RPM: Random Points Matching for Pair-wise Face-Similarity

M. Saquib Sarfraz[1]
saquib.sarfraz@kit.edu
Muhammad Adnan Siddique[2]
siddique@ifu.baug.ethz.ch
Rainer Stiefelhagen[1]
rainer.stiefelhagen@kit.edu

[1] Institute for Anthropomatics
Karlsruhe Institute of Technology
Karlsruhe, Germany

[2] Swiss Federal Institute of Technology (ETH) Zurich
Zurich, Switzerland

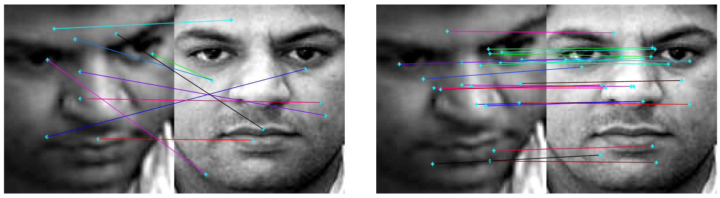
Pair-wise face verification strives to determine whether two probe images belong to the same person. With the increasing demand of face recognition/verification in real-world applications, the challenges confronted by the verification algorithms include not only dealing with large pose and lighting variations, but also very different imaging conditions e.g. different imaging device or resolution, sensor noise, varied expressions, lens distortion and occlusions.

Face verification algorithms generally require image representation followed by matching. For the representation, the current methods typically rely on aligning the two images such that the effect of pose and other geometric deformations are minimized. A feature analysis is then performed for matching, either globally on the whole face, or locally on some points typically found by a fiducial point detector. The over restrictive requirement of image alignment, especially for global feature analysis, is not trivial to meet. Besides the need to know the pose a priori, it requires 2D warps or 3D model-fitting which in turn depends on reliably detected fiducial points. Therefore, the performance of a verification algorithm is inevitably limited by the possibility of detecting reliable fiducial points.

In general, even the state-of-the-art fiducial or key point detectors do not guarantee to find all the required fiducial points. SIFT and similar detectors may fail to provide useful points for matching, especially when the probe images exhibit wide imaging differences. The points found do not guarantee to correspond to the same physical location in the two images and thus result in poor matching correspondence. Figure 1 a. depicts this problem by showing the matches found by employing the SIFT key point detector and matching engine. The image pair shown is comprised of a high resolution frontal gallery image and the probe face detected from a poor resolution surveillance video of the same individual. Due to the underlying imaging differences, the SIFT detector has failed to provide good corresponding points belonging to the same physical location on the two images. Fiducial point detectors pose similar behaviour, particularly in case of wide pose differences (e.g. when matching a frontal and profile view).

In this paper we shift from this over reliance of key point detectors and propose a novel method to automatically obtain well-matched points across the two images, subsequently referred to as the homologous points. we argue that this key/fiducial point detector may not guarantee a good correspondence in the presence of large pose and other imaging variations especially when the underling image modalities are different due to , for example, different image sensors. This is largely because these detectors try to find the specific regions on the face and thus become image specific. We on the other hand propose to find the corresponding point pair (homologous points) that falls on the same physical location from a dense set of randomly initiated points. These pairs can directly provide useful image similarity information. We use these points to establish direct pair-wise similarity, nonetheless, all of the current methods may benefit by using these newly established correspondences.

The procedure is based on randomly initializing enough points on the detected face windows and then follow up with a powerful mutual information assisted matching strategy. The approach lifts the restrictions, noted in the preceding paragraph, of the current key/fiducial point detectors as it is more probable to yield well matched pairs of points (on the same physical location on the face) from the randomly initialized points by using a meaningful iterative matching strategy. The procedure is such that it automatically finds pairs of points that not only correspond well but are also meaningful in establishing the similarity of the face in the image pair. To finally match the image pair we propose a new similarity metric that combines the useful parameters found during random point matching and the similarity scores computed by using a local descriptor on these homologous points.



(a) (b)

Figure 1: Point features matching in a pair of positive images: (a) with SIFT detector and descriptor; (b) with RPM (our approach).

Our approach is to match the candidate regions 'C' centred on each of the random points in the pair by using normalized Mutual Information (nMI).

$$m^i = \arg\max_j \mathrm{nMI}\left(C_A^i, C_B^j\right) \tag{1}$$

where m^i is the matched point. From among the set of these matched points the one that offers the highest MI is retained.

$$\begin{gathered} i_r = \arg\max_{i_k} \mathrm{nMI}\left(C_A^{i_k}, C_B^{m^i}\right) \\ i_r \rightarrow m^{i_r} \end{gathered} \tag{2}$$

We then use a recursive model fitting and an area based optimization to refine these estimates.

$$\Delta T_{xy}^{i_r} = \arg\max_{(\Delta T_{xy})} \mathrm{nMI}\left(C_A^{i_r}, C_B^{m^i + \Delta T_{xy}}\right) \tag{3}$$

The matched pairs retained at the end are the homologous points. We then combine the number of these homologous points 'n', the variance of the model parameters σ (obtained during matching) and the local similarities 'L' computed using a local descriptor on these points, in a novel similarity metric referred to as Q-score.

$$Q = \exp\left(\frac{(L+1)\times n}{\sigma + n}\right) \tag{4}$$

We tested our approach on images that exhibit wide pose and other imaging differences using two challenging datasets, FacePix and LFW. Our proposal shows an improvement over state-of-the-art methods that computes a direct pair-wise similarity.

The RPM formulation, to the best of our knowledge, is the first attempt at generating reliable homologous points among a pair of face images from completely random points, obviating the need to invoke a key or fiducial point detector as traditionally required before performing a local or global pair-wise feature analysis. It lends us useful metrics for pair-wise face similarity verification, such as the proposed Q-score which performs better than some of the state-of-the-art methods. Moreover, our work helps in setting the new perspective that the non-trivial requirement of image alignment (for an appropriate representation prior to matching) can be circumvented.

We aim to extend this work in two directions. In this paper, the different metrics originating from the RPM formulation have been used to build a direct measure of similarity. We intend to investigate the potential use of these metrics in training a classifier as well. Secondly, as the RPM formulation provides well-matched homologous points even in case of wide variations in imaging conditions, it can be used for pose-estimation and image alignment itself. We expect that the RPM formulation may provide a general solution to various machine vision problems.