A Review on Face Photo-Sketch Synthesis
Madhuri Shinde1 N.M.Shahane2
1,M.E. Student 2Associate Professor
1,2Department of Computer Engineering

K. K. W. I. E. R., Nashik Savitribai Phule Pune University, Maharashtra, India.

Abstract— Face sketch synthesis is a process of creating a sketch from available face photograph. This process is widely used in law enforcement and digital entertainment. If some factors except the facial are not present in training set, but they are present in photograph like hairpin, hair style, glasses, then it becomes more difficult to handle. And existing methods fails if the photographs are of different background and size. This results to poor quality of sketches. For that purpose different approaches are proposed. And this paper contains a review of all existing approaches were used for face sketch synthesis.

Key words: Face photograph, Face sketch, Face sketch synthesis

I. INTRODUCTION

Applications of face sketch synthesis are law enforcement, digital entertainment that contains suspect searching, animation of movies or cartoons, missing child and security purpose etc. Challenging part of the synthesis method is that to work well in uncontrolled conditions and the images with different background and size. 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]. Viewed sketch, forensic sketch, composite sketch and caricature sketch these are different types of sketches [12]. Viewed sketch means artist have already seen the face before sketching. Forensic sketch means artist draws the sketch as per eyewitness describes the details. Composite sketch is generated by using specific software. And the caricature sketch means some facial features are exaggerated. Face sketch and photograph are of different modality. So it cannot be matched directly. For that purpose, three different cross-model approaches:

1) To synthesize one modality from another and then they can be compare directly;

2) To learn feature representations which are variant to person identity and invariant to image modality than raw pixels; and

3) Projection of both views into a common space in order to compare directly.

Face image can have different representations which includes analytic, component-based, global holistic, patch-based. Analytic representation detects fiducial points which are invariant to modality. But does not work well with facial expression, texture information is not detected. Component-based representation detects face part like eyes, nose and mouth. Global holistic representation takes whole face image is considered with a single vector. But it is sensitive to pose variation, facial expression and alignment. Last one is patch-based representation; in this divided patches are considered.

II. RELATED WORK

Q. Liu et. al [1] proposed Locally Linear Embedding (LLE) method. It does pseudo-sketch synthesis. It gives mapping relation between sketches and photographs. With this LLE method, nonlinear dimension is reduced from high-dimensional data. If the number of neighbors K is too large then blur appears. If patch size is too large then some facial details will be disappeared and if patch size is too small then noise occurs. So the kernel based nonlinear discriminant analysis (KNDA) is used. It will be used for sketch recognition from pseudo-sketches. But at the fixed scale, the local patches are independently synthesized and the face structures in large scale, so it becomes difficult.

X. Gao and J Li [2, 3] introduced algorithm which synthesize sketch automatically. Nonlinear relationship is constructed between sketch and its corresponding photo with the help of embedded hidden Markov model (EHMM). Then pseudo-sketches will be generated of given photo. It has moderate computational complexity. Also it has ability to extract 2-D facial features except for hair region. But due to noise, quality decreases of generated pseudo-sketch. So the E-HMM inversion (HMMI) is presented to improve the quality. With the help of Selective ensemble strategy pseudo sketches are fused to get finer face pseudo-sketch. Generalization ability is improved. To synthesized final pseudo-sketch, n number of pseudo-sketches and their weights are fused. With this E-HMMI approach, performance is improved but more time is required as this is an iterative algorithm and for optimal searching also.

X Wang and X Tang [4] proposed the scale of the face structure depends on the size of patches. If the patch is small in size and if shadow is added then some facial details (structures) will be missing like eyebrow, bridge of nose. At different scale, facial structures are acquired via Multiscale Markov Random Fields model. It is to be assumed that faces are taken in normal lighting, with front pose, with no expression and occlusion. Initially all photographs and sketches should be translated, rotated and scaled. Face photograph can be of color or gray images. If photograph is in RGB color then it will converted into Luv color space as it better correlates to change in color. With this method, sketches can be well synthesized even with different hair styles. But still small deformation is occurred in the synthesized image. If the input photo is different from above condition, then it will not work well.

W. Zhang et. al [5] proposed a algorithm which can synthesize a sketch of a photograph taken in different lighting condition and pose variation with the help of a multiscale MRF model. If lighting and pose variants then distortion and artifacts occurs. To reduce it shape priors are included. Candidate of sketch patches are searched by...
different patch descriptor. To match neighboring sketch patches, both intensity and gradient compatibility are used which will be smoothen patches. Edges are detected and enhanced by filter called Difference-of-Gaussians (DoG) and get synthesized sketch with better facial details specifically on the nose and the eyebrows. Effect of variation lightening is relatively less.

L. Chang et. al [6] proposed sparse representations which 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 11 - norm minimization problem with Lasoo; it gives its corresponding sketch image. A coupled dictionary is generated by sparse coding with pairs of photograph and sketch patch. Sparse coefficient of image patches of test photograph is calculated as per photograph element present in coupled dictionary. Coefficient with the sketch elements from the coupled dictionary are taken to get sketch patch. And finally face sketch is created from the final obtained sketch patches.

X. Gao [7] proposed two steps sparse neighbor selection (SNS) and enhancement of sparse representation (SRE). With SNS, initial estimation for the pseudo-image is obtained. And with the SRE, relation between sketch patches and photograph patches is mapped. And synthesized photographs and sketches are obtained with enhanced quality. But with this SRE, little noise is introduced. Deformation occurs for mouth, beard, moustache, hair, chin area. Synthesized sketches are look like photos, lacks sketch styles.

H. Zhou et. al [8] proposed Markov Random Fields (MRF) cannot synthesis new sketch patches. It might happen that it cannot find a suitable patch like patch for eyes, mouth. In case of solving this, it is NP-hard problem. New sketch patches are synthesized by using Markov Weight Fields (MWF) and it also preserves large scale features. In order to have guaranteed optimal solution, convex quadratic programming (QP) problem need to be devised. Cascade decomposition method is used to solve QP problem efficiently.

Y. Song [9] proposed exemplar based method gives high quality sketches. But due to matching process from large amount of data, it is computationally difficult. If linear combination of patch is placed densely, then synthesized sketch will contain some amount of noise. So the Spatial Sketch Denoising (SSD) method is used. It 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. When patches are matched correctly with different poses and with increased search range, then sketches are better generated with huge computational load. If SSD algorithm is integrated with PatchMatch (PM) method, it will work well without increasing computational load.

S. Zhang et. Al [10] proposed and face sketches are synthesized with the help of combining image patches and prior knowledge. A photograph patch dictionary is learned from training photograph patches which reduces the dimension of the image patch. Searching nearest photo patch for raw test photo patch, it takes more time and high memory cost. So photo patches are replaced by their sparse coefficients and they find their nearest neighbors. By using prior knowledge i.e. patch intensity and patch gradient, it refines final candidate photo 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.

A. Singh and G. Nandi [11] proposed a technique of face recognition using bilateral symmetry. They propose half face recognition technique as user face has bilateral symmetry. For face recognition, Principle Component Analysis technique is used. This half faces recognition technique increases the efficiency of the system.

III. ANALYSIS

Existing methods are not able to handle the face images captured with uncontrolled conditions, with different background, and size and they are not working well if the non-facial factors like hair-pin, hair style and glasses are not included in training set. This hamper the quality of sketches generated using their methods. Also in that approaches there are assumptions about the picture capturing process e.g. pictures are taken in standard lighting condition hence the same kind of images were not recognized because of variation in lighting condition. In some systems effective sketch synthesis is done though light variation is there in the original picture, but in that case it is a time consuming process.

IV. CONCLUSION

As per the overview of some of the existing systems it can be analyzed that with the help of sparse representation and greedy search, the face sketch synthesis method can be used to solve the problems mentioned in analysis. It works well in controlled conditions and can handle non-facial factors with different image backgrounds, image size, alignments which are not present in the training set. But still it loses some facial features like texture around nose because it matches incorrect patches. This happens because of having different lightning conditions and pose variation. Greedy method is efficient and it has generalization ability to synthesize the final face sketch. But this approach requires training the whole dataset to collect the similarity between different patches of candidate photo and hence it is very time consuming task. This problem can be overcome by using bilateral symmetry which requires only half face and hence it tends to achieve high efficiency of the system with less time.

REFERENCES


