FaceMatch: Visual Search for Pictures of Missing Persons During a Disaster Event

Eugene Borovikov, Szilárd Vajda, Girish Lingappa, Sameer Antani, Michael Gill, George R. Thoma

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

Abstract

Face detection determines the location and the size of a human face in an arbitrary digital image using:
- low-level image features (Haar, LBPP)
- high-level facial landmarks: eye, nose, mouth, ear, chin, etc.
- skin color

Multi-scale window technique
Ada boost classifier cascade (Viola-Jones)

Face matching subsystem uses
- Haar-like features
- multi-scale window technique
- Ada boost classifier cascade (Viola-Jones)

Near-Duplicate Detection

Near-Duplicate Experiments

Overall Challenges:
- pictures may contain 1 or more faces
- face-like objects (cats and dogs faces)
- query/database images may be of suboptimal quality due to:
  - partially occluded or damaged faces
  - presence of duplicates and near-duplicates
  - inconsistence due to facial hair, glasses, jewelry, aging

Method
- Haar wavelet based descriptor: most significant wavelet coeffs'
- real-valued distance measure in [0,1], with 0 = perfect match
- low threshold for near-duplicate detection
- champion selection: highest resolution

Number of near-duplicates in data-sets:
- HEPL: 6K near-dups in 15K images
- PL: 4K near-dups in 12K images

We have also experimented with generating about 800 near-duplicates from a set of 132 unique images by scaling \( s = 0.5 \) and \( s = 2 \), and cropping \( \frac{5}{6} \times \frac{5}{6} \). Our near-duplicate detector is most sensitive to rotations and cropping, detecting very few of those, while detecting most of the scaled near-duplicates correctly. This result was rather expected, given the Haar wavelet nature of the detector.

Advantages:
- high accuracy for skin detection (91%)
- skin maps assist the face finding
- enhance skin region intensities

Face Detection

Solution

Face Detection

With no modifications, Viola-Jones face detector misses about half of the HEPL faces, and about 20% of the missed ones are typically too small for matching.

Accuracy results of our FaceFinder:
- HEPL-380 (300 suitable images): R = 79%, P = 83%, F = 79%
- HEPL-4K (4,000 images including noise): F = 53%
- PL-700m (700 faces skipped by Viola-Jones) + skin map: 46% Recall boost, 22% False Positives rate

The major factors hurting the accuracy are:
- Lighting = 5.3%
- Low Quality = 8.8%
- Occlusion = 10.8%
- Color = 0.6%

Combination = 9.5%
Small Faces = 20.8%
Other = 44.2%

Face Matching

Near-Duplicate Experiments

Advantages:
- near-duplicate detector is most sensitive to rotations and cropping
- detecting most of the scaled near-duplicates correctly

Face Matching Experiments

PL: 4K near-dups in 12K images
HEPL: 6K near-dups in 15K images
HEPL-4K mod
HEPL-62mod

Accuracy (F-score) figures are reported in the table.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>HAAR</th>
<th>SIFT</th>
<th>SURF</th>
<th>ORB</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEPL-380</td>
<td>0.99</td>
<td>0.98</td>
<td>0.96</td>
<td>0.95</td>
</tr>
<tr>
<td>HEPL-4K mod</td>
<td>0.44</td>
<td>0.81</td>
<td>0.55</td>
<td>0.56</td>
</tr>
</tbody>
</table>

We have also experimented with combining the descriptor match distances by using a generalized geometric mean, and found that a combination of HAAR*ORB*SIFT*SURF produce a better F-score (by about 5%) than either of the individual descriptors. Inclusion of a weak descriptor tends to hurt the ensemble.

Conclusion

Having a goal to enable image based query capability in the People Locator (PL) system, we studied several image matching and face recognition methods, evaluated a few state-of-the-art systems on existing data-sets and developed core tools for image near-duplicate detection, face detection and face matching.

The near-duplicate image detector tool (Jacobs et al., 1999) helps the DLI administrator to clean up large group near-duplicate images in the PL database. The face detection capability relies on Viola-Jones object detection method, improved by the skin detection techniques. The face matching subsystem uses Haar, SIFT, SURF and ORB descriptors in an ensemble to capitalize on the strengths of its constituents and results in higher accuracy figures than any of the individual descriptors.