BOOSTING LOCAL DESCRIPTORS FOR MATCHING COMPOSITE AND DIGITAL FACE IMAGES

Paritosh Mittal, Aishwarya Jain, Richa Singh and Mayank Vatsa

IIIT-Delhi, India

ABSTRACT

Sketch recognition is one of the most challenging applications of face recognition. Due to the incorrectness of features in the witness description, standard face recognition algorithms are generally not applicable to matching sketches with digital face images. This research designs a patch based face recognition algorithm that generates patches around fiducial features and extracts local information from these patches using Daisy descriptor. The information extracted from these patches are then efficiently matched using GentleBoostKO algorithm. The experiments performed on the PRIP composite face image database show that the proposed algorithm yields promising results and outperforms existing state-of-the-art algorithms and a commercial system.

***Index Terms*—** Sketch recognition, composite faces, daisy descriptor

1. INTRODUCTION

In face recognition literature, researchers have generally focused on matching two images captured using digital cameras. However, different law enforcement applications have led to other challenging applications of face recognition such as matching digital face images with surveillance camera images, scanned images, and sketches. This research focuses on matching sketches with digital face images.

Automatic sketch recognition algorithms have generally utilized the sketches drawn by artists with paper and pencil. Researchers have proposed several discriminative and generative algorithms for matching hand drawn sketches with digital images [1, 2, 3]. Discriminative algorithms extract invariant features such as Multiscale Local Binary Patterns [2] or Multiscale Circular Weber Local Descriptor [3] from sketch and digital face images for matching. Generative algorithms generate sketch representation of digital image or vice versa to reduce the difference between the two and then use the generated representation for matching [1].

Several law enforcement agencies across the world are now using softwares to generate the composites according to witness description. Some real world examples of composite and digital images are shown in Fig. 1[1]. Since the composite images are generally drawn based on witness description, it is not very accurate and also lack the detailed texture and minute facial features. Very limited research has been performed for automating this process. Han et al. [4] proposed a component based representation approach for matching composites with digital images. The algorithm extracts facial landmarks using active shape models and then features are extracted for every component using multiscale local binary patterns. The features extracted from corresponding components of both composite and digital images are matched followed by score fusion and normalization to generate the matching result. The algorithm also uses gender information as an indexing parameter.

[1]Images are taken from Harrisburg police case and Frontline Stories.

In this research, we propose a patch based framework that extracts multiscale circular patches around the fiducial features and then apply descriptors on these patches. The information obtained from individual components are individually insufficient to make a decision. Therefore, they are treated as weak classifiers and then combined using boosting. The performance of the proposed framework is evaluated using two descriptors and two boosting algorithms. The results computed on the PRIP composite face database shows that the algorithm shows promising results and outperforms a commercial-off-the-shelf (COTS) face recognition system.

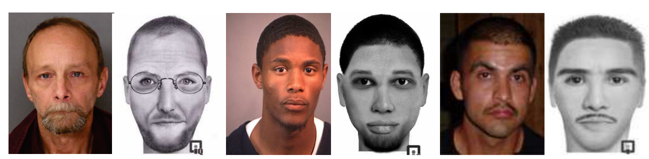

Fig. 1. Composite and digital face images corresponding to some real world cases.

2. PROPOSED MULTISCALE CIRCULAR FEATURE EXTRACTION AND BOOSTING ALGORITHM

Fig. 2 illustrates the steps involved in the proposed algorithm. The main intuition behind the proposed algorithm is that the corresponding digital face images and composites can vary in terms of exact feature size and shape but the local information around fiducial features remains unchanged. It relies on the observation that witnesses may not be able to provide exact details of the suspect but he/she is generally able to capture the key facial characteristics. The proposed algorithm utilizes the descriptor extracted from local patches to encode these key characteristics. The information obtained from these patches is combined using a boosting algorithm. The details of the algorithm are described below.

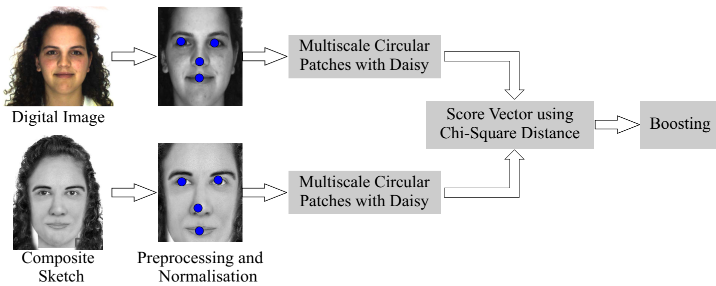

Fig. 2. Steps involved in the proposed composite sketch matching recognition algorithm.

1. *Preprocessing*: Digital images and composites vary in both intrinsic content and image properties. For example, digital images usually have background information, skin texture, and varying resolution whereas the composite sketches can have some markers (specially product logos). Therefore, a preprocessing procedure is used to normalize the digital images and composite sketches. If the input image is a color image (as in the case of digital), it is converted into gray scale and the facial region is detected using the Viola Jones face detector [5]. The detector also provides the eye and mouth coordinates which are then used for geometric normalization and affine registration for fixed eyes and mouth locations. The size of the preprocessed and normalized images are 300×350.

2. *Circular Patch Extraction*: Multiscale circular patches are extracted from both digital images and composite sketches.

    - First, the face detection algorithm is used to detect four points P_1, P_2, P_3, and P_4 i.e., two eyes and one mouth, and the centroid of the triangle they form, respectively. The centroid is a good heuristic for the tip of the nose.
    - Euclidean distance D between the eye points is measured.
    - A circular patch of radius $\frac{D}{4}$ centered at $P_1 = (x_1, y_1)$ is extracted.
    - Next, circular patches of radius $\frac{D}{4}$ centered at $(x_1 + \frac{D}{4}, y_1)$, $(x_1 + \frac{D}{2}, y_1)$, $(x_1 + \frac{3D}{4}, y_1), \cdots$ proceeding right until the image boundary are extracted.
    - Similarly, patches centered at $(x_1 - \frac{D}{4}, y_1)$, $(x_1 - \frac{D}{2}, y_1)$, $(x_1 - \frac{3D}{4}, y_1), \cdots$, proceeding left, centered at $(x_1, y_1 + \frac{D}{4}), \cdots$, proceeding up, and $(x_1, y_1 - \frac{D}{4}), \cdots$, proceeding down are extracted.
    - This process is then repeated for the remaining three points, P_2, P_3 and P_4.
    - Finally, the entire algorithm is repeated using patch radius of $\frac{D}{2}$ and $\frac{3D}{4}$.

3. *Feature Extraction*: Next step in the proposed algorithm is extracting features for the circular facial patches. Any holistic or texture features can be utilized to encode these patches. In this research, we explored three descriptors namely, EUCLBP [6], PHOG [7], and Daisy [8]. Extended Uniform Circular Binary Patterns (EUCLBP) [6] is a variant of LBP and has shown promising results for matching hand drawn sketches with digital images. Daisy descriptor [8] is invariant to illumination variations. This is required since the nature of a composite sketch image is inherently different from that of a digital image. Daisy descriptor is also robust towards geometric inconsistencies, and therefore provides a representation of the composite sketch image which can be matched with its digital counterpart accurately. Pyramid Histogram of Oriented Gradients (PHOG) [7] is a spatial information based descriptor and is used to extract local features for classification.

    For each patch, features are extracted from both digital images and composites. The corresponding patches of digital image and composite sketch are then matched using χ^2 distance metric and a match score s_i is generated. This provides a match score vector $S = \{s_1, s_2, \cdots, s_i\}$ for $i = 1, \cdots, N$ where N represents the number of patches.

4. *Identification using Boosting*: The first step in the proposed algorithm is utilizing the score vectors to make a decision. Since there is a lack of training samples, it is challenging to train a multiclass classifier. However, a two-class classifier can be trained in verification mode (1:1 matching) and it can be executed multiple times so that the identification results can be obtained (i.e., *identification in verification mode*). Further, each patch can be treated as a weak classifier which may not be sufficient individually but may provide improved performance in a boosting framework. Therefore, a boosting algorithm known as GentleboostKO [9] that can be trained with limited samples is used. During training, score vectors and class labels $\{genuine, impostor\}$ are provided as input to the boosting algorithm. The trained boosting classifier is then tested in the identification mode (by running verification experiments multiple times). For a given probe, all the gallery im-

Fig. 3. Sample images from the PRIP composite sketch database [4]. (a) Digital images, (b) composites generated using FACES by the caucasian user, (c) composites generated using FACES by the asian user, and (d) composites generated by IdentiKit.

ages that are classified as *genuine* are sorted using the $\sum s_i$ values to compute the identification rank.

3. EXPERIMENTAL RESULTS

The performance of the proposed algorithm is evaluated on the PRIP composite sketch database [4]. The PRIP composite sketch database is the only database that is publicly available and contains digital images and composites of 123 faces from the AR face database [10]. The composites are generated using two softwares: FACES [11] and Identikit [12]. To get the variation in composites across users, the composites from FACES software are generated by one asian and one caucasian user. The images from IdentiKit have been prepared by the asian user only. Fig. 3 shows digital images and composites from the PRIP database.

The performance is also compared with EUCLBP based sketch matching algorithm [6], commercial-off-the-shelf recognition system, and a modified version of the content based representation algorithm [4]. The modifications in the CBR algorithm include removing the gender information and normalization/preprocessing is performed using the approach discussed in our algorithm. For training, the database is split in two parts: 20% training and 80% testing. The results are reported with three times random cross validation and average identification accuracy along with Cumulative Match Characteristics (CMC) curves are reported. Some key observations and results are discussed below:

Table 1. Rank-10 identification accuracy (%) obtained on the PRIP composite sketch database [4].

Algorithm	Faces (Am)	Faces (As)	IdentiKit (As)
EUCLBP	10.6	13.0	14.6
Patch-PHOG	29.3	4.1	21.9
Patch-Daisy	30.9	13.8	25.2
COTS	10.6	6.5	8.1
Proposed	**45.8**	**20.2**	**33.7**

- CMC curves in Fig. 4 show that the proposed patch based Daisy features with boosting outperforms COTS, PHOG, and existing EUCLBP based matching algorithm.
- CMC curves in Fig. 4 also support the findings of Han et al. [4] that, for all experiments with different descriptors, composites created by the Caucasian user provide higher rank-10 identification accuracy compared to the Asian user. Similarly, composites created using FACES software yields higher accuracy than IdentiKit.
- Among the three descriptors, Daisy descriptor applied on all the patches and then sum of all χ^2 distances used for decision making outperforms EUCLBP and PHOG. Further, Daisy descriptor also outperforms COTS (same software used by Han et al. [4]) by a significant margin.
- The performance of Gentleboost is compared with other boosting approaches namely Adaboost, AdaboostKO, and Gentleboost [13] and SVM with different kernels and parameters. As shown in Table 1 and Fig. 5, under the same experimental protocol, GentelboostKO outperforms other classifiers by at least 6%. The second best classifier is linear SVM which yields rank-10 identification accuracy of 40%.
- Comparison with the modified CBR approach under the same experimental protocol shows that the proposed algorithm is about 7% better at rank-10 identification accuracy. t-test at 95% confidence interval further suggests that both the algorithms are significantly different.
- The concept of extracting patches and then boosting them to create a stronger classifier shows improved performance; however a larger database with more testing samples are required to make a stronger case.

4. CONCLUSION

Matching composite sketches with digital photos is an interesting law enforcement problem and very limited research has been undertaken. This paper presents an algorithm for multiscale patch based feature extraction and boosting approach

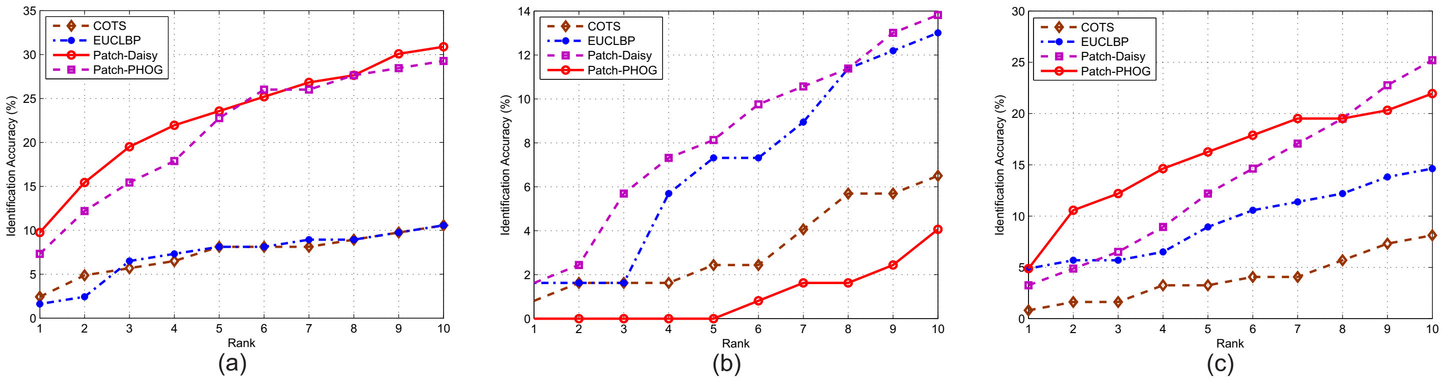

Fig. 4. CMC curves of the existing and proposed algorithms on the PRIP composite sketch database: (a) Faces (Caucasian user), (b) Faces (Asian user), and (c) IdentiKit (Asian user).

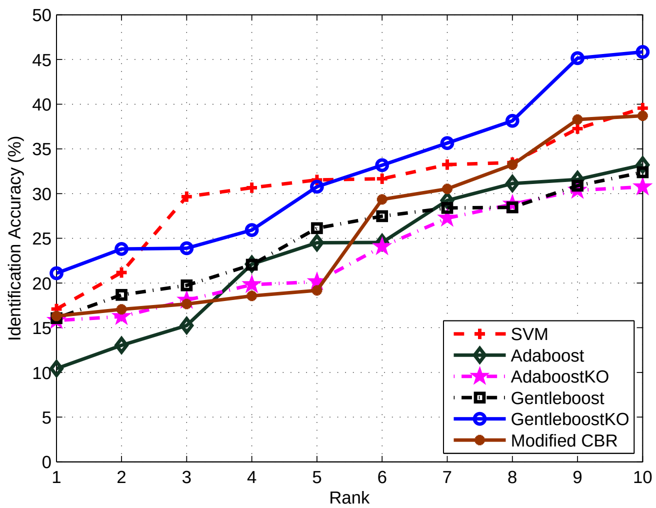

Fig. 5. Comparing the performance of the proposed algorithm with different classifiers and modified CBR algorithm.

for matching composites with digital images. The results suggest that the multiscale patch extraction and boosting helps in improving the identification accuracy. However, we believe that more research is required to build a more sophisticated algorithm but it requires preparing a large composite sketch and digital image database.

5. ACKNOWLEDGEMENT

The authors would like to thank Dr. Anil Jain, MSU for providing the PRIP composite sketch database. The authors also thank the reviewers for their feedback.

6. REFERENCES

[1] X. Wang and X. Tang, "Face photo-sketch synthesis and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 11, pp. 1955–1967, 2009.

[2] B.F. Klare, Z. Li, and A.K. Jain, "Matching forensic sketches to mug shot photos," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 3, pp. 639–646, 2011.

[3] 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*, vol. 7, no. 5, pp. 1522 –1535, 2012.

[4] 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*, vol. 8, no. 1, pp. 191–204, 2013.

[5] P. Viola and M.J. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.

[6] 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*, 2010.

[7] A. Bosch, A. Zisserman, and X. Munoz, "Representing shape with a spatial pyramid kernel," in *Proceedings of the International Conference on Image and Video Retrieval*, 2007, pp. 401–408.

[8] E. Tola, V. Lepetit, and P. Fua, "Daisy: An efficient dense descriptor applied to wide-baseline stereo," *IEEE*

Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 5, pp. 815 – 830, 2010.

[9] L. Wolf and I. Martin, "Robust boosting for learning from few examples," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005, vol. 1, pp. 359 – 364.

[10] A.R. Martinez and R. Benavente, "The AR face database," Tech. Rep., 1998, Computer Vision Center, Technical Report.

[11] "Faces 4.0, iq biometrix," http://www.iqbiometrix.com, 2011.

[12] "Identi-kit, identi-kit solutions," http://www.identikit.net/, 2011.

[13] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: A statistical view of boosting," *The Annals of Statistics*, vol. 38, no. 2, pp. 337–374, 2000.