Matching Age Separated Composite Sketches and Digital Face Images

Tarang Chugh, Himanshu S. Bhatt, Richa Singh, and Mayank Vatsa
IIIT-Delhi, INDIA
{tarang09050,himanshub,rsingh,mayank}@iiitd.ac.in

Abstract

Matching sketches with digital face images is one of the highly sought face recognition applications. An important aspect, though less explored till now, is matching age separated sketches with digital face images. Several law enforcement agencies are progressively using composite sketches for apprehending individuals. This research proposes an algorithm for matching composite sketches with digital face images across different age. It extracts discriminative shape, orientation, and texture features from local regions of a face using image moments and histogram of oriented gradients. The complementary information from these two features is further combined at match score level for efficiently matching composite sketches with digital face images across different age. To study the effects of age variations, this research also presents a composite sketch database with age separated sketches and digital face images. The results on a large gallery experiment suggest that the proposed algorithm efficiently encodes discriminative information from local facial regions useful for matching composite sketches with age separated digital face images.

1. Introduction

Matching sketches with digital face images has been one of the most important cues in apprehending criminals, finding missing individuals, and recognizing individuals when the face is reconstructed as a composite sketch post-mortem. Sketches are primarily a representation of soft and prominent edges around the facial features. Forensic sketches used in law enforcement suffer from a two-fold exaggeration: description of eye-witness and depiction by the sketch artist [5]. Matching sketches with digital face images is a challenging problem as these two represent information in very different forms. Digital images are rich in texture and feature information while sketches are just a blend of soft and prominent edges. Matching sketches with digital face images has recently gained attention from the research community and several algorithms have been proposed for matching sketches with digital face images [5, 10, 14]. However, these approaches mainly emphasize on matching hand-drawn sketches prepared by an artist on a paper based on the description from an eye-witness.

Figure 1. Illustrating the difference in information content among (a) hand-drawn sketches, (b) digital face images, and (c) composite sketches. Digital face images are obtained from the FG-Net database [3].

There has been a recent shift from hand-drawn sketches to composite sketches. Composite sketches are drawn using software tools that facilitate an eyewitness to select different facial components from a wide range of pre-defined templates. An eyewitness based on his/her recollection from the crime scene selects the most resembling facial template for each feature. These tools allow processing each feature individually and then combine all the features to generate a composite sketch. Preparing composite sketches require lesser effort both in terms of cost as well as time as compared to hand-drawn sketches. Figure 1 shows some examples of digital face images along with the corresponding composite sketches and hand-drawn sketches. Composite sketches may not include minute feature details as compared to what an artist can depict in hand-drawn sketches, therefore, composite sketches often look synthetic. Matching composite sketches with digital face images presents an even more challenging problem for automatic face recognition algorithms than matching hand-drawn sketches with digital face images. Recently, Han *et al.* [8] have proposed a component based representation approach for matching composites with digital images using multi-scale local binary pattern. Their algorithm utilizes gender information as

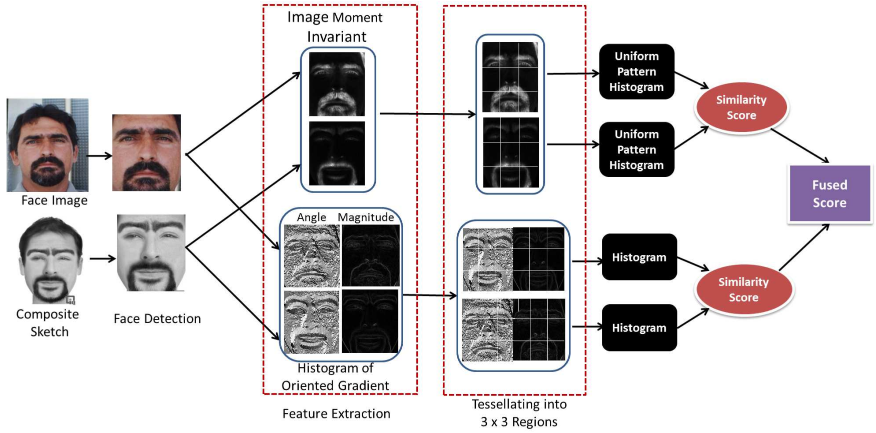

Figure 2. Illustrating the steps involved in the proposed algorithm for matching composite sketches with digital face images.

a soft biometric to reduce the search space in the gallery.

The problem of matching composite sketches with digital face images is not limited to the effects of variations in composite sketches and digital face images. In real world scenarios, it is often required to match composite sketches and digital face images with age variations; e.g., for finding missing individuals and recognizing individuals when the face is reconstructed as a composite sketch after death. Age variations further make this problem arduous as it changes the structural geometry and face texture. Existing research has not yet addressed this challenging problem; albeit, it represents a very realistic scenario in law enforcement applications. This research presents an algorithm to match composite sketches to digital face images with age variations. It encodes discriminative information from the shape, orientation, and texture of local facial regions using image moments [9] and histogram of oriented gradients (HOG) [7]. The histogram descriptors from local facial regions are concatenated to form a global image signature. Matching is performed using Chi square distance measure and the complementary information obtained from the two feature extractors is combined using sum-rule fusion [13] at the match score level. The authors also prepared a composite sketch database which comprises composite sketches and digital face images with age variations. To the best of our knowledge, this is the first research effort that attempts to address the problem of matching composite sketches with digital face images across age variations.

2. Matching Composite Sketches with Digital Face Images

As shown in Figure 2, the proposed algorithm starts with extracting local image moments [9] and histogram of oriented gradients (HOG) [7] from the local regions of both composite sketches and digital face images. Local image moments provide resilience to structural changes in shape and orientation caused due to aging. On the other hand, HOG encodes the discriminative texture information around the fiducial features as a distribution of edge intensities and gradients in local regions surrounding the features. Different stages of the proposed algorithm for matching composite sketches with digital face images are elaborated below.

2.1. Regular Image Moment Invariants

Hu [9] derived a set of regular image moment invariants which has been widely used for contour-based shape recognition due to their invariance to rotation, scaling and translation. It provides information such as localized weighted average of the pixel intensities, centroid, and orientation details of an image. Age variations bring changes in the shape of facial structures [12]; however, the orientation of different features is comparatively less affected or varies coherently [6]. Therefore, it is our assertion that image moments can efficiently encode discriminative information around the prominent features in local regions of sketches and digital face images that is resilient to age variations. Out of the 7 image moments proposed by Hu [9], second image moment M_2 is empirically determined to be the most discriminative feature for representing shape and orientation of local facial regions in composite sketches and digital face images. The second image moment is calculated as:

$$M_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \tag{1}$$

where η_{pq} represents the normalized central moment as:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}}; \quad \gamma = [\frac{(p+q)}{2}] + 1 \tag{2}$$

where μ_{pq} is the central moment. The central image moments are regular image moments which are shifted from the origin to coincide the image centroid $(\bar{x}, \bar{y})$ with the origin. It eliminates the effect of translation as it captures information regarding the spread and structure of an image around the centroid and is not affected by the absolute location. The central image moment μ_{pq} is calculated as:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q I(x, y) \tag{3}$$

where $\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}}$. For a gray-scale image I of size $M \times N$, regular moments of order $(p+q)$ are calculated as:

$$m_{pq} = \sum_{x=1}^{M} \sum_{y=1}^{N} x^p y^q I(x, y) \tag{4}$$

where $I(x, y)$ is the intensity value at image coordinates (x, y) and $p, q = 0, 1, 2, ...n$.

For extracting second image moment invariant, the image is scanned in raster fashion from top-to-bottom and left-to-right. The second order image moment for each pixel in the image is calculated based on 8 neighboring pixels separated by an angular difference of $45°$ starting with $\theta \in (22.5° - 337.5°)$ on a circle of radius $r = 2$ centered at the current pixel. The image moments are calculated for every pixel in image using Eqs. 1-4. The feature image is further tessellated into 3×3 non-overlapping patches and histogram of uniform binary patterns is computed for each local patch, as shown in Figure 2. A binary pattern is called uniform binary pattern if it has at most two bitwise transitions from 0 to 1 or vice-versa. The feature histogram is computed where every uniform pattern has a separate bin and all non-uniform patterns are assigned to a single bin. Concatenation of all the histograms pertaining to each local patch constitutes the image signature. To compare a gallery-probe pair, their features are matched using χ^2 distance.

$$\chi^2(x, \xi) = \sum_i \sum_j \frac{(x_{i,j} - \xi_{i,j})^2}{x_{i,j} + \xi_{i,j}} \tag{5}$$

where x and ξ are the two feature histograms, $i =$ local region, and $j =$ histogram bin.

2.2. Histogram of Oriented Gradients

The gradient of an image provides information about the edges and intensity variations. It can efficiently encode the shape of facial features and flow of pixel intensities in the neighborhood using edge and gradient information. Dalal and Triggs [7] proposed a global descriptor for human (pedestrian) detection in street view images, known as histogram of oriented gradients. HOG is based on the intuition that shape and position of a dominant feature can be understood by the distribution of orientations in local regions of an image. In this research, HOG is used for representing discriminative information from sketches and digital face images by encoding the oriented (directional) gradients that occur with higher frequency. The algorithm starts by dividing an image into small entities, known as cells, for which the oriented gradient is calculated. To remove the effects of varying illumination, the oriented gradients are normalized in larger overlapping regions consisting of many cells called blocks. The image is further tesselated in 3×3 patches and each patch after normalization forms a histogram of 9 bins. These histograms are concatenated together to form a global image descriptor comprising 81 bins. To match two HOG descriptors corresponding to composite sketches and digital face image, χ^2 distance measure is used.

2.3. Sketch Recognition Algorithm Across Age Variations

Since a composite sketch is primarily an edge representation of the actual face, edge preserving approaches that embed the orientation, shape, and texture are used in the proposed algorithm. Moreover, it is our assertion that the shape and orientation of local facial regions vary coherently in scale and translation due to aging [6]. Therefore, the proposed algorithm using histogram of image moments and oriented gradients can efficiently match composite sketches with digital face images along with age variations. The proposed algorithm is summarized below:

- For a given composite sketch and digital face image pair, eye-coordinates are detected using OpenCV‘s boosted cascade of haar-like features. Using the eye-coordinates, rotation is normalized with respect to the horizontal axis and inter-eye distance is fixed to 100 pixels. Finally, the face images are resized to 192×224 pixels.

- The histogram of image moments and HOG are extracted from both composite sketches and digital face images. These feature extractors encode discriminative shape and orientation information along the local facial regions which is used for matching composite sketches and digital images across age variations.

- To match a gallery-probe pair, corresponding descriptors (i.e histogram of image moments and HOG) are compared using the χ^2 distance.

- Complementary information from two descriptors is combined using sum rule fusion [13] of match scores.

  $$S_{fused} = w_0 * s_0 + w_1 * s_1 \tag{6}$$

  where w_0 and w_1 are the weights assigned and s_0 and s_1 are the scores for the histogram of image moments and HOG respectively. In our experiments, w_0 and w_1 are empirically set to 0.5 to yield the best performance.

- In identification mode (1:N), this procedure is applied for each probe and top matches are obtained.

3. Experimental Evaluation

The performance of the proposed algorithm for matching composite sketches with digital face images along the age variations is compared with an existing algorithm and a commercial-off-the-shelf (COTS) face recognition algorithm, Neurotechnology VeriLook [11].

Table 1. Illustrating the details of the composite sketch database prepared by authors.

Category	Digital images	Composite sketches
Young Age Group	361	59
Same Age Group	59	59
Old Age Group	198	59

3.1. Database

The authors prepared a composite sketch database of 59 subjects where each subject has multiple digital face images across different age groups. To prepare composite sketches for the digital face images, a professional composite sketch software, FACES [2], is used. It is a widely used software for drawing composite sketches comprising more than 4,400 facial features. It can be used to create several variations of faces using different combinations of pre-defined facial components. The database comprises a composite sketch of an individual (preferably at a middle age) and digital face images across different age groups. The digital face images in the database are categorized into three groups:

- **Young age group**: Digital face images of an individual at a younger age than the composite sketches are combined into the young age category. Mean age difference between the gallery and probe is 11 years.
- **Same age group**: All the digital images within a time gap of ± 6 months from the composite sketch are combined into the same age category. Mean age of the subjects in this category is 25 years.
- **Old age group**: All digital face images at an older age than the composite sketches are combined into the old age category. Mean age difference between the gallery and probe is 7.6 years.

To prepare the composite sketch database, 45 individuals are selected from the FG-Net aging database [3] that have a good range of age variations ranging from 1 to 65 years. For each of the 45 subjects in the database, a middle-age digital face image is selected and a composite sketch is prepared for that image using the FACES software [2] . Further, composite sketches pertaining to additional 14 subjects are obtained from the internet[1] and the corresponding digital face images are divided into three age categories. In total,

[1]These composite sketches are obtained from the Facebook page [1] of the tool FACES.

(a) (b) (c) (d)

Figure 3. Illustrating samples of (a) composite sketches and digital face images in (b) young, (c) same, and (d) old age groups.

the database comprises a total of 59 composite sketches and 618 digital images across different age categories. Law enforcement applications generally involve matching a composite sketch with a large database of digital face images. To replicate such real world scenario, 4543 digital face images obtained from a law enforcement agency are used to extend the gallery. The details about the database along with the number of images in different age categories are summarized in Table 1.

3.2. Experimental Protocol

The performance of the proposed algorithm is evaluated under three different scenarios as explained below:

- *Experiment 1 (Missing person scenario)*: The gallery comprises digital face images that are captured at a younger age as compared to that of the composite sketch. The gallery consists of 361 digital face images from the composite sketch database. Further, the gallery size is extended to 4904 by adding 4543 digital face images obtained from a law enforcement agency. The probe comprises 59 composite sketches from the composite sketch database. *Such a scenario is applicable in case of missing individuals where digital image at younger age is available in the database and a composite sketch drawn with the help of an eye-witness is used to find the identity of the individual.*
- *Experiment 2 (Recent crime scenario)*: The gallery comprises digital face images from the same age category. In this experiment, the gallery comprises a total of 4602 digital face images, 59 digital face images from the composite sketch database and 4543 images provided by a law enforcement agency. The probe comprises 59 composite sketches. *This scenario is applicable in cases when the police had a recent photo of the individual, for example a photo was captured*

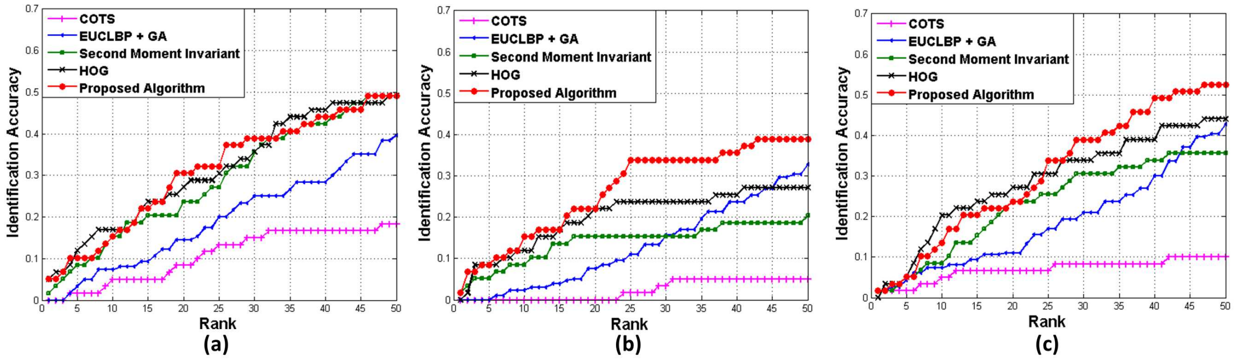

Figure 4. CMC curves for the proposed and existing algorithms for matching composite sketches with digital face images from (a) young, (b) same, and (c) old age category from the composite sketch database. Results can only be compared within age groups.

during booking and the same individual committed another crime within six months and a witness helped in drawing the composite sketch.

- *Experiment 3 (Cold case scenario)*: The gallery comprises digital face images that are captured at an instance later than when the composite sketch is prepared. In this experiment, the gallery comprises a total of 4741 digital face images, 198 images from the composite sketch database and 4543 images as an extended gallery. The probe comprises 59 composite sketches. *Such a scenario is applicable in cases when a composite sketch is prepared in the past and the law enforcement agency requires to match a composite sketch with a recent digital face image of an individual*.

In all three experiments, digital face images are used as gallery and the composite sketch is used as probe. Matching composite sketches with digital face images is more applicable in an identification scenario (1:N matching), therefore, the performance is reported in identification mode. Each experiment has different protocol i.e. different number of subjects in each category; therefore, the results can only be compared within age groups.

3.3. Experimental Results and Analysis

The performance of the proposed algorithm is compared with its individual constituent features, COTS, and a sketch recognition algorithm, referred to as EUCLBP+GA [4]. The key analysis and observations are listed below:

Table 2. Rank-50 identification accuracy of the different algorithms for matching composite sketches with digital face images.

Algorithm	**Rank-50 Accuracy (%)**		
	Exp. 1	**Exp. 2**	**Exp. 3**
COTS	18.3	5.0	10.0
EUCLBP+GA [4]	39.7	32.7	42.7
Second Image Moment	49.2	20.3	35.6
HOG	49.2	27.1	44.1
Proposed	**49.2**	**38.9**	**52.5**

- The Cumulative Match Characteristic (CMC) curves in Figure 4(a) show the rank-50 identification accuracy of existing algorithms for matching composite sketches with digital face images in the young age group (experiment 1). The proposed algorithm performs slightly better that individual constituent feature extractors in cases with young age gallery. The results in Table 2 show that the proposed algorithm yields same rank-50 identification accuracy as HOG and second order image moment. However, the proposed algorithm clearly outperforms EUCLBP+GA [4], and COTS by 10% and 30% at rank-50 respectively.

- The CMC curves in Figure 4(b) compare the performance of different algorithms for matching composite sketches with digital face images in the same age category (experiment 2). For experiment 2, Table 2 shows that the proposed algorithm clearly outperforms individual constituent features, EUCLBP+GA [4], and COTS by at least 6%. It validates our initial assertion that the proposed algorithm, using histogram of image moments and HOG, can efficiently encode the discriminative information from shape and orientation of local facial regions in composite sketches and digital face images. Fusing complementary information from two feature extractors using sum rule fusion at match score level further enhances the identification performance.

- CMC curves in Figure 4(c) show that the proposed algorithm yields better identification performance for matching composite sketches with old age gallery images as compared to other existing algorithms (experiment 3). It outperforms individual constituent features, EUCLBP+GA and COTS by at least 8% which is attributed to the fact that the proposed algorithm extracts features from composite sketches and digital face images that are fairly resilient to variations due to aging.

- The proposed algorithm captures discriminative infor-

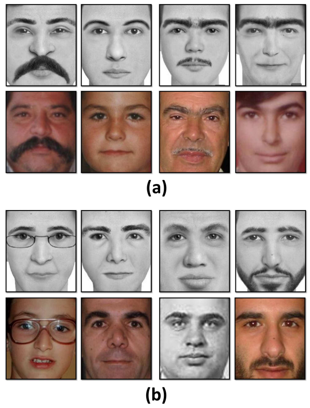

Figure 5. Illustrating sample cases when the proposed approach (a) correctly recognizes and (b) fails to recognize composite sketches with digital face images across age variations.

mation using image moments and HOG which is resilient to variations caused due to aging. It assimilates shape and orientation information from local facial features and its neighboring regions that eliminates the effect of scaling and facial feature drifts due to aging. Therefore, the proposed algorithm efficiently matches composite sketches with digital face images across different age groups.

- Figure 5(a) shows some cases when the proposed algorithm correctly identified the correct digital face image (in rank-50) from different age categories. Figure 5(b) shows cases where the proposed algorithm fails to match the composite sketch to correct digital face image in the gallery. The poor performance can be attributed to the large gallery size, poor representation of minute face details in the composite sketches.

- Finally, on a 2 GHz Intel Duo Core processor with 4 GB RAM under MATLAB programming environment, for a given composite sketch, it requires about 6 seconds to compute histogram of image moments and HOG descriptor and match it to a digital face image.

4. Conclusion

This research presents an algorithm for matching age separated composite sketches with digital face images. In the proposed algorithm, discriminative information is extracted from local facial regions using second order image moments and histogram of oriented gradients. It utilizes the observation that information assimilated from the shape, orientation, and texture of local facial regions provides resilience for matching composite sketches with digital face images of different age. Fusing complementary information from two features at match score level further enhances the identification performance. Experimental results on a composite sketch database, prepared by the authors show the efficacy of the proposed algorithm as compared to an existing sketch matching algorithm and a commercial face recognition system. As a future work, we plan to include gender and race demographic information to enhance the matching performance by reducing the search space.

References

[1] *Face Sketch Software : Famous Faces. Available at-https://www.facebook.com/FACESketchSoftware*.

[2] *FACES: The ultimate composite software. Available at-http://www.facesid.com*.

[3] FG-Net. Availabe at-http://www.fgnet.rsunit.com/.

[4] H. S. Bhatt, S. Bharadwaj, R. Singh, and M. Vatsa. On matching sketches with digital face images. In *Proceedings of International Conference on Biometrics: Theory, Applications and Systems*, pages 1–7, 2010.

[5] H. S. Bhatt, S. Bharadwaj, R. Singh, and M. Vatsa. Memetically optimized MCWLD for matching sketches with digital face images. *IEEE Transactions on Information Forensics and Security*, 5(5):1522–1535, 2012.

[6] S. Biswas, G. Aggarwal, N. Ramanathan, and R. Chellappa. A non-generative approach for face recognition across aging. In *Proceedings of International Conference on Biometrics: Theory, Applications and Systems*, pages 1–6, 2008.

[7] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Proceedings of International Conference on Computer Vision & Pattern Recognition*, pages 886–893, 2005.

[8] H. Han, B. F. Klare, K. Bonnen, and A. K. Jain. Matching composite sketches to face photos:A component-based approach. *IEEE Transactions on Information Forensics and Security*, 8(1):191–204, 2013.

[9] M.-K. Hu. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, 8(2):179–187, 1962.

[10] B. Klare, L. Zhifeng, and A. Jain. Matching forensic sketches to mug shot photos. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(3):639–646, 2011.

[11] Neurotechnology VeriLook SDK : available at http://www.neurotechnology.com/verilook.html.

[12] U. Park, Y. Tong, and A. Jain. Age-invariant face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(5):947–954, 2010.

[13] A. Ross and A. Jain. Information fusion in biometrics. *Pattern Recognition Letters*, 24(13):2115–2125, 2003.

[14] W. Zhang, X. Wang, and X. Tang. Coupled information-theoretic encoding for face photo-sketch recognition. In *Proceedings of International Conference on Computer Vision and Pattern Recognition*, pages 513–520, 2011.