Inter-modality Face Sketch Recognition

Hamed Kiani Galoogahi
School of Computing
National University of Singapore
Singapore
hkiani@comp.nus.edu.sg

Terence Sim
School of Computing
National University of Singapore
Singapore
tsim@comp.nus.edu.sg

Abstract—**Automatic face sketch recognition plays an important role in law enforcement. Recently, various methods have been proposed to address the problem of face sketch recognition by matching face photos and sketches, which are of different modalities. However, their performance is strongly affected by the modality difference between sketches and photos. In this paper, we propose a new face descriptor based on gradient orientations to reduce the modality difference in feature extraction stage, called Histogram of Averaged Oriented Gradients (HAOG). Experiments on CUFS database show that the new descriptor outperforms the state-of-the-art approaches.**

Keywords-face sketch recognition; inter-modality; histogram of oriented gradients.

I. INTRODUCTION

Face sketch recognition has received significant attention in recent years, due to its vital role in law enforcement. In many crime scenes, the only available information is the verbal description of suspect's face given by eye-witness; which can be used by police artist to draw a face sketch. This sketch is then used to recognize the suspect's identity. Generally speaking, face sketch recognition can be defined as matching a query face sketch against a gallery of face photos with known identities.

The major challenge of face sketch recognition is matching images of different modalities [1]. Basically, a face photo is captured by a digital camera, while a face sketch is drawn by artist with different level of information. Even for the same human subject, the face photo and its sketch might be different. The face shape might be exaggerated by artist or texture might be lost or replaced by some artistic rendering. This problem will be more exacerbated for forensic investigations, when the eye-witness cannot exactly remember the detail of suspect's face.

To date, various methods have been proposed for face sketch recognition which can be categorized into two general approaches: intra-modality and inter-modality approaches. Intra-modality approaches try to synthesize pseudo photo (sketch) from input sketch (photo) to perform face sketch recognition in the same modality (sketch or photo). Indeed, photo/sketch synthesis is an inevitable preprocessing step in these approaches. However, the performance of these methods is highly relied on the accuracy of photo/sketch synthesis, which might be even harder than recognition problems.

On the other hand, face sketch recognition is directly performed by inter-modality approaches in different modalities by extracting discriminative features which are invariant to photo and sketch modalities. However, most of these approaches use some common features which are not originally designed to address the problem of modality difference. Therefore, a modality-invariant feature is eagerly needed for face sketch recognition to particularly deal with the presence of modality difference between face photos and sketches.

In this paper, we propose a new face descriptor which called Histogram of Averaged 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.

II. RELATED WORKS

Most of the existing works synthesize pseudo photos (sketches) form input sketches (photos) into a same modality which is followed by intra-modality face recognition. Tang and Wang [8,9] proposed a face sketch synthesize and recognition method by applying Eigen-transformation on the entire image of a given face. Similarly, Liu et al. [10] proposed a patch based nonlinear face sketch synthesis and recognition method inspired by local linear embedding. This approach performs Eigen-transformation on local patches instead of the entire image.

Later, Wang and Tang [11] improved the method of [10] by modeling the spatial relation of local patches using multi scale Markov random filed (MRF). As another approach based on Markov model, Zhong et al. [12] proposed a method based on embedded hidden Markov model (E-HMM) and selective ensemble strategy to model the nonlinear relationship between photos and sketches.

A major limitation of the above methods is that the accuracy of these works is highly dependent on the results of photo-sketch synthesis, i.e. imperfect synthesis results can lead to poor recognition. Therefore, some recent works have focused to reduce the modality difference in feature extraction stage instead of transforming into same modality.

The first feature based method was proposed by Klare and Jain [13]. In this approach, dense SIFT descriptors [14] are directly extracted from local patches to reconstruct a holistic image representation. For each local patch, a 128-dimentional SIFT descriptor is calculated. The holistic image representation is obtained by accumulating the local SIFT descriptors. Direct sketch-photo matching was performed by a simple 1-NN classifier.

Moreover, Klare et al. [15] proposed local feature based discriminant analysis (LFDA) to match forensic sketches to mug shot photos. Photos and sketches are represented by two different types of features: SIFT descriptors and multi local binary patterns (MLBP). Then, multiple discriminant projections on partitioned vectors of the features are used to extract discriminative features. Despite the high accuracy achieved by this method, the modality difference between sketches and photos has not been solved by LFDA. Since, the SIFT and MLBP are not robust against modality difference in face sketch recognition problem [1].

Recently, Zhang et al. [1] presented a new face descriptor based on coupled information-theoretic encoding to extract modality-invariant descriptor. In this work, coupled information-theoretic projection was introduced to maximize the mutual information between the encoded photo and sketch of same subject. This method is the state-of-the-art in face sketch recognition.

III. PROPOSED METHOD

Our method is based on representing face photos and sketches by a new gradient orientations based descriptor, Histogram of Averaged Oriented Gradients (HAOG). The goal of HAOG is to reduce the modality difference between face photos and sketches in feature extraction stage.

A. Motivation

The modality gap (difference) between face sketch and photo is caused by the difference of visual information which can be perceived from face sketch and photo. The visual information of face consists of both face shape and texture.

A face has an oval shape comprising of components such as eyebrows, eyes, lips, nose, chin and ears which are coherently related to each other in a spatial configuration. The visual information of face shape refers to general shape of face and spatial configuration of components (e.g. mouth is below nose, eyebrows are above eyes, etc.), not the size of facial components, width/height of face and the exact distance among components which can be used as similarity or dissimilarity measures for face sketch matching. Therefore, the amount of shape information extracted from face sketches is same as those extracted from photos. As a result, face shape is not involved in modality gap.

On the other hand, visual information of face texture is related to face appearance rather than face shape. The face appearance contains coarse and fine textures belong to facial components and facial skin, respectively. The coarse texture represents the boundaries of facial components with high contrast. Contrarily, the fine texture represents low contrast details of face skin including flaws, moles, wrinkles as well as any shadow/reflection caused by varying lighting condition.

From texture point of view, the modality gap can be explored for coarse and fine textures separately. Coarse texture belongs to boundaries of face components; which are vital for artists to draw face sketches. Moreover, these boundaries make the face components in face photos to be distinguishable from face skin. Therefore, they are steadily present in face sketches and photos. Thus, the amount of modality gap is not affected significantly by coarse texture.

However, the modality gap caused by fine texture is emphasized in both sketches and photos. Fine texture in face photo is represented by low contrast details which might be lost in corresponding face sketch. In addition, fine texture rendered by artists in face sketch might be biased by the artist who draws the sketch or even the sketching style/tool which he or she uses. Therefore, a face sketch might contain rendered textures which are not compatible with those in corresponding face photo. In the other words, the visual information represented by the fine texture of face sketches is much different from those of face photos. That leads to high amount of modality gap.

The discriminative power of image features extracted from gradient orientations (e.g. HOG) is highly dependent on presence of modality gap. Moreover, we discussed that the modality gap between face sketches and photos is principally caused by fine texture not coarse texture. Therefore, it can be concluded that the robustness of face features against different modalities can be increased by extracting features from gradient orientations of coarse texture not fine texture.

One solution to extract face features based on coarse texture is to use edge-preserving smoothing operators, e.g. Bilateral filters, to decompose face photos and sketches into coarse and fine layers [16]. Then, the coarse layer which contains coarse texture can be used to represent face photos and sketches by features such as HOG. However, there are two major challenges that make using Bilateral operators infeasible for this purpose. First, there is not a clear and exact definition of layers scale to separate fine and coarse layers precisely. Second, there are some range and domain parameters which have to be tuned precisely to have control over sketches/photos decomposition [16].

Another alternative is to use both fine and coarse textures in feature extraction process, but emphasizing coarse texture much more than fine texture while computing orientation features. Obviously, coarse texture has much greater gradient magnitudes in compare to fine texture, due to higher contrast. Therefore, emphasizing coarse texture can be typically preformed by voting squared magnitudes of gradient into histogram of orientations. Moreover, we will further show that squaring the gradient magnitudes is equivalent to doubling the angles of gradient orientations; which can be used to compute averaged orientation of gradients. Thus, the squared magnitudes and doubled orientations can be used to form histogram of averaged oriented gradients that emphasizes (weakness) coarse (fine) texture in feature extraction stage.

B. HAOG Descriptors

Given a gray level image *I(x,y)*, gradient vectors can be calculated by taking the partial derivatives of image intensity in Cartesian coordinates as:

$$\begin{bmatrix} g_x(x,y) \\ g_y(x,y) \end{bmatrix} = \begin{bmatrix} \frac{\partial I(x,y)}{\partial x} \\ \frac{\partial I(x,y)}{\partial y} \end{bmatrix} \quad (1)$$

For gradient vectors given by $[g_x, g_y]^T$, gradient magnitude ρ and orientation $\varphi \in [-\pi, \pi)$ are defined as:

$$\rho = (g_x^2 + g_y^2)^{0.5} \quad (2)$$

$$\varphi = \tan^{-1}\left(\frac{g_y}{g_x}\right) \quad (3)$$

Since local gradients around each pixel might be opposite to each other, calculating the averaged oriented gradient $\overline{\varphi}$ by directly averaging the orientation angles at local neighborhood of each pixel is infeasible; the opposite gradients at both sides of an edge are likely to cancel each other. To overcome this problem, Kass and Witkin [17] proposed to double the gradient angles before averaging. Since, if φ and $(\varphi + \pi)$ be two opposite orientations, after doubling, 2φ and $(2\varphi + 2\pi) = 2\varphi$ point in the same direction. Practically, 2φ is the orientation of squared gradient vector $[g_{sx}, g_{sy}]^T$ with magnitude ρ^2. According to trigonometric identities [18]:

$$\begin{bmatrix} g_{sx} \\ g_{sy} \end{bmatrix} = \begin{bmatrix} \rho^2 \cos 2\varphi \\ \rho^2 \sin 2\varphi \end{bmatrix} = \begin{bmatrix} \rho^2(\cos^2\varphi - \sin^2\varphi) \\ \rho^2(2\sin\varphi\cos\varphi) \end{bmatrix} = \begin{bmatrix} g_x^2 - g_y^2 \\ 2g_y g_x \end{bmatrix} \quad (4)$$

The average of squared gradient for each pixel in a local neighborhood by a window of size *W* can be calculated as:

$$\begin{bmatrix} \bar{g}_{sx} \\ \bar{g}_{sy} \end{bmatrix} = \begin{bmatrix} \sum_W g_{sx} \\ \sum_W g_{sy} \end{bmatrix} = \begin{bmatrix} \sum_W (g_x^2 - g_y^2) \\ \sum_W 2g_y g_x \end{bmatrix} \quad (5)$$

thus, the average of doubled gradient orientation and squared magnitude for each pixel are given by:

$$\overline{\varphi} = \tan^{-1}\left(\frac{\bar{g}_{sy}}{\bar{g}_{sx}}\right), \overline{\varphi} \in [-\pi, \pi) \quad (6)$$

$$\overline{\rho} = \sum_W (g_{sy}^2 + g_{sx}^2)^{0.5} = \sum_W \rho^2 \quad (7)$$

Given an image *I* (sketch or photo) in gray level, the HAOG descriptor is computed as follows. First, the image is divided into a set of overlapped local patches by sliding a window of size $s \times s$ through the image. This process starts from the upper left corner of the image and scans the whole image by moving the sliding window δ pixels each time.

Suppose that the set of local patches denoted by $P = \{p_i, 1 \le i \le N\}, N = |P|$. For each pixel in local patch $p_i \in P$, we calculate the averaged orientation $\overline{\varphi}$ and magnitude $\overline{\rho}$ using (1-7). The range of averaged orientations $[-\pi, \pi)$, is quantized into *b* bins. The histogram of averaged orientations for patch p_i, h_i, is computed by accumulating the averaged magnitudes $\overline{\rho}$ of pixels whose averaged orientations fall within each of the possible *b* bins. The concatenation of the histograms h_i represents the final HAOG descriptor, $\mathcal{H}(I)$. Note that each histogram h_i is normalized by its *L2*-norms before concatenation.

Let $I = \{\iota_j, 1 \le j \le M\}$ denotes a set of face photos, $\iota_j: \mathbb{R}^2 \to \mathbb{R}$. Given a query face sketch $\varsigma: \mathbb{R}^2 \to \mathbb{R}$, face sketch recognition is performed by:

$$\hat{\iota_j} \leftarrow \underset{I=\{\iota_j, 1 \le j \le M\}}{\operatorname{argmin}} \chi^2\left(\mathcal{H}(\iota_j), \mathcal{H}(\varsigma)\right) \quad (8)$$

where $\mathcal{H}(.)$ returns the HAOG descriptor and χ^2 is *Chi-Square* distance between two histograms defined as:

$$\chi^2(x,y) = \sum_k \frac{(x_k - y_k)^2}{(x_k + y_k)} \quad (9)$$

Indeed, reducing modality gap by emphasizing strong gradients of coarse texture is performed in two stages: computing the average of gradient orientations and histograms voting. According to (5) and (6), averaged oriented gradients are calculated by summation of squared gradients in a local neighborhood of each pixel. This has the effect that strong gradients in pixel's neighborhood have higher influence in calculating averaged gradient orientations than weaker gradients. In addition, the averaged squared magnitudes of gradients, (7), assign higher votes to the gradients with stronger magnitudes rather than those with weaker magnitudes.

Fig. 1 illustrates that squaring gradient magnitudes corresponded to doubled orientations emphasizes (weakens) stronger (smaller) gradients. The first column represents a face photo (top) and its corresponding sketch (bottom) drawn by an artist based on verbal descriptions by an eye-witness [19]. The original gradient magnitudes and the corresponding average of squared magnitudes are represented in second and third columns, respectively. According to Fig. 1, the second and third columns are similar in that both represent the strong gradients of coarse texture. However, the magnitudes of fine texture are significantly reduced in the latter. Magnitudes are scaled in [0,255] for display purpose.

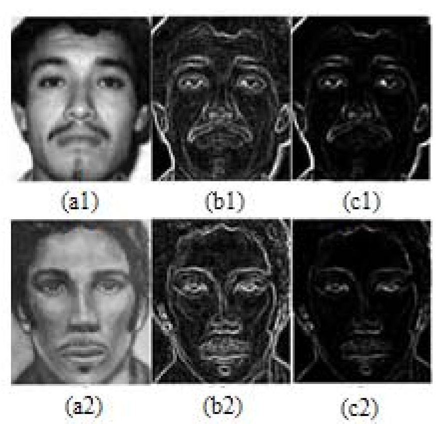

Figure 1. (a1) Face photo, (a2) Face sketch, (b1,b2) Gradient magnitudes of (a1,a2), Squared gradient magnitudes of (a1,a2).

Fig. 2 shows three local patches (size of 32×32 pixels) which are randomly selected from a face photo and its corresponding sketch. All local patches contain both coarse and fine textures. HAOG and HOG descriptors for each pair of local patches in first row are represented in second and third rows of Fig. 2 (b,c,d), respectively. Note that sketch and photo histograms in HOG and HAOG are represented by red and blue colors, respectively. The *Chi-Square* distance χ^2, (9), is used to evaluate the dissimilarity of HOG and HAOG descriptors extracted from corresponding pairs of local patches. According to Fig. 2, χ^2 distances between HAOG descriptors of corresponding local patches are significantly smaller than those of the HOG descriptors. That indicates the modality gap is remarkably reduced by HAOG descriptors.

IV. EXPERIMENTS AND RESULTS

In order to demonstrate the effectiveness of HAOG for face sketch recognition, we conducted two different experiments using standard CUHK face sketch database (CUFS) [11]. This database contains 606 pairs of corresponding photos/sketches of 606 different subjects. The photos are selected form three different face databases: (1) 188 faces from the Chinese University of Hong Kong (CUHK) student database, (2) 123 faces from the AR database [20], and (3) 295 faces from the XM2VTS database [21]. All photos in CUFS database are frontal face image with neutral expression taken under a normal illumination condition. For each photo, there is a sketch drawn by an artist by looking at its corresponding photo. In order to align all the face images together, the face images are translated, rotated and scaled such that the two eye centers of all images are fixed at positions (125,75) and (125,125). All images are then cropped to a same size of 250×200 pixels. Some examples of face photo/sketch pairs are shown in Fig. 3.

The first experiment was designed to compare the discriminative power of HAOG and HOG descriptors based on Fisher criterion [22]. In addition, the performance of face sketch recognition by HOG and HAOG descriptors is evaluated in this experiment. Feature extraction and matching are similar to that presented in Section 3. We directly use gradient magnitudes ρ and orientations φ, (2) and (3), to compute HOG descriptors. While, HAOG descriptors are calculated based on the proposed average of gradient orientations $\overline{\varphi}$ and squared gradient magnitudes $\overline{\rho}$, (6) and (7). This experiment is performed using all 606 photo/sketch pairs with different size of local patches ($s \times s$, $s = 8, 16, 32$ pixels). Sliding step δ and histogram bins *b* are set to $s/2$ and 9, respectively. The size of local patches s and sliding step δ are two parameters to divide face images into overlapped local patches while forming HAOG and HOG descriptors.

The normalized match scores using HOG and HAOG are separately divided into same-class match scores (*SCMS*, obtained by matching sketches and photos of same subjects) and different-class match scores (*DCMS*, obtained by matching sketches and photos belong to different subjects). The discriminative power of HOG and HAOG descriptors is measured by Fisher criterion as [22]:

$$J = \frac{(m_1 - m_2)^2}{v_1^2 + v_2^2} \tag{10}$$

where *m* represents a mean, v^2 represents a variance and the subscripts denote the two classes of matching scores, *SCMS* and *DCMS*. The Fisher criterion is calculated for matching scores obtained by both HOG and HAOG descriptors as J_{HOG} and J_{HAOG}, respectively. The greater Fisher criterion is, the discriminative the feature is.

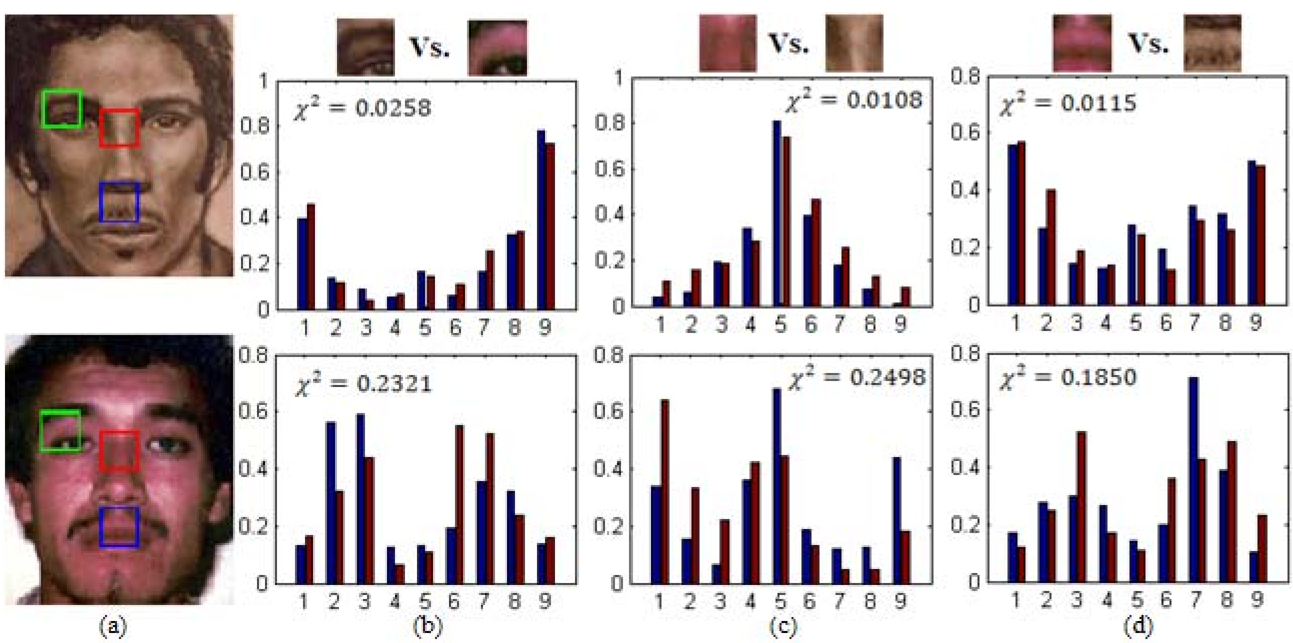

Figure 2. Face sketch (top), photo (bottom), (b,c,d) local patches (first row), HAOG descriptors (second row) and HOG descriptors (third row).

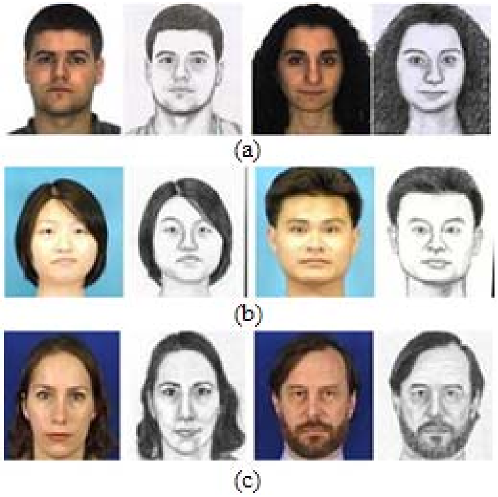

Figure 3. Examples of face photo/sketch pairs of CUFS database (a) AR database [20], (b) CUHK database and (c) XM2GTS database [21].

The results of the first experiment are illustrated in Fig. 4 and Table 1. Fig. 4 shows Fisher criterion versus patch size for HOG and HAOG descriptors. According to Fig. 4, the Fisher criterion of HAOG is significantly greater than HOG descriptor for all three sizes of patches. That states HAOG descriptor is more discriminative than HOG descriptor for face sketch recognition in presence of modality gap.

Furthermore, the recognition accuracy of HOG and HAOG descriptors is represented in Table 1. Obviously, HAOG descriptors achieved higher accuracy than HOG for all three sizes of local patches. That supports the fact that reducing modality gap in feature extraction stage improves the performance of face sketch matching.

In the second experiment, the performance of HAOG is compared with three state-of-the-art approaches which have been proposed for face sketch recognition. The parameters of these approaches are tuned according to their reference papers. For HAOG, patch size $s \times s$ and sliding step δ are chosen as 8×8 and 4, respectively.

LFDA [15]. A discriminate projection is learnt by integrating different LBP and dense SIFT descriptors in a single feature vector. LBP descriptors are computed with four radii $r = \{1,3,5,7\}$ and 8 neighboring points. SIFT descriptor is constructed by overlapping local patches which is parameterized by local patch size $s = 32$ and a displacement of $\delta = 16$. A 1-NN classifier on all 606 photo/sketch pairs is used for photo-sketch matching.

CITE [1]. CITE is characterized by five trees in a CITE forest and 256 nodes for each tree. 306 and 300 photo/sketch pairs are selected for training and testing, respectively. The accuracy obtained by averaging over five random splits.

MRF+RS-LDA [11]. The 606 photo/sketch pairs are randomly divided into three subsets: subset 1, subset 2 and subset 3 with 153, 153 and 300 photo/sketch pairs, respectively. First, subset 1 is used to train MRF-based photo synthesis [11]. Then, 153 face sketches belong to subset 2 are synthesized to 153 pseudo photos using the trained photo synthesis. RS-LDA [23] face recognition is trained by 153 pseudo photos and their corresponding face photos in subset 2. Finally, 300 photo/sketch pairs in subset 3 are used for face sketch recognition in testing stage.

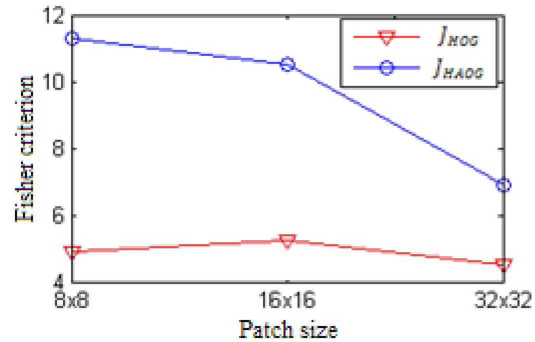

Figure 4. Fisher criterion versus patch size for HOG and HAOG descriptors.

TABLE I. RECOGNITION ACCURACY VERSUS PATCH SIZE OF HOG AND HAOG DESCRIPTORS.

	Patch size		
	8x8	***16x16***	***32x32***
HOG	92.16 %	89.03 %	80.65 %
HAOG	100 %	100 %	94.52 %

The results of the second experiment are shown in Table 2. According to Table 2, our method obtained the highest performance followed by CITE which was deliberately designed to reduce the modality gap in feature extraction stage. Although the problem of modality gap is not addressed by LFDA, its accuracy is promising due to fusing various features with different configurations. The approach based on MRF photo synthesis and RS-LDA face recognition achieved the lowest recognition rate, because its performance is highly relied on the result of photo synthesis.

V. CONCLUSION

We proposed a new face descriptor, called HAOG, for inter-modality face sketch recognition. The HAOG is inspired by the fact that the modality gap between face sketch and photo can be significantly reduced by emphasizing coarse texture of facial components in feature extraction stage. That means face photos and sketches can be directly matched in different modalities, with no necessity of photo/sketch synthesis.

The new method obtained 100% recognition accuracy on the CUFS database with slight degree of shape degradation/exaggeration in face sketches. However, recognizing face sketch with degraded/exaggerated shape is another major challenge which is not addressed in this paper.

Future work will focus to deal with face sketch recognition problem in presence of shape degradation, particularly in real-world applications when the eye-witness cannot exactly remember the detail of suspect's face. Moreover, a face sketch database drawn based on recollection of eye-witness is a necessity for future researches.

TABLE II. RECOGNITION PERFORMANCE. THE NEW METHOD VERSUS THREE MAJOR RECENT WORKS.

MRF+RS-LDA [11]	**LFDA [15]**	**CITE [1]**	**HAOG**
96.30%	99.47%	99.87%	100%

ACKNOWLEDGMENT

We acknowledge the generous support of NUS grant #R -252 -000-451-112 for this work.

REFERENCES

[1] W. Zhang, X. Wang and X. Tang, "Coupled information-theoretic encoding for face photo-sketch recognition", in Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 513–520, 2011.

[2] N. Dalal, B. Triggs, "Histogram of oriented gradients for human detection", in Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 2, pp. 886-893, 2005.

[3] Q. Zhu, M.C. Yeh, K.T. Cheng, S. Avidan, "Fast human detection using a cascade of histograms of oriented gradients", in Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1491–1498, 2006.

[4] M. Bertozzi, A. Broggi, M. Delrose, M. Felisa, A. Rakotomamonjy, F. Suard, "A pedestrian detector using histograms of oriented gradients and a support vector machine classifier", in Proc. of Intelligent Transportation Systems Conference, pp. 143–148, 2007.

[5] O. Deniz, G. Bueno, J. Salido, F. Torre, "Face recognition using histogram of oriented gradients", Pattern Recognition Letters, vol. 32, pp. 1598–1603, 2011.

[6] A. Albiol, D. Monzo, A. Martin, J. Sastre, A. Albiol, "Face recognition using HOG-EBGM", Pattern Recognition Letters, vol. 29, pp. 1537–1543, 2008.

[7] E. R. Davies, "The effect of noise on edge orientation computations", Pattern Recognition Letters, vol. 6, 1987.

[8] X. Tang and X. Wang, "Face sketch synthesis and recognition", in Proc. of European Conference on Computer Vision (ECCV), pp. 687–694, 2003.

[9] X.Tang and X. Wang, "Face sketch recognition", IEEE Trans. Circuits and Systems for Video Technology, vol. 14, no. 1, pp. 50–57, 2004.

[10] Q. Liu, X. Tang, H. Jin, H. Lu and S. Ma, "A nonlinear approach for face sketch synthesis and recognition", in Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1005–1010, 2005.

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

[12] J. Zhong, X. Gao and C. Tian, "Face sketch synthesis using e-hmm and selective ensemble", in Proc. of IEEE Conf. on acoustics, Speech and Signal Processing, 2007.

[13] B. Klare and A. Jain, "Sketch to photo matching: a feature-based approach", in Proc. of SPIE Conf. on Biometric Technology for Human Identification, 2010.

[14] D. Lowe, "Distinctive image features from scale-invariant keypoints", International Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, 2004.

[15] B. Klare, Z. Li and A. Jain, "Matching forensic sketches to mug shot photos", IEEE Trans. Pattern Analysis & Machine Intelligence, vol. 33, no. 3, pp. 639–646, 2011.

[16] Z. Farbman, R. Fattal, D. Lischinski, R. Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation", ACM Trans. of Graphics, vol. 27, no. 3, pp. 67:1–67:9, 2008.

[17] M. Kass and A. Witkin, "Analyzing oriented pattern", Computer. Vision, Graphics and Image Processing, vol. 37, pp. 362–385, 1987.

[18] A. M. Bazen and S. H. Gerez, "Systematic methods for the computation of the directional fields and singular points of fingerprint", IEEE Trans. Pattern Analysis & Machine Intelligence, vol. 24, pp. 905–919, 2005.

[19] L. Gibson, "Forensic Art Essentials", Elsevier, 2008.

[20] A. M. Martinez, and R. Benavente, "The AR face database," CVC Technical Report #24, June 1998.

[21] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre, "XM2VTSDB: The extended of M2VTS database," in Proc. of International Conference on Audio- and Video-Based Person Authentication, pp. 72–77, 1999.

[22] D.J. Hand, Discrimination and Classifcation, Wiley, New York, 1981.

[23] X. Wang and X. Tang, "Random sampling for subspace face recognition', International Journal of Computer Vision (IJCV), vol. 70, no. 1, pp. 91–104, 2006.