerable progress during the last decade, in several cases approaching human-level accuracy in face detection and verification tasks (Naruniec, 2010; Tan, Chen, Zhou, & Zhang, 2006; Zhang & Zhang, 2010), especially in well-controlled environments such as studios.

Modern web-based FR solutions (e.g. in facebook.com or plus.google.com) may work well with limited face datasets (e.g. user circles, family albums) that tend to contain tagged pictures of the same few individuals (e.g. family and friends) with multiple photos per person, which allows for user-specific recognition model training. Our experience did not provide us with an abundance of publicly available single image per person (SIPP) face image retrieval systems that can work effectively using no training with millions of unconstrained face images, presenting many challenges for such systems in practice, e.g. disaster recovery:

DOI: 10.4018/IJCVIP.2017040102

Copyright © 2017, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
No constraints on gallery or query pictures, as in Figure 1;
• Often suboptimal quality images for query and gallery;
• Dataset size: web-scale collections with many near-duplicates¹;
• Large inconsistency in query and gallery face appearance.

Many of those challenges are being addressed by modern FR systems thanks to the emergence of labeled datasets with constrained-free images (Beveridge et al., 2013; Huang, Ramesh, Berg, & Learned-Miller, 2007; Kemelmacher-Shlizerman, Seitz, Miller, & Brossard, 2016) utilized for various competitions. Development of such challenging datasets presents a great opportunity to assess capabilities of the existing systems on the real-world data, and then improve them or develop some new capabilities, ultimately approaching a human-level visual matching accuracy.

Typical FR systems would approach the face recognition problem in one of the two formulations (Zhou et al., 2014): verification (answering if photos depict the same person) or identification (suggest the person ID by visual similarity to the query image). Such systems usually require some sort of model training, using multiple photos per individual. They would typically work with a set of visual features extracted from images by learning a measure of visual similarity, modeling human visual perception of faces. While modern automatic face classification and verification methods can work fairly well on good quality (fairly well lit, sharp, 80x80 pixels or better) face images, their performance degrades quite rapidly as the image quality drops (e.g. due to blurring, scaling, re-compression, etc.) causing significant degeneration of the visual attributes (Scheirer, Kumar, Iyer, Belhumeur, & Boult, 2013) they rely on.

We approach our face matching task as a face image retrieval (FIR) problem: given a query image, we aim at returning visually similar faces from a dynamically changing photo gallery, thus effectively reducing the search space from several thousands to about 20 likely candidates conveniently displayed on a single page. Our open-set approach practically out-rules person-specific training and uses the accuracy evaluation methods (e.g. top-N hit rate) that are typical of CBIR, rather than those typical of FR, e.g. Receiver Operating Characteristic (ROC), although the latter could also be used for compatibility reasons (Fawcett, 2006).

Our SIPP face image retrieval methodology has been deployed in a real-world face retrieval system (FaceMatch), detecting and matching faces in arbitrary orientation or lighting conditions. Handling large scale photo collections, our system requires no training while dealing with any open sets of images. FaceMatch R&D project attempts to solve most of the challenges mentioned, providing the following functionality:

• Semi-automatic annotation for faces and landmarks;
• Accurate face detection robust to scale and rotation;
• Image descriptors ensemble for improved face image match.

Figure 1. Unconstrained images may be challenging to modern face recognition systems
Facial Expression Analysis Using 3D Range Images

Development of Ultrasound Imaging CAD Tool for Assessment of Pathologies in Supraspinatus Tendon