FaceMatch: real-world face image retrieval

Eugene Borovikov and Szilárd Vajda

U.S. National Library of Medicine
8600 Rockville Pike, Bethesda, MD 20894
FaceMatch@NIH.gov

Abstract. We researched and developed a practical methodology for face and image retrieval (FIR) based on optimally weighted image descriptor ensemble. We describe a single-image-per-person (SIPP) face image retrieval system for real-world applications that include large photo collection search, person location in disaster scenarios, semi-automatic image data annotation, etc. Our system provides efficient means for face detection, matching and annotation, working with unconstrained digital photos of variable quality, requiring no time-consuming training, yet showing a commercial performance level at its sub-tasks. Our system benefits public by providing practical FIR technology, annotated image data and web-services to a real-world family reunification system.

Keywords: face recognition, image retrieval, family reunification

1 Introduction

Recent advances in content based image retrieval (CBIR) technology have produced many meaningful image-based web-scale search techniques[10], and several web search engines (e.g. google.com/insidesearch, bing.com/images, yandex.com/images) now provide such capabilities. The recent decade has also seen a considerable progress in the face recognition (FR) technology, in some cases approaching human-level accuracy in face detection and verification tasks, especially in the controlled environments [24,30,33].

Fig. 1. Unconstrained images present challenges to face recognition systems.

The modern web-based FR solutions (e.g. facebook.com, plus.google.com) work well in limited users circles that tend to contain tagged pictures of the same few individuals (e.g. family and friends) with multiple shots per person, which allows for recognition model training. There are still very few publicly available *single image per person* (SIPP) training-less face image retrieval systems that can work effectively with millions of faces pictured in unconstrained settings, presenting many challenges for such systems in practice, e.g. disaster recovery:

- data-set size: millions of photos, many near-duplicates[1]
- no constraints on uploaded or query pictures, as in Fig. 1
- often suboptimal quality query and database images
- inconsistency in query/gallery face appearance.

Many of those challenges are being addressed by the modern FR systems thanks to the emergence of labeled datasets with unconstrained images [7,8,14] utilized for various competitions.

Typical FR systems would approach the face recognition problem in one of the two formulations[34]: *verification* (photos depict the same person), and *identification* (pick the closest in appearance pictures to the query image). Such systems usually require some form of model training, using multiple photos per individual. They would typically work with a set of visual features extracted from images, imposing (or learning) a measure of visual proximity, modeling human visual perception of faces. While modern automatic face classification/verification methods can work fairly well on good quality face images (fairly well lit, sharp, 80×80 pixels or better), their performance degrades quite rapidly as the image quality drops (e.g. due to burring, scaling, re-compression, etc.) causing a significant degeneration of the visual attributes[27] they rely on.

We formulate our face matching problem as a *face image retrieval* (FIR) problem: given a query photo, return visually similar faces (preferably of the same individual) from a dynamically changing photo collection, thus efficiently reducing the user's search space from many thousands to just tens of likely candidates. This dynamic, open set approach essentially demands the method to be *training-less* and uses the accuracy evaluation methods that are more typical of CBIR (e.g. top-N hit-rate, defined in Section 3.2), rather than those used in FR (e.g. receiver operating characteristic[12], ROC).

We implement our SIPP face image retrieval methodology in a real-world *face retrieval* system (FaceMatch), putting no restrictions on input images, detecting and matching faces in arbitrary poses or lighting conditions. Handling large scale photo collections, our system requires no training while dealing with any open sets of images. Our FaceMatch R&D effort addresses many of the mentioned challenges and provides 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.

We present FaceMatch evaluation results for the face detection, matching and retrieval tasks, using several publicly available data sets, some of which were

[1] visually almost identical

annotated at our laboratory. The resulting image retrieval system is geared towards face detection and matching, but it can be used for generic object/scene detection and matching, providing a rich set of tools for practical large-scale image collection management.

In what follows, we discuss our data repository and present the major components of our FaceMatch (FM) system. In each section, we review the relevant publications, describe our approach, and present our experimental results.

2 Image data collections

Our approach to face image processing and retrieval is data driven, automatically extracting and weighting features from the data, based on statistics. Annotated image repositories provide ground truth (GT) for the accuracy evaluation and optimization of individual components, e.g. skin mapping for face localization.

Image annotations for *face detection* typically consists of localized face regions (and optionally face landmarks: eyes, nose, mouth, and ears), optional gender and age groups, and some skin patches. Such annotations are done semi-automatically, providing the human annotator with initial face/landmark localization, which can be manually corrected or completed.

Ground truth for *face matching and retrieval* involves labeling face images with face/person ID[2] that are used to assess the quality of retrieval accuracy. Our system targets unconstrained image data-sets, e.g. photos from natural disaster events collected by People Locator (PL). PL data-set consists of 40 thousand weakly text-labeled mostly color, low quality images, some of which are shown in Fig. 1. PL image repository is changing over time, as disasters happen[31].

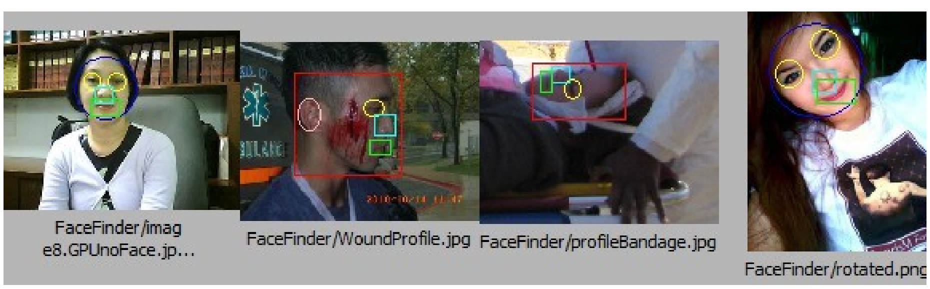


Fig. 2. Face and landmarks annotation examples

To help organize PL repository, we have developed several cross-platform image processing and annotation tools to

- reduce data by removing near-duplicates,
- outline faces, profiles as rectangular regions,

[2]unique alpha-numerical sequence, *not* revealing the true person identity

– localize facial features: eyes, nose, mouth, and ears,
– extract skin patches from the skin-exposed regions.

These tools were used to partially annotate various image collections with the correct face/profile locations and facial landmarks (eyes, nose, mouth, and ear), as shown in Fig. 2. Using our web-based and desktop annotation tools, our team annotated several thousand PL images, producing:

PL-Faces consists of 2882 low resolution, color PL images, with 3/4 of face regions being frontal and about 1/4 are profile views. The average face and profile diameters are 40 and 50 pixels respectively.

HEPL-500 is a subset of PL containing 500 images from 2011 Haiti earthquake, containing a large variety of faces. Some of them are over-exposed, blurry or occluded as shown in Fig. 1.

The images were selected to include a large range of skin tones, environments, cameras, resolutions, lighting conditions. Some of the images contain multiple human subjects. The quality of the images varies significantly in illumination, resolution and sharpness.

The annotated PL data-sets are freely available for research purposes. Additional meta-data annotation (e.g. ethnicity, age/group, gender) is also available for some sets. The annotated repository is regularly updated and used for improving face detection and matching performance. We also utilize some publicly available datasets depicting humans in unconstrained environments for algorithm evaluation and tuning:

CalTech Faces set [1] consists of 450 frontal views of 29 subjects, which are taken under varying lighting and background conditions.

Indian Faces set [19] contains 676 face images of 61 individuals (male and female), shot in a studio, exhibiting large variations in head pose, face expression, and lighting.

ColorFERET set [2] contains 2413 facial images of 856 individuals showing frontal and left/right profile head pose variation, optional glasses, and various facial expressions.

Face Detection Data Set and Benchmark (FDDB) set [18] contains 2845 images with 5171 unconstrained faces.

Lehigh Faces set C1 contains 512 images obtained trough our collaboration with Lehigh University[20] containing unconstrained images of celebrities, exhibiting wide variations in background and pose, with mostly light skin tones. Set C2 is similar to C1, containing 550 images, but with a greater variety in faces and their sizes.

For some of the mentioned sets (e.g. CalTech and Indian Faces) we have provided landmark annotations in addition to the supplied head/face regions. Our experiments use those sets to test FaceMatch performance, and the evaluation results are presented in the respective sections.

3 Face matching

A typical face recognition (FR) system addresses the problem of face matching in one of two formulations: *verification* $1 : 1$ (is the same person depicted in two photos) or *identification* $1 : N$ (find the depicted person in a fixed set of enrolled faces). Our FaceMatch (FM) system understands *face matching* as *single image per person* (SIPP) face retrieval approach utilized for interactive-time searches in large dynamic collections of face images shot in unconstrained environments, i.e. arbitrary resolution, scale, illumination, etc. Thus it is different from *verification* (as our decision is not binary) or *identification* (our image sets are dynamic).

3.1 Background

In the last couple of decades, the face recognition (FR) community has considerably advanced the filed and produced a large number of great papers. Here, we review the research that is most relevant to our approach, describing the methods we drew upon and utilized in the implementation of our FaceMatch.

Face recognition (FR) in general conditions remains to be an open problem that's being researched actively [4,16,28]. Beham[6] gives a good overview of FR techniques and divides them in the following major groups (holistic, feature-based, and soft-computing), providing normalized accuracy (NA) figures, pointing out their advantages and drawbacks.

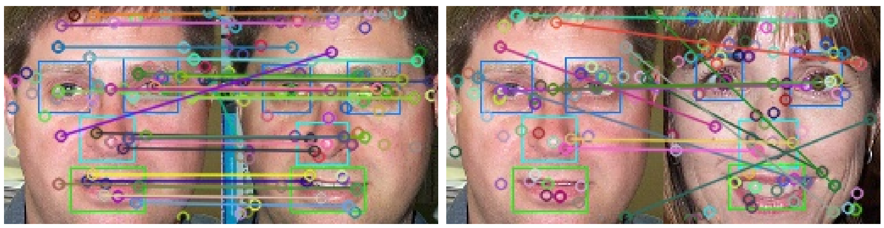
Unconstrained, *single image per person* (SIPP) face retrieval from a large, dynamically changing (open-set) reference gallery basically requires its face matching to be training-less, robust to pose, occlusion, expression, lighting, and fast, i.e. essentially modeling human perception of unfamiliar faces from a single photo and utilizing some fast approximate indexing for efficiency.

Several very promising methods [13,30,32,17] have been proposed over the past decade, and more recent papers describe systems that are as accurate as a human[29] at the face verification task or sometimes even better[22]. This kind of accuracy typically implies (deep) learning systems with a substantial training stage using hundreds or thousands shots per person, and their matching time may still be not very practical for large scale interactive searches.

Wolf et al.[32] presented an interesting approach to face matching called the one-shot similarity kernel, using a special similarity measure to produce some impressive face matching results on Labeled Faces in the Wild (LFW) collection[14]. We cannot utilize this approach directly, as it requires some training with the background examples.

3.2 Face image retrieval

Given a dynamically changing repository of images, we propose a methodology for scalable visual search, effectively solving the face image retrieval problem. Face matching queries can be performed after the face/profile regions in the image collection are localized and their descriptors are indexed. The proposed method accommodates wide variations in face appearance mentioned in section 1. Given a query face image, the goal is to match it against the repository of the existing face descriptors, and output a list of likely face candidates ordered by similarity. The matching technique cannot assume that many faces of the same subject are present in the database, and it needs to be robust to illumination, scale and affine transformations.



(a) same person, different photos (b) different faces

Fig. 3. SIFT based matching performance of the system on two example faces: observing more correct correspondences for the same person

Among the training-less single-descriptor face matching methods, we decided to focus on rotation and scale invariant key-spot descriptor based matching (e.g. SIFT[21], SURF[5], and ORB[26]), compare them with the holistic descriptors (color HAAR[15] and LBPH[3]), and consider an weighted ensemble of them.

Fig. 3 presents two unrestricted key-spot matching examples with SIFT descriptors. The left pair shows matches between two different photos of the same person: the number of correctly matched locations is relatively high. The right pair shows the faces that belong to different people: there are evidently fewer sensible matching locations, e.g. note the non-matched key-spots at the chin location of the faces. Experiments with several datasets revealed that

- single descriptor is insufficient for accurate retrieval,
- some key-spot matches need to be filtered as outliers,
- face landmarks help filter and weigh the matches.

Having several image descriptors per face (HAAR, SURF, SIFT, ORB, LBPH), we experimented with similarity distance-based and similarity rank-based feature combination strategies. The combinations used individual distances $d_i \in [0, 1]$ (or ranks) and descriptor matching confidence weights $w_i \in [0, 1]$:

DIST: *weighted distance product* $d = \prod {d_i}^{w_i}$
RANK: *rank-based* combination based on Borda count[11].

Evidently, the *decreasing confidence radical* is a particular case of the *weighted distance product*, but it skips the need to specify the weights explicitly, and uses the inverse confidence ordering of the descriptors to compute its distance combination efficiently. The weighted descriptor ensemble hence allows:

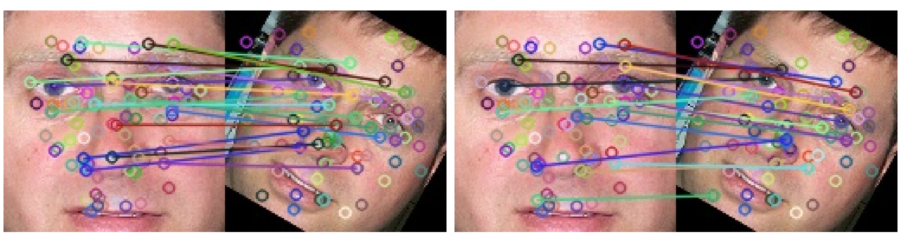
- combination of holistic with the key-spot based image descriptors,
- utilization of color along with the texture information,
- optimal descriptor weighting procedure according to the matching confidence

The optimization procedures are performed using the non-linear simplex[25] method maximizing the retrieval accuracy expressed as F-score or *hit rate*, i.e. the frequency of retrieving the correct subject given a probe photo in a top-N query, i.e. for a set of query images $\mathcal{Q}$, define the hit rate for top-N matches as

$$HitRate(N) = HitCount(N, \mathcal{Q})/|\mathcal{Q}|, \tag{1}$$

where $HitCount(,)$ is a function that counts the successful top-N matches using the query set of size $|\mathcal{Q}|$.

Boosting key-spot matching accuracy As Fig. 3 suggests, there may be some key-spot mis-matches, that may in turn cause some false hits in face image queries. To improve matching confidence, our key-spot descriptor matching scheme includes the descriptor symmetric match *cross-check* to ensure that best match relationship works both ways. Our matching scheme also removes the descriptor matches whose distance is greater than two minimum distances across the matching pool, but still can produce some false hits.



(a) RANSAC scale and rotation (b) MEADOW scale and rotation

Fig. 4. Spurious SURF key-spot match filtering to ensure geometric consistency

To further imporove the key-spot descriptor matching accuracy, we filter out the outliers among the two-way descriptor matches via the inter-view homography[9] based RANdom SAmple Consensus (RANSAC) algorithm [35]. This iterative statistical method computes and uses an affine transform between two images (homography) of the same (or similar) object to assert the key-spot consensus. It works quite well for the near-frontal views of in-plane rotated and scaled faces, as shown in Fig. 4(a), but it may slow down the face matching process because of is iterative nature and having to estimate the homography matrix at each iteration.

Increasing key-spot matching speed As quicker alternative to RANSAC, we researched and developed MEdian-based Anomalous Distance Outliers Weeding (MEADOW) method. As the name suggests, the method weeds out the key-spot outliers, i.e. matches with too unlikely geometric distances between the corresponding key-spots. Compared to RANSAC, MEADOW is intended to be

- more efficient: no iterative estimation of homography
- less constrained: no key-spot co-planarity assumption

MEADOW is expected to be less accurate than RANSAC in general, but for practical face image matching applications, their accuracies are comparable.

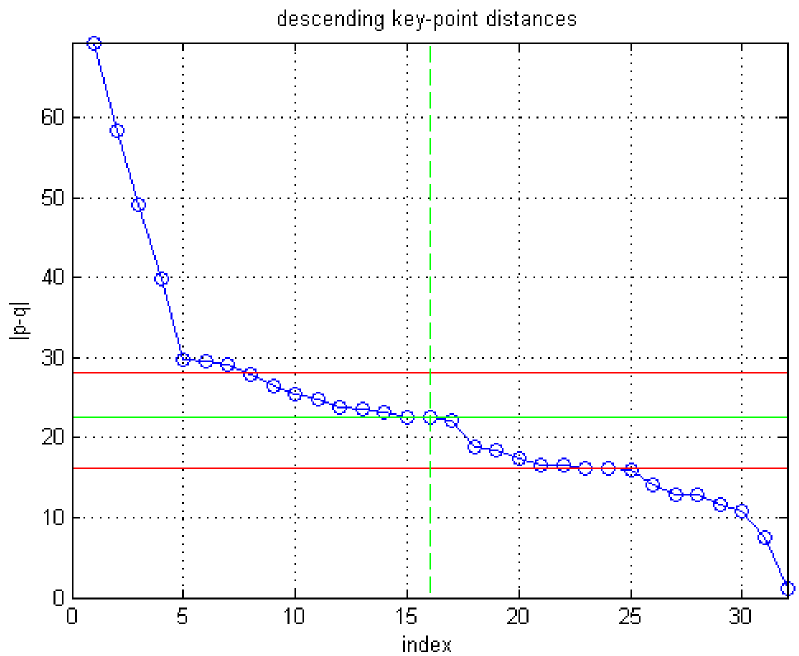


Fig. 5. MEADOW filters distance outliers above and below the median deviation lines (red) with respect to the sample's median (green).

In MEADOW, for each two-way descriptor match, we compute the Euclidean distance between their key-points (not descriptors) p and q, and we discard that match as a false positive, if that distance $D = |p - q|$ is an outlier among all the distances in the match sample: $|D - M| > T$, as shown in Fig. 5, where M is the sample's distance median (dashed green lines), and T is computed as a median deviation from M. MEADOW is a simpler (than RANSAC) method for filtering out the largest outliers from a sample, which is what we intend for the key-spot distances to ensure the key-spot geometric consistency. As we can see in Fig. 4(b), MEADOW efficiently handles the outliers, filtering out most of the false matches, typically five times faster than RANSAC, resulting in a similar matching accuracy.

Distance normalization The *cumulative difference* between two face images is computed as a median (also to be robust to the outliers) among their closest descriptor matches:

$$\hat{d}(f,g) = \mu_{i=1,\ldots,I}\{\min_{j=1,\ldots,J} \tilde{d}(s_i(F), s_j(G))\} \tag{2}$$

where μ stands for the *median*, $\tilde{d}$ is the matching distance between descriptors $s_i(F)$ and $s_j(G)$ computed for their images F and G, with the (filtered) descriptor match counts I and J respectively.

Since a difference measure (e.g. $\hat{d}$) may range in $[0,\infty)$, it can be *normalized* to $[0,1)$ via the monotone increasing and smooth arctangent mapping

$$d(x) = \arctan(ax)/a \tag{3}$$

where $a = \frac{\pi}{2}$, which maps an infinite range to a unit, and behaves quite linearly near 0, having $d(0) = 0$ and $d'(0) = 1$.

For easier distance thresholding, one can scale the normalized distance to satisfy some perceptual similarity constraints, depending on the descriptor and the image set $\delta(x) = d(\alpha x)$ by picking $\alpha > 0$ such that $\delta(x) \leq 0.5$ for the similar faces, and $\delta(x) > 0.5$ for the dissimilar ones, but that would involve some human judgment and semi-manual grouping of similar faces. Clearly, this monotone distance normalization approach applies to any distance measure.

Descriptor search space partitioning While dealing with large unconstrained face image datasets (over 40K images), our system, to be practical, needs to retrieve face images within interactive (about 1 second) turn-around time intervals. To accomplish that we researched and developed the *attribute bucketing* strategy and utilized the *approximate nearest neighbor* (FLANN) searches[23].

We have noticed that our image typically carry gender and age-group meta-information, which allowed us to partition the search space into a number of age and gender groups (called buckets), which we could query in parallel using multi-threading. This allowed us to optimize our query turn-around times by a factor of 9 or more, especially when we introduced sub-bucketing within groups.

Utilization of FLANN resulted in the additional (five-fold on average) queries speed-up with a small penalty (a couple of percentage points) to the retrieval accuracy and a small one-time clustering overhead during the index loading and incremental update. Overall, the face image query turn-around times are kept under a second for our image data-sets. Provided enough of the multi-core power it should be scalable to the web-scale sets of millions of images.

3.3 Experiments

Due to the sources of our target image collections, we very rarely have more than one picture of the same person. Hence, in our face retrieval evaluations, we had to rely on a mixture of datasets, e.g. the CalTech Faces data mixed with some typical PL photos, described in Section 2.

Color-aware face matching For the color-aware face matching experiments, we considered IndianFacesDB and ColorFERET datasets (described in Section 2), containing color images of male and female faces with good variations in lighting, pose, and expression.

Table 1. Color-aware face matching top-1 hit rates

	IndianFaces		ColorFERET	
descriptor	alone	+ CW	alone	+ CW
CW	0.52	-	0.78	-
SIFT	0.61	**0.66**	0.91	**0.95**
SURF	0.75	**0.78**	0.96	**0.98**
SURF+SIFT	0.76	**0.79**	0.97	**0.98**

We observe that our CW descriptor alone is a weaker matcher than any of the key-point based descriptors, but it considerably improves the query hit rates, when included in the ensemble with the stronger (but color-blind) descriptors. This behavior suggests that bringing color-awareness to the descriptor ensemble helps improve the face matching performance on color images.

Descriptor ensemble matching For the FaceMatch overall visual feature ensemble (with optimally weighted descriptors), the top-N hit rate accuracy results on the available benchmark datasets are summarized in Table 2 in comparison with the commercial face matching engine FaceSDK.

Table 2. FaceSDK (FSDK) vs. FaceMatch (FM) hit rate accuracy in top-N queries

	CalTech		ColorFERET		IndianFacesDB	
top-N	FSDK	FM	FSDK	FM	FSDK	FM
1	.98	.98	.74	.98	.69	.79
3	.99	.98	.75	.98	.73	.85
5	.99	.99	.75	.99	.76	.87
10	.99	.99	.76	.99	.79	.90
20	.99	**1.0**	.76	**1.0**	.83	**.92**

On the relatively easy CalTech dataset (with large, mostly frontal faces), accuracy figures of both FaceMatch and FaceSDK are predictably high and close to each other. On the more challenging (than CalTech) ColorFERET benchmark dataset with considerable variations in head pose and lighting, FaceSDK clearly yields to FaceMatch, which performs just as well as it does on CalTech, reaching the statistically guaranteed retrieval of the correct person within top 20 retrieved records. The accuracy on even more challenging (than CalTech or ColorFERET) IndianFacesDB dataset is noticeably lower for both competitors probably due to some extreme head pose variations, but FaceMatch clearly outperforms FaceSDK, providing the 92% likelihood of retrieving the right person in top 20 visual query results.

4 System

Our system is cross-platform and production-level. The core FaceMatch (FM) imaging code is written in C++. It relies on open source libraries (e.g. STL, OpenCV, OpenMP) and is packaged as a shared library. This makes it deployable for desktop applications or over as web services. The key focus during the web integration was to ensure the top performance across all FaceMatch operations, e.g. list, ingest, query and remove. Our design takes advantage of multi-core architectures by exploiting task level and functional parallelism inside all critical modules. For instance, the web service can answer multiple queries while ingesting or removing descriptors.

The FaceMatch (FM) services are currently utilized in a real-world family re-unification system, which adds a visual modality to the otherwise text based searches. The user is free to browse the database by inspecting the details of the retrieved records and optionally re-submitting queries using the retrieved faces as examples. The output of the FaceMatch module can be optionally fused with the text query results for an increased query accuracy.

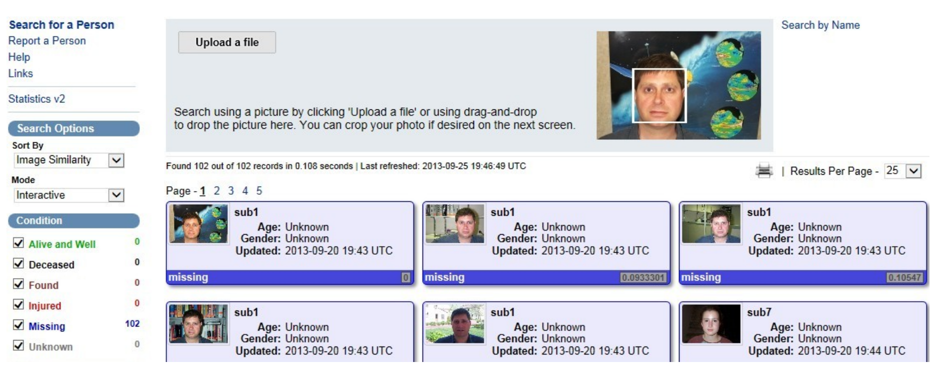


Fig. 6. FaceMatch sample visual query results on the CalTech+PL data

A sample visual query results are shown in Fig. 6, and we can see how the system retrieves the faces similar to the query in the ascending distance-to-query order, observing the same person photos being at the top of the result set.

5 Summary

Targeting a practical system handling web-scale photo collections with real-world images, we researched and developed a *single-image-per-person* (SIPP) query-by-photo methodology (FaceMatch) working with unconstrained images of variable quality, implemented it as a cross-platform software library, exposing its face image retrieval functionality via web-services, which can be consumed by real-world applications, such as efficient photo collection search for the disasters management, missing children location, and law enforcement organizations.

5.1 Applications

FaceMatch services currently benefit public by providing its SIPP face image query capability to a real-world family re-unification system, which includes a set of mobile applications and a website accepting missing/found people records and answering (visual) queries about them. Since text-based records may be inexact or incomplete, queries by photo greatly improve the user experience and provide the visual modality reducing the search space.

With tens of thousands records in a real-world collection, FaceMatch can reduce the user browsing set of most likely candidates to about 20 with the user-friendly query turn-around time of about a second. FaceMatch is quite robust to *cross-cultural face matching*, retrieving visually matching records much faster (and in some cases more accurately) than a human emergency coordinator under stress in case of an emergency. This helps save time and effort for the disaster event managers, health emergency coordinators and people who search for their missing relatives.

FaceMatch also provides its robust face detection and general image matching services, which have been successfully utilized at detecting pictures with faces in a multi-million document collection of medical documents. Our rapid image search services are utilized in identifying visual near-duplicates, efficiently reducing image collections, e.g. by 40% for PL, thus making the system more user-friendly by nearly halving the query turn-around time and presenting non-duplicate results.

5.2 Methodology

We researched and developed several image matching and face recognition methods, evaluated a few state-of-the-art systems on available datasets, developed a software library for: (i) image near-duplicate detection, (ii) general image queries, (iii) robust face detection, (iv) efficient face matching. The major features that make FaceMatch practical for the real-world face image retrieval:

- *unconstrained* images handling,
- *training-less* single-image-per-person (SIPP) approach,
- *cross-platform* approach to the implementation.

Our technology matches the performance of the leading open-source and commercial solutions. We have made several important improvements to the existing methods and developed some new ones:

Face detection was improved by using human skin tone information and facial landmarks along with default (color-blind) face detection algorithm. The skin regions are mapped using an artificial neural network (ANN). On public data sets, our face detector was more accurate than the available state-of-the-art engines, both commercial and open-source.

Face matching utilized a SIPP approach using weighted image descriptor ensemble to optimize the matching accuracy without training. Our MEADOW key-point filtering, attribute bucketing and FLANN indexing helped speed-up queries up to 20-times (compared to the linear search), keeping turn-around time within one second for a typical real-world collection.

We have annotated thousands of face images in the PL dataset with face, profile and landmark regions. The annotated datasets are public domain and can be made available upon request.

5.3 Prospects

We plan to expand the number of applications for our public FaceMatch services. Those may include visual search by photo for missing children, pets, criminal suspects, as well as detecting disaster and crime scenes. From the technical prospective, our R&D team is actively engaged in research and development that lead to (i) robust gender, age and ethnicity estimation, and (ii) general object and animal detection/matching.

We are currently researching the *human-in-the-loop* (HiL) approach for naturally merging face image retrieval with annotation, making both more efficient via semi-supervised and incremental machine learning techniques as well as via more natural human-computer interactions, which may include the development of more convenient game-like visual annotation tools, and use of crowd-sourcing for developing more comprehensive testing and evaluations data sets, including video, because mobile technology tends to generate an increasing amount of moving pictures often with characteristic audio tracks, quite useful for practical face and object image retrieval.

Acknowledgments. This research is supported by the Intramural Research Program of the National Library of Medicine at National Institutes of Health. We thank our colleagues from Lister Hill National Center for Biomedical Communications (LHNCBC) Michael C. Bonifant, Girish Lingappa, Sema Candemir, Le Lan and Tehseen Sabeer for their technical assistance during the experiments. Our special thanks to Michael Gill, Sameer Antani and George Thoma for their valuable advice and general guidance.

References

1. Caltech frontal face dataset. http://www.vision.caltech.edu/archive.html
2. The color FERET database. http://www.nist.gov/itl/iad/ig/colorferet.cfm
3. Ahonen, T., Hadid, A., Pietikäinen, M.: Face recognition with local binary patterns. In: Proceedings of the European Conference on Computer Vision, pp. 469–481. Springer (2004)
4. Azeem, A., Sharif, M., Raza, M., Murtaza, M.: A survey: face recognition techniques under partial occlusion. Int. Arab J. Inf. Technol. 11(1), 1–10 (2014)
5. Bay, H., Tuytelaars, T., Gool, L.V.: SURF: Speeded up robust features. In: European Conference on Computer Vision. pp. 404–417 (2006)
6. Beham, M.P., Roomi, S.M.M.: A review of face recognition methods. International Journal of Pattern Recognition and Artificial Intelligence 27(04), 1356005 (2013), http://www.worldscientific.com/doi/abs/10.1142/S0218001413560053
7. Beveridge, J., Phillips, P., Bolme, D., Draper, B., Givens, G., Lui, Y.M., Teli, M., Zhang, H., Scruggs, W., Bowyer, K., Flynn, P., Cheng, S.: The challenge of face recognition from digital point-and-shoot cameras. In: Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on. pp. 1–8 (Sept 2013)
8. Borovikov, E., Vajda, S., Lingappa, G., Antani, S., Thoma, G.: Face matching for post-disaster family reunification. In: IEEE International Conference on Healthcare Informatics. pp. 131–140 (Sept 2013)
9. Chum, O., Pajdla, T., Sturm, P.: The geometric error for homographies. Computer Vision and Image Understanding 97(1), 86 – 102 (2005), http://www.sciencedirect.com/science/article/pii/S1077314204000244
10. Dharani, T., Aroquiaraj, I.: A survey on content based image retrieval. In: Pattern Recognition, Informatics and Mobile Engineering (PRIME), 2013 International Conference on. pp. 485–490 (Feb 2013)
11. van Erp, M., Schomaker, L.: Variants of the borda count method for combining ranked classifier hypotheses. In: 7th International Workshop on Frontiers in Handwriting Recognition. pp. 443–452 (2000)
12. Fawcett, T.: An introduction to ROC analysis. Pattern Recognition Letters 27(8), 861 – 874 (2006), http://www.sciencedirect.com/science/article/pii/S016786550500303X, ROC Analysis in Pattern Recognition
13. Gao, Y., Qi, Y.: Robust visual similarity retrieval in single model face databases. Pattern Recogn. 38(7), 1009–1020 (Jul 2005)
14. Huang, G.B., Ramesh, M., Berg, T., Learned-Miller, E.: Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Tech. rep., University of Massachusetts, Amherst (2007)
15. Jacobs, C.E., Finkelstein, A., Salesin, D.H.: Fast multiresolution image querying. In: Proceedings of the 22nd annual conference on Computer graphics and interactive techniques. pp. 277–286. ACM, New York, NY, USA (1995)
16. Jafri, R., Arabnia, H.R.: A survey of face recognition techniques. JiPS 5(2), 41–68 (2009)
17. Jain, A., Klare, B., Park, U.: Face matching and retrieval in forensics applications. MultiMedia, IEEE 19(1), 20–20 (Jan 2012)
18. Jain, V., Learned-Miller, E.: Fddb: A benchmark for face detection in unconstrained settings. Tech. Rep. UM-CS-2010-009, University of Massachusetts, Amherst (2010)

19. Jain, V., Mukherjee, A.: The indian face database (2002), http://vis-www.cs.umass.edu/~vidit/IndianFaceDatabase/
20. Kim, E., Huang, X., Heflin, J.: Finding VIPs - a visual image persons search using a content property reasoner and web ontology. In: Multimedia and Expo (ICME), 2011 IEEE International Conference on. pp. 1 –7 (july 2011)
21. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60(2), 91–110 (2004)
22. Lu, C., Tang, X.: Surpassing human-level face verification performance on LFW with GaussianFace. CoRR abs/1404.3840 (2014)
23. Muja, M., Lowe, D.G.: Fast approximate nearest neighbors with automatic algorithm configuration. In: International Conference on Computer Vision Theory and Applications. pp. 331–340 (2009)
24. Naruniec, J.: A survey on facial features detection. International Journal of Electronics and Telecommunications 56(3), 267–272 (2010)
25. Nelder, J.A., Mead, R.: A simplex method for function minimization. The computer journal 7(4), 308–313 (1965)
26. Rublee, E., Rabaud, V., Konolige, K., Bradski, G.: ORB: An efficient alternative to SIFT or SURF. In: IEEE International Conference on Computer Vision. pp. 2564 –2571 (November 2011)
27. Scheirer, W.J., Kumar, N., Iyer, V.N., Belhumeur, P.N., Boult, T.E.: How reliable are your visual attributes? In: Proceedits of SPIE. vol. 8712, pp. 87120Q–87120Q–12 (2013)
28. Sharif, M., Mohsin, S., Javed, M.Y.: A survey: Face recognition techniques. Research Journal of Applied Sciences, Engineering and Technology 4(23), 4979–4990 (December 2012)
29. Taigman, Y., Yang, M., Ranzato, M.A., Wolf, L.: Deepface: Closing the gap to human-level performance in face verification. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2014)
30. Tan, X., Chen, S., Zhou, Z.H., Zhang, F.: Face recognition from a single image per person: A survey. Pattern Recognition 39(9), 1725–1745 (2006)
31. Thoma, G., Antani, S., Gill, M., Pearson, G., Neve, L.: People Locator: a system for family reunification. IT Professional 14, 13–21 (2012)
32. Wolf, L., Hassner, T., Taigman, Y.: The one-shot similarity kernel. In: In International Conference on Computer Vision. pp. 897–902 (September 2009), http://www.openu.ac.il/home/hassner/projects/Ossk
33. Zhang, C., Zhang, Z.: A survey of recent advances in face detection. Tech. rep., Microsoft (2010)
34. Zhou, H., Mian, A., Wei, L., Creighton, D., Hossny, M., Nahavandi, S.: Recent advances on singlemodal and multimodal face recognition: A survey. Human-Machine Systems, IEEE Transactions on 44(6), 701–716 (Dec 2014)
35. Zuliani, M., Kenney, C.S., Manjunath, B.S.: The multiRANSAC algorithm and its application to detect planar homographies. In: IEEE International Conference on Image Processing (September 2005), http://vision.ece.ucsb.edu/publications/05ICIPMarco.pdf