FaceMatch: a visual search system for family reunification during disasters

Eugene Borovikov, Michael C Bonifant, Girish Lingappa, Michael Gill, George Thoma
Communications Engineering Branch, Lister Hill National Center for Biomedical Communications, U.S. National Library of Medicine, Bethesda, MD, 20894, USA

PEOPLE LOCATOR®

Introduction

The objective of our system (FaceMatch) is to provide a fast and robust visual search capability for those who are looking for missing people in the aftermath of calamities, natural or otherwise. We propose a Single-Image-per-Person (SIP) approach to face image matching using an *ensemble* of weighted face image descriptors, taking advantage of their matching strengths. FaceMatch utilizes many visual features including color, texture and shape to provide robust face *detection*, *matching* and image *near-duplicate detection* capabilities. To test our system, we considered several well-known benchmark face image collections. Our results are comparable to, and in some cases, better than those of leading open source and commercial systems, especially for low quality images, which is usually the case in disaster scenarios. FaceMatch is integrated in a real-world family reunification system called Lost Person Finder (LPF). It has been used during several hospital drills and actual disaster events, providing an efficient visual search modality; extending the text-based search facility; and benefiting hospitals, national/international organizations, and other entities closely involved in handling such events.

Challenges

- low quality, suboptimal lighting
- pictures may contain 0 or more faces
- face-like objects (animal/cartoon faces)
- presence of duplicates and near-duplicates
- face images may be hard to match due to
  - partially occluded or damaged faces
  - presence of facial hair, glasses, jewelry
  - person natural aging
  - source photograph degradation

Near-Duplicate Detection

Description

An image data-set may contain many near-duplicate images due to multiple postings of the same photograph rescaled or re-compressed. Such near-duplicates need to be identified and grouped. Each group would be represented by the highest quality image. We solve this by

- color wavelet descriptor: scale- and compress-robust
- real-valued distance measure in $[0, 1]$, with $0 =$ perfect match
- tight threshold for *near*-duplicate detection
- champion selection: highest resolution, lower compression
- using 128×128 YIQ color images: gray-scale compatible

Experiments

Detect near-duplicate images in our data

data-set		near-duplicates		
name	size	# of	% of	proc.time
HEPL	15K	6K	40	5 min
PL	12K	4K	30	4 min

Image matching on generated near-dups

distortion	Recall	Precision	F-score
rotation	0.69	0.62	0.65
crop	0.71	0.70	0.71
scale	0.99	0.99	0.99

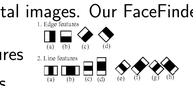
We have also experimented with generating 792 near-duplicates from a set of 132 unique images by scaling ($s = 0.5, 2$), rotating ($\alpha = \pm\pi/12$) and cropping ($c = 0.8, 0.65$). Our near-duplicate detector is most sensitive to rotations and cropping, detecting very few of those, while detecting most of the scaled near-duplicates correctly. This result was rather expected, given the Haar wavelet nature of the detector.

Face Detection

Description

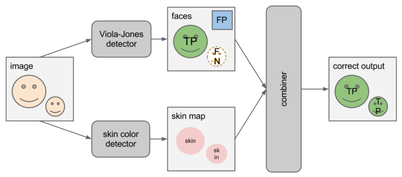
A reliable *face detector* is necessary for any face matching application, as it determines the locations and sizes of human faces in digital images. Our FaceFinder achieves this goal via

- Haar-like gray-scale features
- major 30-degree rotations



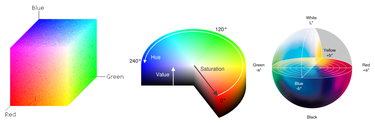
- skin color mapping in RGB, HSV, Lab spaces
- color based landmarks (eye, nose, mouth) detection
- artificial neural net (ANN) landmark verifier
- correcting minor rotations using eye line

ViolaJones+SkinMap+Landmarks



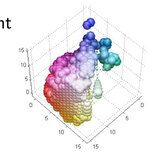
skin mapping

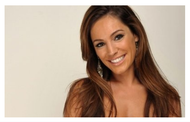
- skin color color from annotation
- estimate color models: ANN or histogram
- extended color space: [RGB,HSV,Lab]

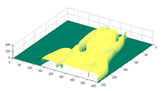


combiner module

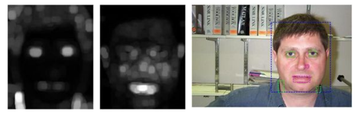
- recover false negatives (FN)
  - skinmap driven enhancement
  - color landmark detection
- reject false positives (FP)
  - skinmap region integration
  - landmark positioning
- GPU based acceleration







landmark detection



- eye/mouth maps derived from luma/chroma bands
- major peaks are landmark candidates
- false positives eliminated by ANN landmark verifier

Experiments

Base Viola-Jones face detector misses about 50% of PL faces. The data-sets we experimented with:

HEPL-500 : 500 images from Haiti earthquake 2010

Lehigh-512: 512 images of celebrities

FDDB-2K : 2845 images with 5171 faces

FaceFinder (FF) outperforms some commercial detectors (iOS, FaceSDK) and leading open-source detectors (e.g. Zhu-Ramanan) by correcting (right) its baseline (left)

set	method	R	P	F
HEPL-500	ViolaJones	0.76	0.87	0.81
	FaceFinder	**0.77**	0.89	**0.83**
	iOS	0.68	0.87	0.76
	FaceSDK	0.73	0.87	0.79
	Zhu-Ramanan	0.33	**0.92**	0.49
Lehigh-512	ViolaJones	**0.95**	0.81	0.88
	FaceFinder	**0.95**	**0.94**	**0.95**
	iOS	**0.95**	0.92	0.94
	FaceSDK	0.93	0.91	0.92
	Zhu-Ramanan	0.83	0.91	0.87
FDDB-2K	ViolaJones	0.60	0.80	0.69
	FaceFinder	**0.74**	**0.88**	**0.80**
	iOS	0.63	0.76	0.69
	FaceSDK	0.64	0.85	0.73
	Zhu-Ramanan	0.61	0.79	0.69

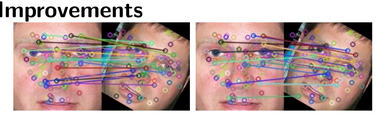
Face Matching

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 face matcher has not seen before. FaceMatch is robust to accommodate wide variations in the appearance, and it is precise to eliminate many false positives.

Solution

- pre-localized face/profile and landmarks
- ensemble approach to *trainingless matching*
- HAAR/SIFT/SURF/ORB/RSILC descriptors
- features weights by confidence
- scale/rotation invariant metrics
- FLANN indexing for better scaling
- real-valued distance in $[0, 1]$
  - $0 =$ perfect match
  - $1 =$ complete mismatch

Improvements

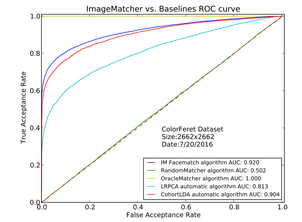


- RANSAC and MEADOW key-spot filtering
- candidate list re-ranking based on
  DIST : $d = \prod d_i^{w_i}$, with $d_i, w_i \in [0, 1]$
  RANK : Borda count weighted re-ranking
- stronger descriptors weigh more
- downplay weak matches via salience maps

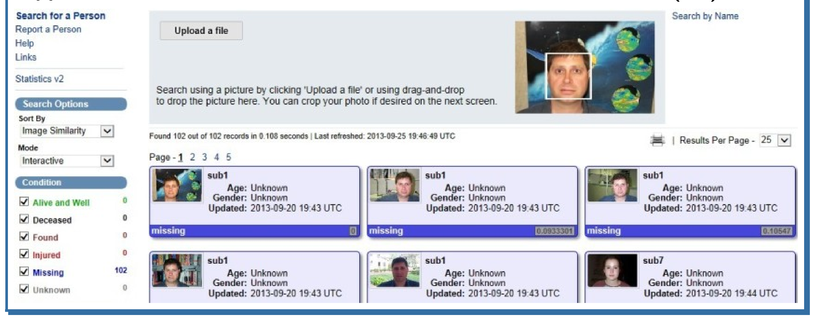
Experiments

Data-sets: Caltech (450 color images), ColorFERET (2662 color images), IndianFaces (676 color images). Compared with commercial and open systems. Accuracy as top-N hit rates.

	CalTech		ColorFERET		IndianFaces	
top-N	FSDK	FM	FSDK	FM	FSDK	FM
1	.98	.98	.74	.98	.69	.79
5	.99	.99	.75	.99	.76	.87
10	.99	.99	.76	.99	.79	.90
20	.99	**1.0**	.76	**1.0**	.83	**.92**



Application: FaceMatch web services for PEOPLE LOCATOR (PL)®



Conclusion

We provided query-by-image capability to the PEOPLE LOCATOR (PL)® system, evaluated several 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 than modern commercial and and open-source FR systems.

NIH U.S. National Library of Medicine

2016 NIH Research Festival, Bethesda, MD, USA

NATIONAL LIBRARY OF MEDICINE