A Sketch Based Face Matching Approach for IR and Optical Images Using CFDA

Sujata M. Kullarkar, Anjali R. Mahajan
Department of Computer Science & Engineering
Priyadarshini Institute of Engineering & Technology
Nagpur, India
e-mail: sujatakullarkar@gmail.com
armahajan@rediffmail.com

Abstract— In biometrics research and industry, it is critical yet a challenge to match infrared face images, optical face images to sketches. The most challenging issue in heterogeneous face recognition is that face images associated with the same person but taken with different imaging devices may be mismatched due to the great discrepancy between the different image modalities (optical, infrared and sketch), which is referred to as modality gap. The major difficulty lies in the fact that a great discrepancy exists between the infrared face image, corresponding optical face image and sketch image because they are captured by different imaging devices. In this paper our focus is on the approach which defines cross modality face reorganization problems such as sketch-photo and high-low resolution face matching. In this approach, a new learning-based face descriptor is first proposed to extract the common features from heterogeneous face images (infrared face images and optical face images and sketch images), and an effective matching method is then applied to the resulting features to obtain the final decision.

Keywords - *Modality Gap, Sketches, Optical Face Image, Infrared Face Images, Heterogeneous Face Recognition.*

*****

I. Introduction

It is critical yet to match face images with different modalities. The main difficulty lies in the fact that great discrepancy exists between different imaging devices like infrared and optical imaging devices and sketches. Sketches are either drawn by artists or created with computer software following the verbal description provided by an eyewitness or the victim. The traditional method of identifying suspects is slow and tedious and may not lead to apprehension of the suspect. Hence, there is a need for a method that can automatically and quickly match facial sketches to gallery of mug shot databases. Our proposed method is useful in identifying not only the sketch images with face sketches but also the suspects that deals with different image modalities. In biometrics research and industry, it is critical yet a challenge to match infrared face images and optical face images to sketches. The major difficulty lies in the fact that a great discrepancy exists between the infrared face image and corresponding optical face image because they are captured by different devices (optical imaging device and infrared imaging device).

Moreover, traditional optical imaging devices require appropriate illumination conditions to work properly, which is difficult to achieve satisfactorily in practical face recognition applications. To combat low illumination at nights or indoors, infrared imaging devices have been widely applied to many automatic face recognition (ARF) systems. The task of infrared-based ARF systems is to match a probe face image taken with the infrared imaging device to a gallery of face images taken with the optical imaging device, which is considered to be an important application of heterogeneous face recognition [1] (also known as *cross-modality face recognition*). The most challenging issue in heterogeneous face recognition is that face images associated with the same person but taken with different devices may be mismatched due to the great discrepancy between the different image modalities (optical, infrared and sketch), which is referred to as *modality gap*. The infrared photos are usually blurred, low contrast, and have significantly different gray distribution compared to the optical photos. In the discriminant analysis common feature (CFDA) approach, a new learning-based feature descriptor is first developed to learn a set of optimal hyper-planes to quantize continuous vector space into discrete partitions for common feature extraction, and an effective discriminant analysis technique is then applied for feature classification. We conduct extensive experiments on two large and challenging optical, infrared and face sketch datasets to investigate the effectiveness of our new approach.

II. Overview of Existing Methods

To combat the great modality gap, a variety of algorithms have been proposed in the literature which can be summarized as belonging to the following three categories.

The first category of approaches is to convert an image from one modality to the other by synthesizing a pseudo-image from the query image such that the matching process can be done within the same modality. For example in order to perform the match between two image modalities say face image and face sketch, there is a need to convert the face image into sketch then perform the matching between two face sketch depending on their feature measures. For instance, [2] applies a holistic mapping to convert a photo image into a corresponding sketch image, and in [3]–[5] the authors used local patch-based mappings to convert images from one modality to the other for sketch photo recognition. In [6] the author synthesized VIS face images from NIR face images with pose rectification.

The second category of approaches is to design an appropriate representation that is insensitive to the modalities of images. For example, [7] used SIFT feature descriptors and multi-scale local binary patterns to represent both the sketch and photo images. Zhang et al. [8] proposed a learning based algorithm to capture discriminative local face structures and effectively match photo and sketch. In [9], author designed a multi-scale common feature descriptor to combat the large intra-class difference incurred by the modality (VIS-NIR) difference.

The third category of approaches is to compare the heterogeneous images on a common subspace where the modality difference is believed to be minimized. For example, in [10] author applied the Bilinear Model (BLM) by Singular Value Decomposition (SVD) to develop a common content (associated with identity) space for a set of different styles (corresponding to modalities). In [11], author used the Canonical Correlation Analysis (CCA) technique to construct a common subspace where the correlations between infrared and optical images can be maximized. In [12], author applied the CCA to cross-pose face recognition. In [13], author applied the Partial Least Squares (PLS) method to derive a linear subspace in which cross-modality images are highly correlated, while at the same time preserving variances more effectively than the previous CCA method. In [16], author proposed an effective subspace learning framework called coupled discriminant analysis for heterogeneous face recognition. In [17], a generic HFR framework was proposed in which both probe and gallery images are represented in terms of non-linear kernel similarities to a collection of prototype face images to enhance heterogeneous face recognition accuracy.

The probe face image has to go through three different stages. 1) Face Detection 2) Feature Extraction 3) Evaluation of Faces. In the first step face is detected from the input face image based on the key features that we have decided for feature detection purpose. In the second step the decided features has to be extracted from detected face and in the third step based on the measure of features extracted from the detected face, matching process is applied to evaluate the matching face images for the corresponding input probe image. Most of the techniques mentioned in the literature differ in the feature extraction stage. In [2] Eigen Transformation method is applied on the input face image detected face in order to separate shape and texture information. By transforming the photo image into sketch significantly thus allow effective matching between the two and then Bayesian classifier is used to recognize the probing sketch from the synthesized pseudo sketches. In [3] author used component based face sketch recognition, where first a full Sobel edge detection filter is used, then the resultant image is enhanced using histogram equalization. After that high-pass filtering is applied to obtain the clean image. In [4] author proposed a new face descriptor called Histogram of Average Oriented Gradients (HAOG). HAOG is inspired by the fact that orientations of stronger gradients (gradients of facial components with high contrast, e.g. eyes, eyebrows, ears, mouth and nose) are more modality invariant than weaker gradients (gradients of fine texture, wrinkles and shadows with low contrast). Thus, the modality difference between face photos and sketches can be reduced by emphasizing features extracted from stronger gradients. HAOG is different from Histogram of oriented gradients (HOG) [2] which is computed entirely on images with both fine and coarse texture [2-6], whereas HAOG is extracted only on coarse texture where we believe is more robust against modality difference. In [4] author proposed a system for forensic face sketch recognition by a computer vision approach like Two-Dimensional Discrete Cosine Transform (2D-DCT) and the Self-Organizing Map (SOM) Neural Network simulated in MATLAB.

III. Proposed Methedology

In our new approach we are dealing with three different modalities infrared face image, optical face image and sketch image. To combat the modality gap we proposed a new approach called Common Feature Descriminant Analysis where a new learning based face descriptor is first developed where the vectors of continuous space is converted to discrete code representation in order to convert the image into an encoded image. Vectors of continuous space is converted to the decimal code with the help of pixel normalization techniques like K-min and Random Forest Algorithms where, center pixel value is normalized with respect the neighboring pixels.

Vector quantization is an effective technique that has been widely used to create discrete code representations for object recognition. An Image can be turned into an encoded image by converting each pixel into a specific code using vector quantization technique. We design a hyperplane based encoding method for effective feature representation for heterogeneous face images.

In feature extraction stage we use our CFDA approach for image encoding purpose. For image encoding purpose the image has to go through pipeline for feature extraction. For each pixel, we first sample its five d-neighbor (Radii = d) pixels for each direction (the figure shows one of four directions, with blue arrows), and then subtract the center pixel value. Finally the centered vector is normalized into the unit L2-norm to form the associated pixel vector of that direction. Each pixel is associated with four vectors, forming four sets of training vectors that are used to train four encoders. Each encoder consists of two sets of mutually orthogonal hyper-planes (we only show one for illustration), which divide the vector space into four partitions. Vectors of each direction are encoded into a 2-bit value, according to the partition in which the vector lies (i.e. 00 for the first partition, 01 for the second partition, and so forth). Finally, the four 2-bit values are concatenated to form an 8-bit value that will be converted into a decimal value (from 0 to 255) as the code. With the face image encoded the image we can use densely sampling technique in order to extract the features. For this the whole encoded image is divided into a set of overlapping patches with the size c x c pixels (the step between adjacent patches is s). Then compute the histogram over each patch of the frequency of each code occurring which gives a feature vector for each patch. Concatenate the outputs of each patch into a long vector to form the final face feature. The matching framework involves two levels of subspace analysis. In the first level, the large feature vector is first divided into multiple segments of smaller feature vectors. Discriminant analysis is performed separately on each segment to extract the discriminant features. The goal for the first level is to generate more discriminative projections to reduce intraclass variations and avoid over-fitting. In the second level, projected features from all the segments are then combined, with PCA for efficient recognition.

The CFDA approach is proposed specifically for handling the optical-infrared face recognition problem. In the feature extraction stage, a learning-based feature descriptor is developed to maximize the correlations between the optical face images and corresponding infrared images. In this way, the modality gap between the two kinds of face images can be significantly reduced; hence, it is expected that the resulting features will be well-suited to the optical-infrared face recognition problem.

Our new feature descriptor differs significantly from State-of-the-art descriptors, such as the widely used HOG and LBP in the literature. Instead of encoding the images using a handcrafted encoding scheme, our feature descriptor learns a new encoding scheme to encode the common micro-structure of both the optical and infrared face images for effective feature representation. Our experimental results also support the effectiveness of our new descriptor over state-of-the-art descriptors.

Our new feature descriptor also differs significantly from the CITE (coupled information tree encoding) in [8]. The major difference between them is summarized as follows. (1) CITE is inherently a tree based encoding method, while our feature descriptor is inherently a binary encoding scheme. (2) Unlike CITE, we encode a pixel with four directions to make full use of the geometry information. This also reflects the significant difference between them.

IV. Conclusion

In this proposed paper we introduced a new approach called common feature discriminant analysis (CFDA), for matching infrared face images to optical face images and sketches. In CFDA, we will first develop a new descriptor to effectively represent optical, infrared face images and sketches to reduce the modality gap, and then a two-level matching method will be subsequently applied for fast and effective matching as a part of our proposed system. Extensive experiments on two large and challenging optical-infrared face datasets will be used to find the significant improvement of our new approach over the state-of-the-art.

V. Acknowledgment

I sincerely acknowledge the teaching staffs of the department and authors of the corresponding references for their guidance and being an inevitable part of this review article.

References

[1] *Zhifeng Li, Senior Member, IEEE, Dihong Gong, Yu Qiao, Senior Member, IEEE, and zacheng Tao, Senior Member*," Common Feature Discriminant Analysis for Matching Infrared Face Images to Optical Face Images," june 2014.

[2] *Xiaoou Tang and Xiaogang Wang,"* Face Sketch Synthesis and Recognition", IEEE International Conference on Computer Vision, 2003.

[3] *Xiaoou Tang, Senior Member, IEEE, and Xiaogang Wang, Student Member, IEEE,"*Face Sketch Recognition,"IEEE Transactions On Circuits And Systems For Video Technology, Vol. 14, No. 1, January 2004.

[4] *Setiawan Hadi, Iping Supriana Suwardi, and Farid Wazdi,* "Face sketch Recognition System to Support Security Investigation", Cryptology and Information Security Conference, 2005.

[5] *Hamed Kiani Galoogahi, Terence Sim*, "Inter-modality Face sketch Recognition", IEEE International Conference On Information Security, 2011.

[6] *Vineet Srivastava*, "Forensic Face Sketch-Photo Recognition Using Computer Vision", International Journal on Recent and Innovation Trends in computing and Communication, MAR 2013.

[7] *X. Tang and X. Wang,* "Face sketch recognition," IEEE Trans. Circuits Syst. Video Technol., vol. 14, no. 1, pp. 50–57, Jan. 2004.

[8] *B. Xiao, X. Gao, D. Tao, Y. Yuan, and J. Li,* "Photo-sketch synthesis and recognition based on subspace learning," Neurocomputing, vol. 73, pp. 840–852, Jan. 2010.

[9] *X. Wang and X. Tang, "Face photo-sketch synthesis and recognition*," IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 11, pp. 1955–1967, Nov. 2009.

[10] *Q. Liu, X. Tang, H. Jin, H. Lu, and S. Ma,* "Nonlinear approach for face sketch synthesis and recognition," in Proc. IEEE Comput. Soc. Conf. CVPR, Jun. 2005, pp. 1005–1010.

[11] *P. Xiong, L. Huang, and C. Liu,* "A method for heterogeneous face image synthesis," in Proc. 5th IAPR ICB, Apr. 2012, pp. 1–6.

[12] *B. Klare, Z. Li, and A. K. Jain,* "Matching forensic sketches to mug shot photos," IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 3, pp. 639–646, Mar. 2011.

[13] *W. Zhang, X. Wang, and X. Tang,* "Coupled information-theoretic encoding for face photo-sketch recognition," in Proc. IEEE Conf. CVPR, Jun. 2011, pp. 513–520.

[14] *S. Liu, D. Yi, Z. Lei, and S. Z. Li,* "Heterogeneous face image matching using multi-scale features," in Proc. 5th IAPR ICB, Apr. 2012,pp. 79–84.

[15] *J. B. Tenenbaum and W. T. Freeman,* "Separating style and content with bilinear models," Neural Comput., vol. 12, no. 6, pp. 1247–1283, 2000.

[16] *D. Yi, R. Liu, R.-F. Chu, Z. Lei, and S. Z. Li,* "Face matching between near infrared and visible light images," in Proc. Int. Conf. Adv.Biometrics, 2007, pp. 523–530.

[17] *A. Li, S. Shan, X. Chen, and W. Gao,* "Maximizing intra-individual correlations for face recognition across pose differences," in Proc. IEEE Conf. CVPR, Jun. 2009, pp. 605–611. A. Sharma and D. W. Jacobs, "Bypassing synthesis: PLS for face recognition with pose, low-resolution and sketch," in Proc. IEEE Conf. CVPR, Jun. 2011, pp. 593–600.

[18] *W. Yang, D. Yi, Z. Lei, J. Sang, and S. Z. Li,* "2D–3D face matching using CCA," in Proc. 8th IEEE Int. Conf. Autom. Face Gesture Recognit., Sep. 2008, pp. 1–6.

[19] *S. Liao, D. Yi, Z. Lei, R. Qin, and S. Li,* "Heterogeneous face recognition from local structures of normalized appearance," in Proc. 3rd ICB, 2009, pp. 209–218.

[20] *Z. Lei, S. Liao, A. K. Jain, and S. Z. Li,* "Coupled discriminant analysis for heterogeneous face recognition," IEEE Trans. Inf. Forensics Security, vol. 7, no. 6, pp. 1707–1716, Dec. 2012.

[21] *B. Klare and A. K. Jain,* "Heterogeneous face recognition using kernel prototype similarities," IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 6, pp. 1410–1422, Jun. 2013.

[22] *F. Nicolo and N. A. Schmid,* "Long range cross-spectral face recognition: Matching SWIR against visible light images," IEEE Trans. Inf. Forensics Security, vol. 7, no. 6, pp. 1717–1726, Dec. 2012. Z. Lei, C. Zhou, D. Yi, A. K. Jain, and S. Z.

ISSN: 2321-8169
132– 135

Li, "An improved coupled spectral regression for heterogeneous face recognition," in Proc. 5th IAPR ICB, 2012, pp. 7–12.

[23] *F. Jurie and B. Triggs,* "Creating efficient codebooks for visual recognition," in Proc. 10th IEEE ICCV, vol. 1. Oct. 2005,pp.604–610.

[24] *Huda Abdulaali Abdulbaqi, Ghazali Sulong, Soukaena Hassan Hashem,* "A Sketch Based Image Retrieval", 2012.S. Z. Li, "Heterogeneous face biometrics," in Encyclopedia of Biometrics. New York, NY, USA: Springer-Verlag, 2009.

[25] *X. Wang and X. Tang, "Face photo-sketch synthesis and recognition,"* IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 11, pp. 1955–1967, Nov. 2009.

[26] *Z. Lei, S. Liao, A. K. Jain, and S. Z. Li,* "Coupled discriminant analysis for heterogeneous face recognition," IEEE Trans. Inf. Forensics Security, vol. 7, no. 6, pp. 1707–1716, Dec. 2012.

[27] *B. Klare and A. K. Jain,* "Heterogeneous face recognition using kernel prototype similarities," IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 6, pp. 1410–1422, Jun. 2013.

[28] *F. Nicolo and N. A. Schmid,* "Long range cross-spectral face recognition: Matching SWIR against visible light images," IEEE Trans. Inf. Forensics Security, vol. 7, no. 6, pp. 1717–1726, Dec. 2012.

[29] *Z. Lei, C. Zhou, D. Yi, A. K. Jain, and S. Z. Li,* "An improved coupled spectral regression for heterogeneous face recognition," in Proc. 5th IAPR ICB, 2012, pp. 7–12.