International Journal of Engineering Sciences &Research Technology

(A Peer Reviewed Online Journal)

Impact Factor: 5.164

IJESRT

RESEARCHERID
THOMSON REUTERS

[Aict 2019]
ICTM Value: 3.00

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

A SURVEY ON FACIAL IMAGE RECOGNITION AND RECONSTRUCTION

Reshma M S*[1] & Sabeena M [2]
PG Scholar, Assistant Professor
* Dept. of Electronics and Communication Engineering, College of Engineering Karunagappally, Kollam, India

DOI: 10.5281/zenodo.2629249

ABSTRACT

Human face is essential information for robust applications. Facial Image Recognition and Reconstruction is always challenging in computer vision. The facial images are often present in very low resolution due to environmental influence and imaging equipment limitation. This can be solved by using many face super resolution techniques. Facial image recognition methods can be vision oriented or recognition oriented methods. Many works have been done in the progress of automatic Facial image Recognition and Reconstruction. This has many applications in security, health and human computer interaction. In the recent years, there is an increase in the studies of Facial Image Recognition and Reconstruction system. This survey presents a review of approaches in Facial Image Recognition and Reconstruction using different technologies.

KEYWORDS: Facial Image Recognition, Reconstruction, Face Super Resolution, Low Resolution, Face Super Resolution

1. INTRODUCTION

Facial image recognition and reconstruction is currently a very active area of research. Facial images provide important information for human visual perception as well as computer analysis. Facial image analysis has been widely studied for many years, with specific research problems including face detection, recognition, expression analysis and animation. However, the performance of most existing systems is affected by the resolution of facial images, like low-resolution (LR) facial images are captured by surveillance cameras. When the distance between the human and the camera is large, this limits the performance of face recognition systems. In order to solve LR face recognition problems, many face super resolution techniques have been used. The traditional methods usually employ super resolution (SR) as a preprocessing step to get the high resolution (HR) image and then pass the super resolved facial image to some face recognition and reconstruction system.

One way to increase the resolution is interpolation. However, the performance of the interpolation algorithm is poor. Many more effective SR algorithms have been proposed during the past researches. Face super resolution (FSR) methods can be divided into vision oriented and recognition-oriented methods. Vision-oriented methods focus on obtaining good visual effects by image reconstruction while recognition-oriented techniques aim at achieving high recognition accuracy on the LR face images. Vision-oriented typical methods include manifold-based [1, 2, 3], dictionary-based [4, 5] and regression based methods [6, 7, 8, 9, 10].The recognition-oriented techniques are sparse-representation-based methods [11, 12] and deep-learning based [13, 14, 15, 16] methods. The challenging task of estimating a high resolution (HR) image from its low-resolution (LR) is referred to as super-resolution (SR). SR received a great attention from the computer vision research community and has a wide range of applications.

The first of facial detection system has been developed since in early 1970's. But due to the limitation of the computation, system can't be satisfied the requirement of users, it is to identify passport photograph real time. At the beginning of 1990's M. A. Kerin [17] was the first to propose face recognition using a digital neural network with self organizing capacity. With the improvement in deep learning methods, recent studies have concentrated on the application of deep Neural Networks to specifically perform Facial Image Recognition and

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Reconstruction. Instead of giving an extensive overview on all previous methodologies on Facial Image Recognition and Reconstruction, this survey concentrates on some typical methods mentioned above.

2. FACIAL IMAGE RECOGNITION AND RECONSTRUCTION APPROACHES

Based on the methodologies used for handling Facial image Recognition and Reconstruction approaches can be classified into vision oriented and recognition-oriented approaches. Which is again classified as manifold-based approach, dictionary-based approach, regression based approach, sparse-representation-based approach and deep-learning based approach.

A. MANIFOLD BASED APPROACH

The impact of super-resolution problems in facial images are overcome by using a set of training examples. Hong Chang [1] proposed a novel method for solving single-image super-resolution problems by training the examples. Proposed a general way of using the training examples, so multiple training examples can contribute the high-resolution image. The results show that the method is very flexible and gives good empirical results.

Xiaoqiang Lu [2] proposed a method; double sparsity regularized manifold learning (DSRML). The novel sparse coding technique is formulated by extending the locally linear embedding (LLE) method. DSRML can preserve the properties of the local geometrical structure by employing the manifold learning. The number of experimental results shows robustness and effectiveness of the proposed SR method.

Changbo Hu [3] proposed a novel approach for modeling, tracking and recognizing facial expressions. This method works on low dimensional expression manifold, it is obtained by Isomap embedding. Active Shape Model (ASM) algorithm is used for image observation. ICondensation tracking is used for the tracking and classification. The results show it is a robust method for facial expression tracking based on the ASM models.

B. DICTIONARY BASED APPROACH

Dictionary learning methods are mainly focus on training an over-complete dictionary in a single feature space for various recognition tasks. In many applications there are two feature spaces: high and low-resolution signal (feature) spaces. Denoting the two spaces as the observation space and the latent space and are tied by some mapping function.

Jianchao Yang [4] proposed a coupled dictionary training method for single-image super-resolution (SR).It is based on the patchwise sparse recovery. There are two coupled dictionaries: the observation dictionary and the latent dictionary. There are many potential applications in both signal processing and computer vision by learning the coupled dictionaries such as compressive sensing [18].The experimental results shows by comparing with the previous joint dictionary training method pixelwise mean square error (MSE) improvement using the coupled dictionary training method. The algorithm improves the recovery accuracy and at the same time removes the recovery artifacts.

Quan Pan [5] proposed a semi-coupled dictionary learning (SCDL) model to solve cross-style image synthesis problems. In SCDL, a pair of dictionaries and a mapping function will be simultaneously learned. The two dictionaries in SCDL will not be fully coupled so that much flexibility can be given to the mapping function.

C. REGRESSION BASED APPROACH

A simple but efficient linear regression based classification (LRC) approach is developed by Naseem, Togneri and Bennamoun [19]. LRC based approach represents a probe image by the linear combination of all the images of a object in the gallery set. The similarity between the probe image and the group of images of the object in the gallery set is represented by the residual error.

Christian Ledig [7] proposed a SRGAN, a generative adversarial network (GAN) for image super-resolution (SR). This is the first framework capable of inferring photo-realistic natural images for the upscaling factors of 4x.For achieving this a perceptual loss function is introduced. Perceptual loss consists of an adversarial loss and a content loss. This paper describes the first very deep ResNet [20] architecture by using the concept of GANs to form a perceptual loss function. A significant gain in perceptual quality is shown in the experimental results.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Yibing Song [8] proposed a two-stage method for face hallucination. The facial component of the input image is generated by using CNNs. Then synthesizing fine-grained facial structures from high resolution training images. For enhancement the details of these structures are transferred. Present a pipeline of Learn to hallucinate face images via Component Generation and Enhancement (LCGE) algorithm. The experiments demonstrate that this method performs favorably against state-of-the-art methods.

Yuan Wang [10] proposed an image-to-image face recognition algorithm that uses Dual Linear Regression based Classification (DLRC). Each face image is first converted into a cluster of images. By shifting the original image a few pixels, each image in the cluster is obtained. The distance between the corresponding image clusters represents the similarity of a pair of face images. Sub images are obtained by partitioning the clusters which are the single image into union of clusters for improving performance. Similarities between corresponding sub-image clusters is measured by DLRC. The experimental result demonstrates that the proposed approach works best in certain simple situations.

D. SPARSE REPRESENTATION BASED APPROACH

In sparse coding approach the training samples are treated as a dictionary. Robust object recognition is done using a sparse representation of a test image. A linear combination of training images is used for its formation. For dynamic range compression in images [21], separation of texture and cartoon content in images [22], inpainting [23], sparsity and overcompleteness can be successfully used. The hard problem that has been extensively investigated in the past few years is the extraction of the sparsest representation.

Sparse coding is the process of computing the representation coefficients based on the dictionary and the given signal. This process is referred as atom decomposition and is typically done by the pursuit algorithm. The simplest algorithms are matching pursuit (MP) and orthogonal matching pursuit (OMP) algorithms. Another well-known pursuit approach is the basis pursuit (BP). The focal underdetermined system solver (FOCUSS) is very similar and the combination of both the BP and the FOCUSS can be motivated based on maximum a posteriori (MAP) estimation.

Michal Aharon [12] presents a new method, the K-SVD algorithm, K-means clustering process. It is an iterative method which alternates between sparse coding of the examples based on the current dictionary and a process. The process is updating the dictionary atoms to better fit the data. This K-SVD algorithm is flexible and can work with any pursuit method.

Jianchao Yang [11] presents sparse representation for each patch of the low-resolution input and using the coefficients of this representation, high-resolution output is generated. This paper focuses on recovering the SR version of a given low-resolution image. Results demonstrate the effectiveness of the sparsity as a prior for patch-based SR for generic and face images.

E. DEEP LEARNING BASED APPROACH

Deep learning is a machine learning approach and these are composed of multiple processing layers. In 1980's deep learning was first theorized. Large amount of labeled data and substantial computing powers are required for deep learning. The neural networks (NNs) has the power that it lies in their ability to approximate any continuous function. A n-layer NN having an input layer and $n - 1$ hidden layers. Deep neural network (DNN) has been widely used in computer vision. This is a powerful tool for uncovering the nonlinear information hidden in the data. The commonly used activation functions are, sigmoid and the Rectified Linear Unit (ReLU). In 1988 the birth of Convolutional neural networks (CNN) or ConvNets has been traced [24].

Zhangyang Wang [13] attempt to solve the Very Low Resolution Recognition (VLRR) problem using deep learning methods. In this paper the advantage of super resolution, domain adaptation and robust regression is taken. There are five models are used in this paper for the comparison, starting from the convolutional neural networks (CNNs) baseline. The effectiveness of the proposed models is evaluated by three different VLRR tasks, which includes face identification, digit recognition and font recognition.

Jin Yamanaka [14] proposed a faster Single Image Super-Resolution (SISR) model with Deep Convolutional neural networks (Deep CNN). Deep CNN have a significant reconstruction performance on single-image super-

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

resolution. This model achieves state-of-the-art reconstruction performance with at least 10 times lower calculation. The feature extractor is the combination of Deep CNNs and Skip connection layers.
Yasen Aizezi [15] proposed an iterative deep learning algorithm (IIDLA) based on convolutional neural network. Pooled convolutional layer (PCL) is adopted by the algorithm and the different nonlinear features of input image set are learned by the convolutional neural network (CNN). The algorithm is proposed for face recognition and object classification of image set task. IIDLA is the theoretical framework of CNN and improves it to certain degree. Coding layer, decoding layer and hidden layer are the main layers of IIDLA. Three types of datasets are collected and the result shows best performance on the evaluation dataset.

Lucy Nwosu[16] proposed a two-channel convolutional neural network in which the Facial Parts (FPs) are used as input to the first convolutional layer and the extracted eyes are used as input to the first channel. The input into the second channel is the mouth. Two separate CNN channels are used in this work, the output of the two channels are converges into a fully connected layer and the result used for classification.

3. ANALYSIS FACIAL IMAGE RECOGNITION AND RECONSTRUCTION

- Manifold based approaches having algorithms that are more powerful than other pattern recognition methods. But there are some practical problems it prevents the algorithms being applied to super-resolution. Almost of the manifold learning methods cannot generate mapping functions for new test images these are absent from a training set.
- Dictionary based approaches having high accuracy and at the same time removes the recovery artifacts. These methods are very flexible and the disadvantage is that most of the systems are not unsupervised; these are not discriminative enough to classify face sets.
- Regression based approaches are simple but effective. Major limitation of the current regression based methods is the shape information is ignored or ineffectively used.
- Sparse representation based approaches are robust but there are still some limitations, they are not quite robust to poses, misalignment variations and expressions, especially in under-sampled cases.
- Deep learning based approaches having high computation speed but have a main disadvantage is that greater number of training samples is required.

Table 1 provides an overview of the Facial image Recognition and Reconstruction approaches with respect to author and classifier.

Table1. An analysis of facial image recognition and reconstruction

Author/Year	Classifier
Hong Chang.2007[1]	Locally linear embedding (LLE)
Xiaoqiang Lu.2013[2]	Double sparsity regularized manifold learning (DSRML)
Jianchao Yang.2012[4]	Patchwise sparse recovery
Quan Pan.2012[5]	Semi-coupled dictionary learning (SCDL)
Christian Ledig.2017[7]	SRGAN
Yibing Song.2017[8]	Learn to hallucinate face images via Component Generation and Enhancement (LCGE)
Yuan Wang.2016[10]	Dual Linear Regression based Classifier (DLRC)
Jianchao Yang.2010[11]	Sparse Representation Classifier(SRC)
Michal Aharon.2006[12]	K-clustering process(K-SVD)
Zhangyang Wang.2016[13]	Deep CNN, SR Pre-training, Partially Coupled Networks

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Jin Yamanaka.2017[14]	Deep CNN
Yasen Aizezi.2018[15]	Convolutional neural network (CNN)
Lucy Nwosu.2017[16]	Two-channel convolutional neural network(CNN)

4. CONCLUSION

This paper reviewed various Facial Image Recognition and Reconstruction approaches. Eventhough, a lot of studies were made in the domain of handling facial images not many researchers have been introduced to completely overcome the problem of Facial Image Recognition. Many authors have worked on Facial Image Recognition and Reconstruction systems under controlled environments and certain specified datasets but with uncontrolled conditions less work has been focused. Most researchers have considered certain type of datasets, but number of studies considering real-world datasets and number of features are still very limited. The sparse coding based approach and dictionary based approach are mainly used for Facial Expression recognition from images. Regression based approach and Deep Neural Networks based approaches can be used for both cases. Sparse coding based approaches and Deep Neural Networks based approaches give more effective results for Facial Image Recognition and Reconstruction.

REFERENCES

[1] H. Chang, D. Y. Yeung, and Y. Xiong, “Super-resolution through neighbor embedding,” in IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2004, pp. 275–282.
[2] X. Lu, Y. Yuan, and P. Yan, “Image super-resolution via double sparsity regularized manifold learning,” IEEE Transactions on Circuits & Systems for Video Technology, vol. 23, no. 12, pp. 2022–2033, 2013.
[3] Changbo Hu, Ya Chang, Rogerio Feris, and Matthew Turk, “Manifold Based Analysis of Facial Expression”, in IEEE Conference on Computer Vision and Pattern Recognition Workshop, 2004
[4] J. Yang, Z. Wang, Z. Lin, S. D. Cohen, and T. S. Huang, “Coupled dictionary training for image super resolution,” IEEE Transactions on Image Processing, vol. 21, no. 8, pp. 3467–78, 2012.
[5] Q. Pan, Y. Liang, L. Zhang, and S. Wang, “Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis,” in IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2012, pp. 2216–2223.
[6] C. Dong, C. L. Chen, K. He, and X. Tang, “Image super-resolution using deep convolutional networks,” IEEE Transactions on Pattern Analysis & Machine Intelligence, vol. 38, no. 2, pp. 295, 2016.
[7] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, and et al., “Photo-realistic single image super-resolution using a generative adversarial network,” in IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016, pp. 4681–4690.
[8] Y. Song, J. Zhang, S. He, L. Bao, and Q. Yang, “Learning to hallucinate face images via component generation and enhancement,” in International Joint Conference on Artificial Intelligence. Morgan Kaufmann, 2017, pp. 4537–4543.
[9] E. Zhou, H. Fan, Z. Cao, Y. Jiang, and Q. Yin, “Learning face hallucination in the wild,” in AAAI Conference on Artificial Intelligence. AAAI Press, 2015, pp. 3871– 3877.
[10] Yuan Wang, Liang Chen, “Image-to-Image Face Recognition Using Dual Linear Regression Based Classification and Electoral College Voting,” IEEE,2016
[11] J. Yang, J. Wright, T. S. Huang, and Y. Ma, “Image super-resolution via sparse representation,” IEEE Transactions on Image Processing, vol. 19, no. 11, pp. 2861, 2010.
[12] Michal Aharon, Michael Elad, and Alfred Bruckstein, “K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation,” IEEE,2006
[13] Z. Wang, S. Chang, Y. Yang, D. Liu, and T. S. Huang, “Studying very low resolution recognition using deep networks,” in IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016, pp. 4792–4800.
[14] Jin Yamanaka, Shigesumi Kuwashima and Takio Kurita, “Fast and Accurate Image Super Resolution by Deep CNN with Skip Connection and Network in Network,” 102-0084, Japan

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

[15] Yasen Aizezi, Anniwaer Jiamali, Ruxianguli Abudurexiti, Xuehua Liu, Jin Du, Liping Ding, "Research on Image Recognition Method Based on Deep Learning Algorithm," IEEE, 2018

[16] Lucy Nwosu, Hui Wang, Jiang Lu, Ishaq Unwala, Xiaokun Yang, Ting Zhang, "Deep Convolutional Neural Network for Facial Expression Recognition using Facial Parts," 2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress

[17] M. A. Kerin, T. J. Stonham, "Face recognition using a digital neural network with self- organising capabilities," 1990 IEEE

[18] E. J. Candès, "Compressive sampling," in Proc. Int. Congr. Math., 2006, p. 1.

[19] I. Naseem, R. Togneri and M. Bennamoun. Linear Regression for Face Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(11): 2106–2112, 2010.

[20] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 770–778, 2016.

[21] R. Gastaud and J. L. Starck, "Dynamic range compression : A new method based on wavelet transform," in Astron. Data Anal. Software Systems Conf., Strasbourg, 2003.

[22] J. L. Starck, M. Elad, and D. L. Donoho, "Image decomposition: Separation of texture from piece-wise smooth content," in SPIE Conf. Signal Image Process.: Wavelet Applicat. Signal Image Process. X, SPIE 48th Annu. Meeting, San Diego, CA, Aug. 3–8, 2003.

[23] M. Elad, J. L. Starck, P. Querre, and D. L. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)," J. Appl. Comput. Harmon. Anal., submitted for publication.

[24] Y. L. Cun, B. Boser, J. S. Denker, R. E. Howard, W. Habbard, L. D. Jackel, and D. Henderson, "Advances in neural information processing systems 2," D. S. Touretzky, Ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1990, ch. Handwritten Digit Recognition with a Back-propagation Network, pp. 396–404.