Fusion of intra- and inter-modality algorithms for face-sketch recognition

Christian Galea and Reuben A. Farrugia

Department of Communications and Computer Engineering, Faculty of ICT,
University of Malta, Msida, Malta
{christian.galea.09,reuben.farrugia}@um.edu.mt

Abstract. Identifying and apprehending suspects by matching sketches created from eyewitness and victim descriptions to mugshot photos is a slow process since law enforcement agencies lack automated methods to perform this task. This paper attempts to tackle this problem by combining Eigentransformation, a global intra-modality approach, with the Eigenpatches local intra-modality technique. These algorithms are then fused with an inter-modality method called Histogram of Averaged Orientation Gradients (HAOG). Simulation results reveal that the intra- and inter- modality algorithms considered in this work provide complementary information since not only does fusion of the global and local intra-modality methods yield better performance than either of the algorithms individually, but fusion with the inter-modality approach yields further improvement to achieve retrieval rates of 94.05% at Rank-100 on 420 photo-sketch pairs. This performance is achieved at Rank-25 when filtering of the gallery using demographic information is carried out.

Keywords: face recognition, synthesis, intra-modality, inter-modality, hand-drawn sketches

1 Introduction

Recollections of eye-witnesses and victims are important to assist law enforcement agencies in identifying and apprehending suspects, especially when there is lack of evidence at the crime scene [5, 10]. Facial composite sketches created using witness descriptions are then disseminated to the public so that any persons recognising the suspect in the sketch may come forward with information that will hopefully lead to an arrest. However, this process is time-consuming, not guaranteed to be successful and inefficient since resources such as mugshot photo galleries available at law enforcement agencies cannot be utilised [10]. Hence, automated methods are required to match sketches with photos.

Several algorithms have been proposed in the literature for this task, which can be categorised as either intra-modality or inter-modality approaches [2]. Algorithms belonging in the first category attempt to reduce the modality gap by transforming a photo to a sketch (or vice versa) so that all images to be compared lie within the same modality. Any kind of face recogniser can then be applied

to match the synthesised pseudo-sketches (photos) with original sketches (photos). However, these algorithms have been criticised for being potentially more complex than the task of face recognition itself [2, 4]. On the other hand, inter-modality approaches perform face sketch recognition by extracting modality-invariant features to compare sketches with photos directly.

To the best of the authors' knowledge, even though numerous algorithms have been proposed for both types of approaches, neither the fusion of intra-modality methods nor the fusion of both intra- and inter-modality algorithms has been considered for photo-sketch recognition. In this paper, the combination of a global intra-modality method with a local intra-modality method is performed along with the fusion of these methods with a reportedly robust inter-modality technique. It is shown that these approaches provide complementary information to each other and provide superior performance when combined, across almost all ranks especially at lower ones.

The rest of this paper is organised as follows: the related work in the area of matching sketches with photos is presented in Sect. 2 followed by a description of the proposed method in Sect. 3. An overview of the experimental protocol used and the evaluation of the proposed method are given in Sect. 4. Conclusions and ideas for future work are finally outlined in Sect. 5.

2 Related Work

Several algorithms have been proposed for matching sketches with photos, although most can be categorised as being intra-modality techniques. One of the first such methods was the Eigentransformation algorithm [14], a global method based on Eigenfaces [15] which synthesises whole faces. A local-based intra-modality method was later proposed by the authors of [11], based on the Locally Linear Embedding (LLE) manifold learning technique such that patches are reconstructed using a linear combination of neighbouring patches. Kernel-based non-linear discriminant analysis (KNDA) is then used to match a probe sketch with the pseudo-sketches. This approach was improved in [16] by modelling spatial relationships among patches by a multi-scale Markov Random Fields (MRF) model. Patches were extended to encompass entire facial components in the Embedded Hidden Markov Model (E-HMM) method proposed in [3], where non-linear mapping between photo/sketch pairs is learned using HMMs for each component. Multiple models are obtained and fused using selective ensemble to improve the system's generalisation ability and thus enable the synthesis of more accurate pseudo-sketches. Recognition between probe sketches and pseudo-sketches is performed using the Principal Component Analysis (PCA) classifier.

Although most approaches have focused on reducing the modality gap, inter-modality methods are being given more attention lately especially by the Biometrics Research Group at Michigan State University [4, 7, 9, 10]. These approaches are based on the extraction of Multi-scale Local Binary Pattern (MLBP) and Scale-Invariant Feature Transform (SIFT) descriptors from patches to directly match sketches and photos in terms of their feature representations. The recently

proposed FaceSketchID system [10] combines the component-based approach proposed in [4], in which MLBP codes were extracted for each facial component, with the holistic-based heterogeneous face recognition algorithm in [9] to obtain a system with performance that was shown to exceed that of either of the two techniques individually. In other words, the fusion of local and global inter-modality algorithms yielded improvements in performance. The authors of [8] also attained performance gains with the fusion of two inter-modality approaches, namely using 'common' and 'direct' matching of multi-scale SIFT descriptors. Fusion with a commercial face recogniser yielded further improvement.

The work presented in this paper differs from the related work since (i) a local intra-modality algorithm based on the computation of a global intra-modality approach on a local level is implemented, (ii) the fusion of two intra-modality methods is investigated, and (iii) the fusion of intra- and inter-modality approaches is also analysed. To the best of the authors' knowledge, these contributions have never been investigated in previous works.

3   Proposed Approach

The proposed approach combines the scores output by three algorithms, namely the Eigentransformation [14] and Eigenpatches intra-modality techniques and the Histogram of Averaged Orientation Gradients (HAOG) [2] inter-modality descriptor. The system flow diagram is shown in Fig. 1. Both Eigentransformation and Eigenpatches use a set of training photo/sketch pairs to synthesise a pseudo-sketch or pseudo-photo, while the HAOG method extracts the inter-modality features for each image in the test set to be compared. The synthesised pseudo-sketches (pseudo-photos) derived from gallery photos (probe sketches) are compared to probe sketches (gallery photos) using the PCA face recogniser while the Chi-Square (χ^2) histogram matching method is used to evaluate the distance between the HAOG descriptors of probe sketches and those of the gallery photos. The resulting scores are then normalized using the *Normalisation module* which are finally fused with the *Fusion module*. A summary of each of these algorithms will now be given.

3.1   Intra-modality approaches

In the traditional Eigenfaces approach [15], face photos are reconstructed by using a weighted summation of a set of eigenvectors $\mathbf{U}_p$ representing faces, consequently called Eigenfaces, as follows:

$$\boldsymbol{P}_r = \mathbf{U}_p \boldsymbol{b}_p \tag{1}$$

where $\boldsymbol{P}_r$ is a column vector representing the reconstructed face photo and $\boldsymbol{b}_p$ is a vector containing the projection coefficients in the eigenvector space. As demonstrated in [14], (1) may be rewritten as follows:

$$\boldsymbol{P}_r = \bar{\boldsymbol{P}} + \sum_{i=1}^{M} \boldsymbol{c}_p^{\{i\}} \boldsymbol{\Phi}^{\{i\}} \tag{2}$$
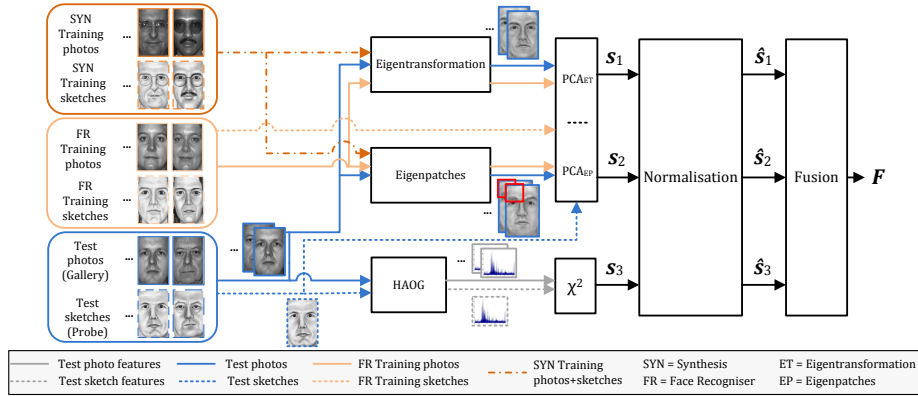
Fig. 1: System flow diagram of the proposed approach, for photo-to-sketch synthesis. Photos and sketches simply switch roles for sketch-to-photo synthesis.

where $\bar{\boldsymbol{P}}$ is the mean face computed over M training face images, $\boldsymbol{\Phi}^{\{i\}} = \boldsymbol{P}^{\{i\}} - \bar{\boldsymbol{P}}$ and $\boldsymbol{c}_p^{\{i\}}$ is a column vector of dimension M representing the contribution of the ith training image $\boldsymbol{P}^{\{i\}}$ in the reconstruction of a test face image computed according to [14]. Hence, (2) shows that a reconstructed photo can be approximated to the original image using a weighted linear addition of the training images [14]. Since a photo and corresponding sketch should also be similar in terms of structure, (2) may be modified such that the training photos $\boldsymbol{P}^{\{i\}}$ are replaced by the corresponding training sketches $\boldsymbol{S}^{\{i\}}$, as follows:

$$\boldsymbol{S}_r = \bar{\boldsymbol{S}} + \sum_{i=1}^{M} \boldsymbol{c}_p^{\{i\}} \boldsymbol{\Psi}^{\{i\}} \tag{3}$$

where $\boldsymbol{S}_r$ is the reconstructed sketch, $\bar{\boldsymbol{S}}$ is the mean sketch, $\boldsymbol{\Psi}^{\{i\}} = \boldsymbol{S}^{\{i\}} - \bar{\boldsymbol{S}}$ and $\boldsymbol{S}^{\{i\}}$ is a column vector representing the ith sketch. This is based on the hypothesis that if a photo contributes more weight to a reconstructed face photo, then the corresponding sketch will also contribute more weight to the reconstructed sketch [14].

The Eigentransformation approach has recently been extended to operate on local patches by the authors of [1] for the area of face super-resolution. This approach may be applied for sketch (or photo) synthesis by synthesising each patch in an image (rather than the whole image) and can be achieved by learning the optimal linear combination of patches found in the same local area of images in the training set. Formally, for each of the n patches in an image, (3) can be rewritten as follows:

$$\boldsymbol{S}_r^{\{j\}} = \bar{\boldsymbol{S}}^{\{j\}} + \sum_{i=1}^{M} \boldsymbol{c}_p^{\{i,j\}} \boldsymbol{\Psi}^{\{i,j\}} \quad \text{for } j = 1, 2, \ldots, n \ . \tag{4}$$

where $\mathbf{S}_r^{\{j\}}$ is the jth patch of the synthesised sketch, $\boldsymbol{\Psi}^{\{i,j\}} = \boldsymbol{S}^{\{i,j\}} - \bar{\boldsymbol{S}}^{\{j\}}$, $\boldsymbol{S}^{\{i,j\}}$ is the ith training sketch of patch j, $\bar{\boldsymbol{S}}^{\{j\}}$ is the jth mean patch and $\boldsymbol{c}_p^{\{i,j\}}$ are the reconstruction weights for the jth patch derived using the ith training face image. To the best of the authors' knowledge, this is the first use of Eigenpatches for face-sketch synthesis. Similar to other approaches involving patch-based operations, the overlap of patches with their neighbours is arbitrarily fixed at half the patch size and a simple averaging operation is done on overlapping areas to combine patches together.

For both methods, the role of sketches and photos is simply interchanged for pseudo-photo synthesis. Although any face recogniser can then be used to match the pseudo-sketches (photos) with the original sketches (photos), PCA (Eigenfaces method) [15] is implemented due to its widespread use in the literature for intra-modality methods. Hence, photo and sketch subspaces are learned and the face images to be compared are projected into these subspaces, similar to the approach in [14].

3.2   Inter-modality approach

The authors of [2] observed that effective inter-modality matching of sketches with photos may be obtained by extracting features only from those regions exhibiting the smallest amount of modality gap, namely those corresponding to coarse texture which represent boundaries of facial components critical in representing face images. Using a histogram whose bins represent orientation gradients for each patch, coarse textures are emphasised by accumulating squared magnitudes in those bins corresponding to the orientations of the patch being considered. In other words, more weighting is applied to those regions having high magnitudes when constructing a histogram of orientation gradients. The histograms of each patch are concatenated to yield the final HAOG descriptor and comparison of histograms is performed by finding the χ^2 distance, which is also the match score between two face images. More information on the HAOG descriptor is provided in [2].

3.3   Fusion

Fusion of the two intra-modality algorithms together and with the HAOG algorithm is performed at the matching score level, by first normalising the scores output from face recogniser using min-max normalisation. If a gallery contains a total of T subjects, then there will be T scores for each probe sketch representing the distance between the probe and all the photos in the gallery. The normalisation of these scores is achieved using min-max normalisation:

$$\hat{\boldsymbol{s}} = \frac{(\boldsymbol{s} - min)}{(max - min)} \tag{5}$$

where $\boldsymbol{s}$ is a vector of dimension T containing the scores output by a face-sketch recognition algorithm for matching a probe with the T gallery images, $\hat{\boldsymbol{s}}$ contains

the corresponding normalised scores, and min and max represent the minimum and maximum values of s, respectively [13].

The sum-of-scores method is then used to fuse the normalised scores together:

$$\boldsymbol{F} = \sum_{k=1}^{N} \hat{\boldsymbol{s}}_k \tag{6}$$

where $\hat{\boldsymbol{s}}_k$ is the score of kth face recognition method, N represents the number of intra- and inter- modality methods considered, and $\boldsymbol{F}$ is the final similarity score between a sketch (pseudo-photo) and the T pseudo-sketches (photos) [13]. Sum-of-scores fusion and min-max normalisation were chosen since they have been shown to provide some of the best results for fusion of multi-biometric systems in [6] and [13].

4   Results

The algorithms described in Sect. 3 are evaluated on the Chinese University of Hong Kong Face Sketch FERET (CUFSF) database [17] [1]. This database contains pairs of photos and viewed hand-drawn sketches of subjects in the Color FERET database [12]. In this paper, 842 subjects are considered.

For the intra-modality approaches, training is required for the transformation and face recognition stages. An approach similar to that used in [16] is adopted, namely $M = 211$ subjects are first used to train the transformation algorithm being considered (Eigentransformation or Eigenpatches). Another 211 subjects are then selected to train the PCA face recogniser, where their photos (sketches) are used along with the corresponding synthesised pseudo-photos (sketches) derived from the sketches (photos) for sketch-to-photo synthesis (photo-to-sketch synthesis). The remaining 420 subjects are then used for testing, i.e. $T = 420$. In addition, the performance of Eigenpatches for varying patch sizes is evaluated to determine the optimal patch size. Patch overlap is set to be half the patch size in all cases. For the HAOG algorithm, the parameters are set to those reported in [2].

Photos and sketches in the testing set form the gallery and probe sets, respectively. For photo-to-sketch synthesis, photos in the gallery are transformed into pseudo-sketches and matched with each sketch in the probe set. For sketch-to-photo synthesis, probe sketches are transformed to pseudo-photos and then matched with photos in the gallery.

Since PCA is used as a face recogniser in the intra-modality approaches, recognition using PCA only (without performing synthesis) is also carried out as a baseline performance measure, where each sketch in the probe set is directly matched with photos in the gallery. For all face recognition experiments carried out with PCA, the distance between the resultant vectors of the images to be compared represents the match score between face images and is found using

---

[1] Available at: `http://mmlab.ie.cuhk.edu.hk/archive/cufsf/`

$L2$-norms (i.e. Euclidean distance). In addition, each image was converted to grayscale, aligned using eye and mouth coordinates such that they are at fixed locations in all images considered and cropped to the same size of $300{\times}190$ pixels.

The subjects in the Color FERET dataset were pre-labelled with gender and race information, enabling filtering of the gallery based on the characteristics of the probe sketch. For example, if the sketch depicts a female subject, then all male subjects can be eliminated from the gallery. Statistics of the demographic distribution for the subjects considered in this study are shown in Table 1.

While face recognisers are typically measured at Rank-1 matching accuracy, the best match is often not the required suspect due to inaccuracies provided by a witness in describing facial details [4, 5]. As a result, the top 100 matches are typically examined with equal importance in forensic investigations and therefore the performance of an automatic system is recommended to be measured at around Rank-100 [4]. Graphs depicting the cumulative rank retrieval rates are thus shown to evaluate the ability of the algorithms in identifying persons.

For brevity, 'photo-to-sketch' synthesis will be denoted as 'P2S' and 'sketch-to-photo' synthesis will be denoted as 'S2P' in the remainder of this paper.

Table 1: Demographic statistics. Labels used are the same as those given in the Color FERET dataset.

Gender		Race	
Male	57.14%	White	63.33%
Female	42.86%	Asian	15.00%
		Black-or-African-American	9.29%
		Hispanic	6.19%
		Asian-Middle-Eastern	4.52%
		Pacific-Islander	0.95%
		Asian-Southern	0.24%
		Other	0.48%

4.1 Eigenpatches patch size

The rank retrieval rates using different values of the patch size are given in Fig. 2. Patch sizes were set to be equal to $2^p \times 2^p$ for $p = 3, 4, \ldots, 7$. Firstly, it can be observed that P2S synthesis is superior to S2P synthesis for all patch sizes at all ranks. This is likely due to the fact that in transforming photos to sketches, a large amount of information is being compressed into a smaller representation since photos tend to be more complex and therefore contain more information. This is a more stable operation than attempting to expand a small representation into a larger one, as in the case of S2P synthesis [14]. As can also be observed from Fig. 3, it is evident that S2P synthesis is unable to recover a face photo

adequately. It should also be noted that the performance tends to improve with increasing patch size. From empirical observations, low patch sizes resulted in high amounts of noise which decreased with larger patch sizes. In addition, at the largest patch size of 128×128 pixels, the patches roughly overlap with the individual facial components. Hence, synthesis is virtually being done for each component and is therefore a contributor to increased quality of the synthesised images. Based on these observations, Eigenpatches with a patch size of 128×128 pixels is used for the remainder of the experiments.
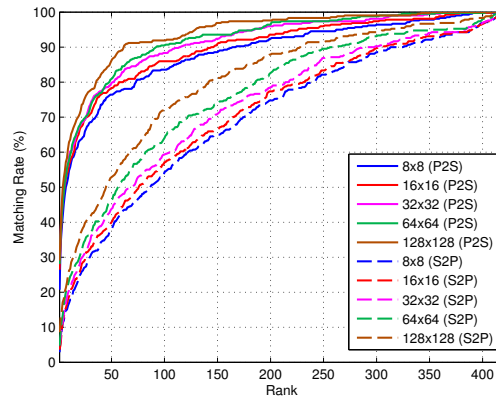

Fig. 2: Recognition rate for varying patch sizes of the Eigenpatches algorithm, using the Color FERET/CUFSF datasets

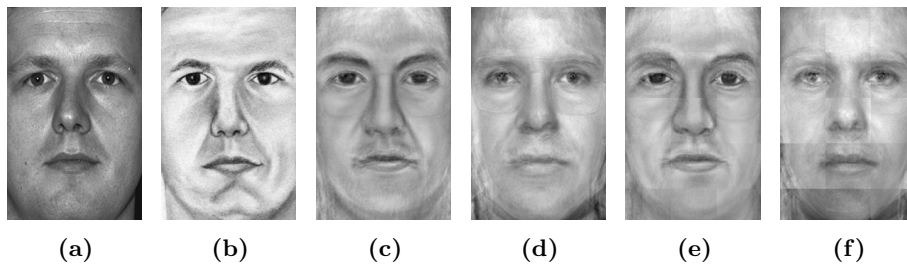(a)   (b)   (c)   (d)   (e)   (f)

Fig. 3: Synthesised images of one subject in the Color FERET/CUFSF datasets: (a) Original photo, (b) Original sketch, (c) Eigentransformation P2S, (d) Eigentransformation S2P, (e) Eigenpatches P2S, (f) Eigenpatches S2P

4.2 Matching rate

This section investigates the performance of the independent intra- and inter-modality methods mentioned in this paper and a combination of these methods. The direct use of PCA (without performing synthesis) is used as a baseline. The Rank retrieval rates are shown in Fig. 4 and Table 2. It should be noted that all algorithms outperform the baseline face recogniser at all ranks. This clearly indicates the need of using approaches designed specifically for person identification via matching of sketches with photos. In particular, the fact that PCA is better able to match synthesised images indicates that the intra-modality approaches are successfully reducing the modality gap between photos and sketches.

The Eigentransformation and Eigenpatches intra-modality approaches yield superior performance to the HAOG inter-modality approach for P2S synthesis, across almost all ranks. For S2P synthesis, the intra-modality approaches are inferior to HAOG especially for Ranks below 150. It can also be noticed that Eigenpatches typically achieves higher matching rates at lower ranks (until roughly Rank-50) for P2S synthesis. This indicates that Eigenpatches is able to construct images which are able to discriminate between persons more reliably due to the utilisation of local features, while Eigentransformation also encodes spatial information which enables it to perform better at higher ranks when faces are being compared at a more global level.

Fusion of the two intra-modality algorithms yields noticeable improvements at most ranks. Even in cases where the performance is less than the best performance of either of the intra-modality algorithms, generally the difference is only marginal and is better than one of the intra-modality algorithms. This holds for both S2P and P2S synthesis and indicates that the two approaches contain complementary information which can improve performance. Given that the Eigentransformation algorithm can be considered as a holistic algorithm whereas the Eigenpatches algorithm can be viewed as a local descriptor, this result is in line with the observations reported in [10] where the fusion of holistic- and component-based inter-modality algorithms yielded the best performance.

Further improvements in matching rates can be achieved with the fusion of the HAOG algorithm. In fact, this yields the best performance of all algorithms considered at almost all ranks, especially at lower ranks (roughly below Rank-100). The gain is more substantial when the intra-modality algorithms using P2S synthesis are fused, likely due to the better performance of these algorithms with respect to when they are used for S2P synthesis. These results not only demonstrate that Eigentransformation, Eigenpatches and HAOG contain complementary information which can be exploited for increased matching rates, but they are also indicative that the combination of both intra- and inter-modality algorithms may be the ideal approach for matching sketches with photographs rather than using *either* intra-modality algorithms *or* inter-modality algorithms.

Filtering using demographic information yields substantial improvements for all algorithms considered. In addition, as demonstrated in Fig. 4e, filtering with race information provides better performance than with gender information. For example, the Rank-1 retrieval rate of the proposed approach improves from

Table 2: Rank retrieval rates for the algorithms considered when evaluated on the CUFSF dataset, without demographic filtering; ET = Eigentransformation, EP = Eigenpatches. The best result for each rank is highlighted in boldface.

Metric	Matching Rate (%) at Rank-N					
	Rank-1	Rank-10	Rank-50	Rank-100	Rank-150	Rank-200
PCA	0.71	5.48	25.00	37.62	46.43	58.57
HAOG	27.14	48.10	68.81	80.95	84.76	89.05
ET (P2S)	24.05	56.90	84.29	91.43	96.19	97.38
ET (S2P)	9.29	30.48	57.62	72.38	82.14	85.71
EP (P2S)	29.05	62.86	86.19	91.90	**96.90**	97.62
EP (S2P)	8.57	26.19	52.86	72.14	81.19	87.86
ET + EP (P2S)	26.90	64.52	**87.38**	94.04	95.71	**98.57**
ET + EP (S2P)	8.81	29.05	58.33	73.81	83.10	89.29
ET + EP + HAOG (P2S)	**44.05**	**71.90**	**87.38**	**94.05**	95.48	97.14
ET + EP + HAOG (S2P)	24.52	51.43	72.14	82.38	88.10	92.62

44.05% to 46.43% using gender information and to 55.71% using race information. This is likely due to the fact that subjects are distributed more evenly in terms of gender. This observation was also reported in other works in the literature, e.g. in [7].

5 Conclusion

This work presented the use of Eigenpatches for intra-modality sketch-to-photo and photo-to-sketch applications. It was demonstrated that Eigenpatches performs similarly to Eigentransformation except at lower ranks for photo-to-sketch synthesis, where it is superior. This indicates that Eigenpatches is able to synthesise images having more accurate local structures which are better able to discriminate between persons. It was also observed that photo-to-sketch synthesis generally provides superior results over both sketch-to-photo synthesis and the HAOG inter-modality approach.

Furthermore, it was shown that the fusion of intra- and inter-modality algorithms yields the best performance across virtually all ranks. This indicates that not only do the local and global intra-modality methods and the inter-modality method considered in this paper provide complementary information, but their combination provides the best performance for matching sketches with photos.

Use of demographic information to filter the gallery yielded substantial improvements to result in a Rank-25 retrieval rate of 94.05%. Future work includes the suppression of blocking artefacts created due to patch averaging for images synthesised with Eigenpatches and the use of software-generated composites which are widely used in forensics.
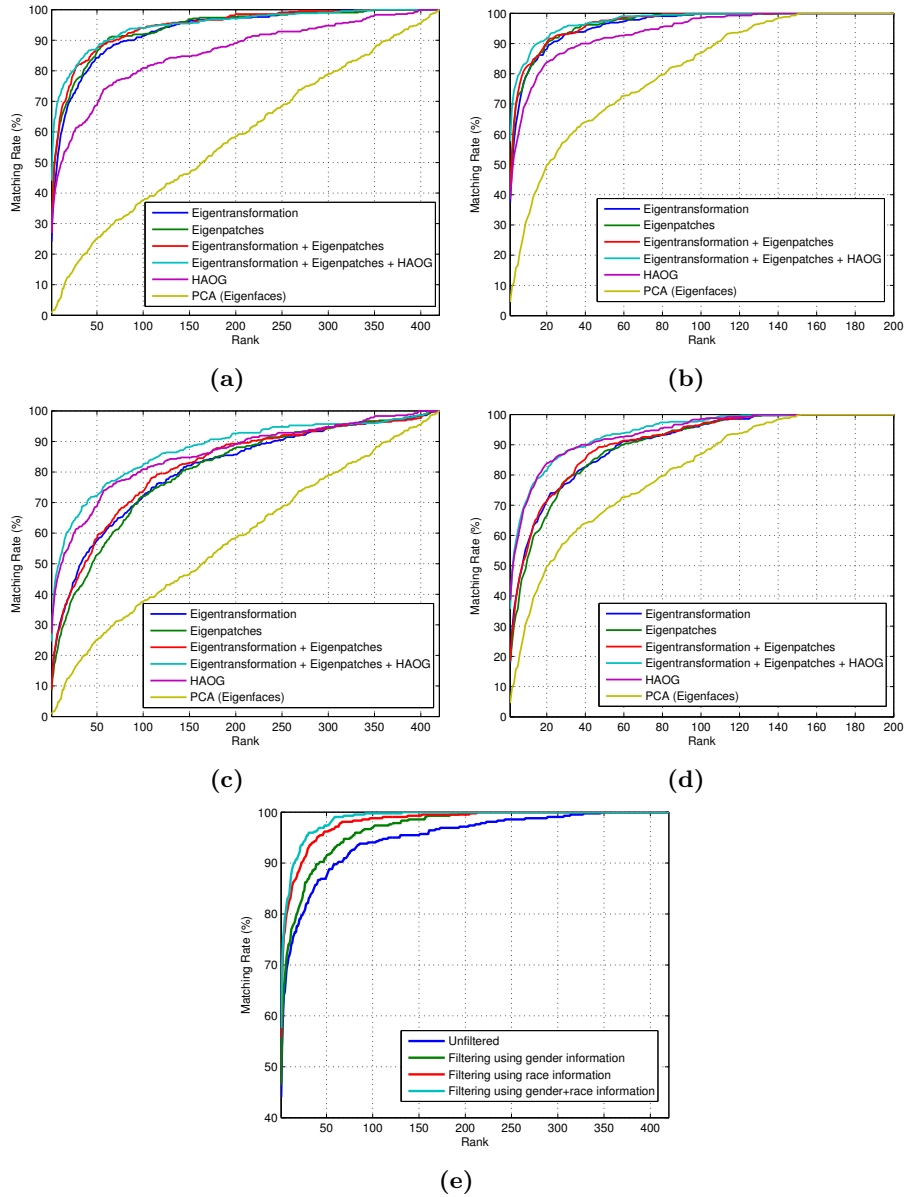
Fig. 4: Matching performance for: (a) P2S synthesis, (b) P2S synthesis with filtering, (c) S2P synthesis, (d) S2P synthesis with filtering, (e) Filtering the gallery using different types of demographic information for P2S synthesis of the proposed method (using the fusion of Eigentransformation, Eigenpatches and HAOG).

References

1. Chen, H.Y., Chien, S.Y.: Eigen-patch: Position-patch based face hallucination using eigen transformation. In: IEEE International Conference on Multimedia and Expo. pp. 1–6 (July 2014)
2. Galoogahi, H., Sim, T.: Inter-modality face sketch recognition. In: IEEE International Conference on Multimedia and Expo. pp. 224–229 (July 2012)
3. Gao, X., Zhong, J., Li, J., Tian, C.: Face Sketch Synthesis Algorithm Based on E-HMM and Selective Ensemble. IEEE Transactions on Circuits and Systems for Video Technology 18(4), 487–496 (April 2008)
4. Han, H., Klare, B.F., Bonnen, K., Jain, A.K.: Matching composite sketches to face photos: A component-based approach. IEEE Transactions on Information Forensics and Security 8(1), 191–204 (Jan 2013)
5. Jain, A.K., Klare, B., Park, U.: Face matching and retrieval in forensics applications. IEEE MultiMedia 19(1), 20–20 (Jan 2012)
6. Jain, A.K., Nandakumar, K., Ross, A.: Score normalization in multimodal biometric systems. Pattern Recognition 38(12), 2270 – 2285 (2005)
7. Klare, B., Li, Z., Jain, A.K.: Matching forensic sketches to mug shot photos. IEEE Transactions on Pattern Analysis and Machine Intelligence 33(3), 639–646 (March 2011)
8. Klare, B., Jain, A.K.: Sketch-to-photo matching: a feature-based approach. In: Proceedings of the SPIE Conference on Biometric Technology for Human Identification VII. vol. 7667, pp. 766702–766702–10 (2010)
9. Klare, B.F., Jain, A.K.: Heterogeneous face recognition using kernel prototype similarities. IEEE Transactions on Pattern Analysis and Machine Intelligence 35(6), 1410–1422 (June 2013)
10. Klum, S.J., Han, H., Klare, B., Jain, A.K.: The FaceSketchID System: Matching Facial Composites to Mugshots. Tech. Rep. MSU-CSE-14-6, Michigan State University (2014)
11. Liu, Q., Tang, X., Jin, H., Lu, H., Ma, S.: A nonlinear approach for face sketch synthesis and recognition. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition. vol. 1, pp. 1005–1010 vol. 1 (June 2005)
12. National Institute of Standards and Technology (NIST): The Color FERET Database version 2, http://www.nist.gov/itl/iad/ig/colorferet.cfm
13. Snelick, R., Indovina, M., Yen, J., Mink, A.: Multimodal biometrics: Issues in design and testing. In: Proceedings of the 5th International Conference on Multimodal Interfaces. pp. 68–72. ICMI '03 (2003)
14. Tang, X., Wang, X.: Face sketch recognition. In: IEEE Transactions on Circuits and Systems for Video Technology. vol. 14, pp. 50–57 (January 2004)
15. Turk, M., Pentland, A.: Eigenfaces for recognition. Journal of Cognitive Neuroscience 3(1), 71–86 (1991)
16. Wang, X., Tang, X.: Face photo-sketch synthesis and recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 31(11), 1955–1967 (Nov 2009)
17. Zhang, W., Wang, X., Tang, X.: Coupled information-theoretic encoding for face photo-sketch recognition. In: 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 513–520 (June 2011)