FaceMatch: Visual Search by Photos of Missing Persons During a Disaster Event

Eugene Borovikov, Szilárd Vajda, Sema Candemir, Girish Lingappa, Sameer Antani, Michael Gill, George Thoma

Communications Engineering Branch, Lister Hill National Center for Biomedical Communications, U.S. National Library of Medicine, Bethesda, MD, 20894, USA

Introduction

We report on our FaceMatching research and development (R&D) that aims to provide robust image near-duplicate detection and face localization/matching on digital photos of variable quality, as an integral part of PEOPLE LOCATOR (PL) developed by NLM as a Web-based system for family reunification in cases of natural or man-made disasters. PL collects photos and brief text meta-data (name, age, etc.) of missing or found persons. Currently supported text queries may be insufficient because text data are often incomplete or inconsistent. Adding an image search capability can significantly benefit the user experience. Face localization is done via skin-tone/landmarks enhanced gray-scale face detector, more accurate than many open source and commercial detectors. Face matching is done via an ensemble of image descriptors (HAAR, LBPH, SIFT, SURF, ORB), using a smart re-ranking procedure. We describe the integration of our face matching system with PL, report on its performance. Unlike other face recognition systems often having many high quality well-illuminated sample images for each person, ours can handle the lack of training examples for individual faces, as those are unlikely in a disaster setting.

Near-Duplicate Detection

A reliable face detector is necessary for any face matching application, as it determines the locations and sizes of human faces in digital images. Our FaceFinder achieves this goal via

- Haar-like gray-scale features
- major 90-degree rotations
- skin color mapping in RGB, HSV, Lab spaces
- color based landmarks (eye, nose, mouth) detection
- artificial neural net (ANN) landmark verifier
- correcting minor rotations using eye line

Experiments

With no modifications, Viola-Jones face detector misplaces about half of the PL faces. About 20% of these are typically too small for matching. The data-sets we experimented with:

- HEPL-500: 500 images from Haiti
- Lehigh-512: 512 celebrities images
- Caltech-450: 450 Caltech faculty faces

Aided by skin mapping and landmark awareness, our FaceFinder outperforms some major commercial detectors (iOS, FaceSDK) and the leading open-source detectors by Viola-Jones and Zhu-Ramanan.

Improvements

- candidate list re-ranking based on MANY: $d = \sqrt{\sum d_i^2}$ with $d_i$ being the most confident (lowest) distance
- strong descriptors weigh more
- downstream weak matches via salience maps

Conclusion

Once the face/profile regions in the image collection are localized and their descriptors are indexed, they can be matched against a query face/profile picture, which may come from an existing (possibly annotated) image, or from a new photograph, that FaceMatch has not seen before. Hence the face matching module needs to be robust to accommodate wide variations in the appearance, and it needs to be fairly exact to eliminate numerous false positive hits.

We experimented with Caltech (450 color images) dataset and HEPL-372 (52 images with 8 synthetic modifications: crop, scale, and rotate). Accuracy (F-score) figures are reported in the table

<table>
<thead>
<tr>
<th>Method</th>
<th>HEPL-500</th>
<th>Lehigh-512</th>
<th>Caltech-450</th>
</tr>
</thead>
<tbody>
<tr>
<td>FaceFinder</td>
<td>0.77</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td>iOS</td>
<td>0.68</td>
<td>0.99</td>
<td>0.98</td>
</tr>
<tr>
<td>FaceSDK</td>
<td>0.73</td>
<td>0.94</td>
<td>0.91</td>
</tr>
<tr>
<td>Zhu-Ramanan</td>
<td>0.33</td>
<td>0.83</td>
<td>0.83</td>
</tr>
</tbody>
</table>

We provide query-by-image capability to the PEOPLE LOCATOR (PL) system, evaluated a few state-of-the-art systems on existing data-sets and developed tools for image annotation and near-duplicate detection. The face detection module improves a gray-scale face detector with the skin/landmark detection techniques. The face matching subsystem uses an ensemble of descriptors to capitalize on the strengths of its constituents, and results in higher accuracy figures than any of the individual descriptors.