Face Sketch Synthesis using Bilateral Symmetry

Shinde Madhuri N.
M. E. Computer (Student)
Department of Computer Engineering
K. K. W. I. E. E. R., Nashik
S. P. P. U.

Shahane Nitin M.
Associate Professor
Department of Computer Engineering
K. K. W. I. E. E. R., Nashik
S. P. P. U.

ABSTRACT
Method of creating sketch from available photograph is known as sketch synthesis. To synthesize face sketch, similarity between different photograph patches is evaluated. Existing patch level method requires huge computational time and space for finding candidate photograph patches. Photograph patches are replaced by sparse coefficients using sparse coding. Sparse representation is used to select candidate photograph patches. Then greedy search is used to measure the similarity between the test photograph patch and the training photograph patch. Final candidate patches are stitched to get synthesized sketch. By using bilateral symmetry, it tends to achieve high efficiency of the system.

Keywords
Bilateral symmetry, Greedy search, Sketch synthesis, Sparse coding.

1. INTRODUCTION
Face sketch synthesis means creating sketches from available photographs. Applications of face sketch synthesis are law enforcement, digital entertainment that contains suspect searching, animation of movies or cartoons, missing child and face sketch ageing. During suspect searching, it might happen that photograph of suspect is not available. To solve this problem sketches are created by the artist with the help of description given by eyewitnesses. The face images can be categorized into photograph and sketch. The photograph is captured with optical imaging equipment or with different sensors [3]. And if good quality of the photograph is captured, then synthesized sketch will be of good quality. But if this condition is not satisfied means photograph is blur, background is so bright or so dark then the generated sketch will be of poor quality. The sketch is a subjective image drawn by artist [3]. Face sketch and photograph are of different modality. If matched directly, it leads to geometrical deformation and error occurs. So it cannot be matched directly [1]. Challenging part of the synthesis method is that to work well in uncontrolled conditions and the images with different background and size. If photograph is available with different lighting condition and blur then the generated sketch will be of poor quality. Image-based method generates strokes and shadings according to edges present in an input image. But it doesn't capture important facial details. Exemplar-based methods [9] reconstruct new sketches from available sketches. So it can produce sketches with high quality but it requires huge computational costs due to the matching process to a large amount of existing data. So to solve this problem, bilateral symmetry technique can be used in which partial left image matching is done first and then right side image matching is done which will reduce matching process time.

In the nest section, it includes the literature survey. Section 3 describes the block diagram and implementation of proposed system. Section 4 includes the results and discussion. Finally section 5 concludes with the summary of this paper.

2. LITERATURE SURVEY
Many existing methods are available for face photo sketch synthesis. Some of them are listed below:

Embedded Hidden Markov Model (E-HMM) [2, 3] is used to construct relationship between sketch and its corresponding photograph. Selective ensemble strategy is used to fuse pseudo sketches to get finer face pseudo-sketch. But due to noise, quality decreases of generated pseudo-sketch. Performance is improved but more time is required as this is an iterative algorithm and for optimal searching also.

With Multiscale Markov Random Fields model [4, 5] sketches can be well synthesized even with different hair styles. But still small deformation is occurred in the synthesized image. If the input photograph is different from above condition, then it will not work well.

Sparse representation [6] is used because it chooses relevantly best samples. And photograph will be represented very well. Consider a test face photograph with help of solving an l1- norm minimization problem with Lasso and gives its corresponding sketch image.

Initial estimation for the pseudo-image is obtained using Sparse Neighbor Selection (SNS) method [7]. And relation between sketch patches and photograph patches is mapped using Enhancement of Sparse Representation (SRE) method but little noise is added. Deformation occurs for mouth, beard, moustache, hair, chin area. Synthesized sketches are look like photos, lacks sketch styles.

Markov Weight Fields (MWF) [8] used to synthesize new sketch patches. As Markov Random Fields (MRF) cannot synthesis new sketch patches. It cannot find a suitable patch like patch for eyes, mouth.

Exemplar based method [9] gives high quality sketches. But matching process from large amount of data is a computationally difficult. Spatial Sketch Denoising (SSD) method eliminates the sketch noise and the facial details are well preserved. If the lighting conditions are different and search range is insufficient then patch matching will be incorrect and huge computational load will occur.

Sparse based representation greedy search [10] method synthesize face sketch with the help of combining image patches and prior knowledge. A photograph patch dictionary is created by randomly sampling patches from training photograph patches which reduces size of dataset. Searching nearest photograph patch for raw test photograph patch, it takes more time and high memory cost. So photograph patches are replaced by their sparse coefficients. By using prior knowledge i.e. patch intensity and gradient, it refines final candidate photograph patches. Final sketch is

synthesized by Bayesian inference. It can handle non-facial factors. It can work for different background, image size, face posture, multiple faces and photographs of old persons also. Due to sparse representation, incorrect patch matching occurs. But still it losses some facial features like textures around nose.

Bilateral symmetry [11] technique used in many applications like face recognition. In this system, face photograph is divided in vertical partition. First left part of face photograph patches is selected and they are matched with training photograph patches to get candidate photograph patches. Then right part of face photograph patches are consider to find candidate photograph patches. This process increases the efficiency of the system.

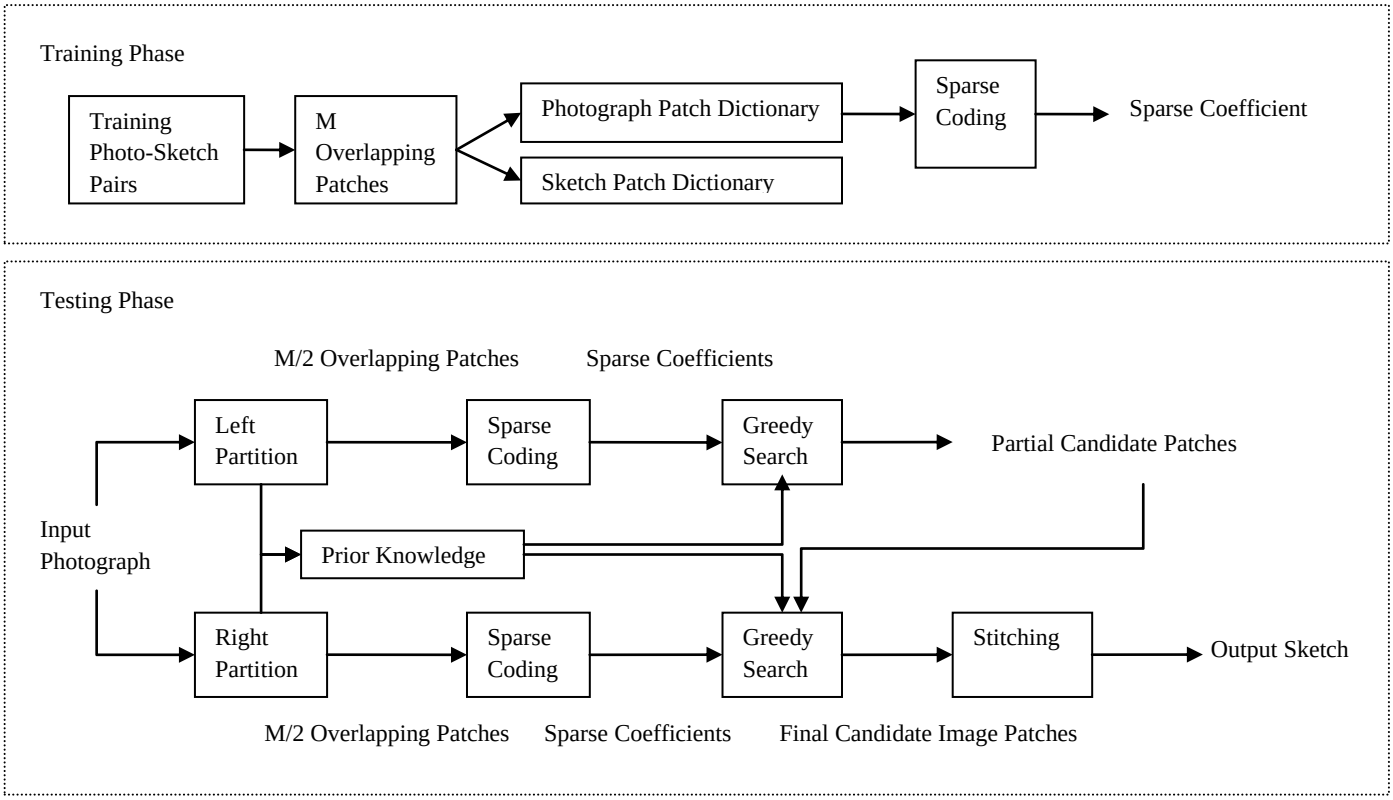

Fig 1: Block diagram of proposed system

3. IMPLEMENTATION DETAILS

This section contain details of proposed system i.e. block diagram and implementation of proposed system.

3.1 Block Diagram

The Block diagram of the system is shown in the Fig. 1. There are two main phases of the system: Training and Testing.

3.1.1 Training Phase

In this phase, training photo-sketch pairs are as input and gives sparse coefficients for all photograph patches.

3.1.1.1 Training Photograph-sketch Pairs

Photograph-sketch pairs are denoted by $(x_1, y_1),..,(x_N, y_N)$ where x_i, $i = 1,..,N$ denotes the i^{th} photograph and y_i, $i = 1,..,N$ denotes its corresponding sketch.

3.1.1.2 M Overlapping Patches

Training photograph-sketch pairs with 256 x 256 size are divided into M overlapping patches of size 64 x 64 with 50 % overlap. Training photograph-sketch patches are

$(\{x_1^1,.., x_1^M\},\{y_1^1,...,y_1^M\}),....(\{x_N^1,...,x_N^M\} ,\{y_N^1,..., y_N^M\})$.

3.1.1.3 Photograph Patch and Sketch Patch Dictionary

Photograph patch dictionary and sketch patch dictionary are used to store overlapping patches. So training photograph patches are $(\{x_1^1,..,x_1^M\},..\},...\{x_N^1,...,x_N^M\})$ and training sketch patches are $(\{y_1^1,...,y_1^M\},....,\{y_N^1,...,y_N^M\})$.

3.1.1.4 Sparse Coding

In sparse coding, patches are represented as weighted linear combinations of some patches from photograph patch dictionary. For sparse coding Orthogonal Matching Pursuit algorithm is used which calculates sparse coefficients values. $\{c_1^1,...,c_1^M,...c_N^1,...,c_N^M\}$ is a sparse coefficient set of the training photograph patches where c_i^j , i = 1,...,*N*; j = 1,...,*M*.

3.1.2 Testing Phase

In this phase, input photograph is used of which sketch is to be generated. Using left and right partition, input photograph is partition to left photograph patches and right photograph patches. And Prior intensity and Prior gradient (Prior knowledge) of patches are used during greedy search.

3.1.2.1 Greedy Search

For this method, sparse coefficient of input photograph patches and of dictionary photograph patches are used. For better selection of patches, prior intensity and gradient knowledge is used. For each input photograph patch, its corresponding sketch patch is simulated. It finds K candidate photograph patches i.e. nearest patches for each input photograph patch then corresponding sketch patches are considered. Patch matching is time consuming process hence

each patch is replaced by sparse coefficient and employs the sparse coefficients for search. As patches are represented by sparse coefficient, the memory require to store them is greatly reduced. Then get partial candidate photograph patches and sketch patches with sparse coefficient. These partial candidate photograph patches and sketch patches with sparse coefficient are fetched for again greedy search to get final candidate photograph patches and sketch patches.

3.1.2.2 Stitching

Greedy search gives final candidate sketch patches. M overlapped sketch patches are stitched together to make a sketch of size 256 x 256. And final sketch is displayed.

3.2 Implementation of proposed system

In the training phase, the system is trained using training dataset and sparse coefficients are generated as the output. In testing phase, a photograph will be taken as an input and the system produces its corresponding sketch. To get good quality of sketch with less process time, face photo-sketch synthesis system using bilateral symmetry approach is proposed.

1. Training Image pairs dataset contains photograph-sketch pair. They are divided into M overlapping patches.
2. Photograph patches are stored in Photograph Patch Dictionary. Sketch patches are stored in Sketch Patch Dictionary.
3. By sparse coding, a sparse coefficient of all photograph patches is calculated.
4. Consider an input photograph of which sketch to be generated.
5. Step 1 and 3 is repeated for input photograph.
6. Sparse coefficients of training image pairs and of left partitioned input photograph and prior knowledge are given to greedy search method as an input. It gives partial candidate image patches with sparse coefficients.
7. Sparse coefficients of partial candidate image patches and of right partitioned of input photograph are given to greedy search method as an input. It gives final candidate image patches.
8. Final candidate sketch patches are stitched to get final sketch.

4. RESULTS & DISCUSSION

4.1 Performance Measures

Structure SIMilarity [12] is used to evaluate visual perception quality by capturing loss of structure in the image. It is a combination of Luminance, Contrast, Structure factors. It can be defined as

$$\text{SSIM}(x,y) = \frac{(2\mu_x \mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (8)$$

Where

x and y are artist's sketch and synthesized sketch respectively.

μ_x , μ_y are average of intensity values of x and y respectively.

σ_x^2, σ_y^2 are the variance of intensity values x and y respectively.

σ_{xy} is the covariance intensity values of x and y.

c_1 and c_2 are two variables to stabilize the division with weak denominator.

4.2 Experimental Setup

This section describes about the Chinese University of Hong kong (CUHK) dataset used by the system for experimentation purpose. CUHK dataset shown in fig. 2.

Fig 2: Chinese University of Hong Kong (CUHK) dataset

The dataset consists of 188 images of photograph-sketch pairs. It includes 54 females and 134 males. This dataset includes face photographs and sketches with different conditions like it can have hairpins, hair style, different background, and different image size. Original photograph possess resolution 1024 x 768 with 144kb size. And original sketch possess resolution 414 x 582 with 38kb size. Photograph and sketches are cropped to 256 x 256. Images are divided to get overlapping Patches of size 64 with $1/2$ overlapping region.

4.3 Result Table

CHUK database is used for this experiment. In this system 88 image pairs are trained which are divided into patch size 32 x 32 and 64 x 64 and sparse coefficients are calculated. For testing 100 images are available. Sparse coefficients of training patches and input photograph patches are matched. And gives final candidate image patches. Later final sketch is stitched from final candidate sketch patches. After this synthesized sketch and available sketch are used to check the similarity by using SSIM parameter.

Table 1. SSIM values on CHUK database

Sr. No.	Test Images	No. of images	Average SSIM	Average SSIM
			32 x 32	64 x 64
01	Female face (Training)	34	0.9827	0.9553
02	Male face (Training)	54	0.9780	0.9678
03	Female face (Testing)	20	0.8479	0.9524
04	Male face (Testing)	80	0.8776	0.9595
		188	0.9215	0.9587

This tested on two different patch sizes 32 x 32 and 64 x 64. From results it was decided experimentally to use patch size 64 x 64 because it is found that patches are incorrectly matched leading to poor quality in case of 32 x 32 patch size.

5. CONCLUSION

Face photo-sketch synthesis method is used in various areas like law enforcement, digital entertainment and face sketch aging. This method works very well in well controlled conditions and with non-facial factors. But the problem occurs when the input is taken with different lighting conditions and poses variations. Some facial features like texture around nose tend to lose because of matching incorrect patches. Proposed system gives better quality of sketch with less time. For time efficiency it is analyzed that by using bilateral symmetry system takes 10 to 15 sec. to generate a sketch. But still there is a scope to improve on this parameter. Further, proposed system is required to carry out experimentation and investigations to check any gender bias or race bias.

6. ACKNOWLEDGMENTS

I would like to express my gratitude to my guide Prof. N. M. Shahane, Associate Professor, Computer Engineering, K.K.W.I.E.E.R., Nashik for giving me a moral support, valuable guidance and encouragement. A special thanks to Prof. Dr. K. N. Nandurkar, Principal and Prof. Dr. S. S. Sane, Head of Department of K.K.W.I.E.E.R. Nashik for their kind support and suggestions. It would not have been possible without the kind support. I would like to extend my sincere thanks to all the faculty members of the department of computer for their help. I am also thankful to my colleagues who encouraged and willingly helped me with their abilities. Lastly, I am also thankful to those who helped me directly or indirectly.

7. REFERENCES

[1] Q. Liu, X. Tang, H. Jin, H. Lu, and S. Ma, "A nonlinear approach for face sketch synthesis and recognition", in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 1005-1010, Jun. 2005.

[2] J. Zhong, X. Gao, and C. Tian, "Face sketch synthesis using E-HMM and selective ensemble", in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., pp. 485-488, Apr. 2007.

[3] X. Gao, J. Zhong, J. Li, and C. Tian, "Face sketch synthesis algorithm based on E-HMM and selective ensemble", IEEE Trans. Circuits Syst. Video Technol., vol. 18, no. 4, pp. 487-496, Apr. 2008.

[4] 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.

[5] W. Zhang, X. Wang, and X. Tang, "Lighting and pose robust face sketch synthesis", in Proc. 11th Eur. Conf. Comput. Vis., pp. 420-433, Sep. 2010.

[6] L. Chang, M. Zhou, Y. Han, and X. Deng, "Face sketch synthesis via sparse representation", in Proc. 20th Int. Conf. Pattern Recognit., pp. 2146-2149, Aug. 2010.

[7] X. Gao, N. Wang, D. Tao, and X. Li, "Face sketch photo synthesis and retrieval using sparse representation", IEEE Trans. Circuits Syst. Video Technol., vol. 22, no. 8, pp. 1213-1226, Aug. 2012.

[8] H. Zhou, Z. Kuang, and K.-Y. K. Wong, "Markov weight fields for face sketch synthesis", in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 1091-1097, Jun. 2012.

[9] Y. Song, L. Bao, Q. Yang, and M.-H. Yang, "Real-time exemplar based face sketch synthesis", in Proc. 13th Eur. Conf. Comput. Vis., pp. 800-813, Sep. 2014.

[10] S. Zhang, X. Gao, N. Wang, J Li, and M Zhang, "Face Sketch Synthesis via Sparse Representation-Based Greedy Search", IEEE Trans. Image Process., Vol. 24, No. 8, pp.2466-2477, August 2015.

[11] A. Singh, G. C. Nandi, "Face Recognition Using Facial Symmetry", ACM , pp. 550-554, 2012.

[12] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Trans. Image Process., vol. 13, no. 4, pp. 600–612, Apr. 2004.

[13] S. Ouyang, T. Hospedales, Y. Song, X. LI, "A Survey on Heterogeneous Face Recognition: Sketch, Infrared, 3D and Low-resolution", ACM Computing Surveys, pp. A1-A35.

[14] Shinde Madhuri N. and Prof. Shahane Nitin M.," A Review on Face Photo-Sketch Synthesis", IJSRD - International Journal for Scientific Research & Development| Vol. 3, Issue 11, 2016 | ISSN (online): 2321-0613.

[15] Shinde Madhuri N. and Prof. Shahane Nitin M., "Face Photo-Sketch Synthesis",cPGCON-2016.