

A study on facial expressions recognition

Jingjing Xu *

*School of Mechanical, Electrical & Information Engineering, Shandong University (Weihai), Weihai 264209, China

17862705803@163.com

Abstract. In terms of communication, postures and facial expressions of such feelings like happiness, anger and sadness play important roles in conveying information. With the development of the technology, recently a number of algorithms dealing with face alignment, face landmark detection, classification, facial landmark localization and pose estimation have been put forward. However, there are a lot of challenges and problems need to be fixed. In this paper, a few technologies have been concluded and analyzed, and they all relate to handling facial expressions recognition and poses like pose-indexed based multi-view method for face alignment, robust facial landmark detection under significant head pose and occlusion, partitioning the input domain for classification, robust statistics face formalization.

1 Introduction

As an important part of artificial intelligence, face recognition has gained much popularity and is able to be applied in many purposes. Many scientists devote themselves in developing methods to refine the performance of face recognition system to an extent like human eyes.

However, there are some drawbacks in these algorithms. Firstly, the method Current cascaded pose regression (CPR) [1-6] used for face alignment is not ideal. When the mean shape is far away from the ground truth, it is not reliable to use the mean shape as an initiation. What's more, when dealing with large pose variations, the conditions referred above are more likely to occur. Secondly, when dealing with facial landmark detection, most of the algorithms cannot handle images with occlusions and large head poses simultaneously [5], [7-8]. Thirdly, with the classification tasks becoming more and more sophisticated, more complex classifiers like Convolutional neural networks [9, 10] are in demand. At last, state-of-the-art landmark localization methods take advantages of available annotated data and method [11-14] will consume a lot of time and money. In addition, 2D landmark localization is not desirable enough for face formalization.

To tackle these problems mentioned above, several more advanced technologies have been brought up. Firstly, Pose-indexed based multi-view method [6] for face alignment could perform better when processing images with various poses, occlusions, or illumination, when compared to other methods. Although using relevant shape as the initiation for face alignment is coarser than using the mean shape, it appears more robust and more accurate. Secondly, [15] raised a robust cascaded regression framework, and it can process images with severe occlusions and big head poses in the meanwhile. Nevertheless, with letting the real-time tracking detection possible and improving the algorithm to solve more sophisticated conditions in real world, it can be refined. Thirdly, instead of introducing complex classifiers, [16] has brought up a great idea that partitions the input domain in order to utilize simple classifiers which are more easy to control to tackle complex problems. On the one hand, by constraining the energy of input domain it can limit the complexity, and it could control the energy



directly. On the other hand, this way to partition can improve the performance of both linear support vector machine and classifiers in every partitioning. The disadvantage of this algorithm is that it does not consider the class distribution directly during the partitioning. At last, a new method named robust statistical face formalization [17] is able to estimate facial formalization reconstruction, facial landmark localization, pose-invariant facial recognition and face verification without constraint conditions. In this paper, a number of algorithms related with face recognition including face alignment, facial landmark detection, classification, facial landmark localization (as shown in Fig. 1) are discussed and analyzed.

The rest of the paper is organized as follows. Section 2 introduces some methods dealing with face recognition, including face alignment, face landmark detection, classification, and facial landmark localization in detail. The results of these methods are presented in Section 3. The challenges of these methods are illustrated in Section 4.

Finally, conclusions are drawn in Section 5.

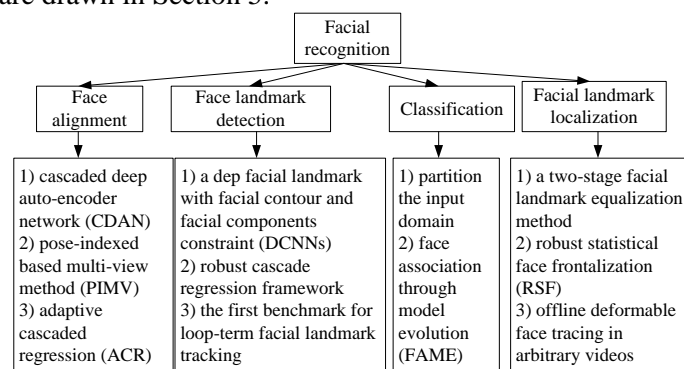


Fig. 1. Some algorithms about facial recognition

2 Techniques and Methods

2.1 Facial alignment

In [6], the author has presented pose-indexed based on multi-view method for face alignment. Firstly, it uses multi-view model based on piece to estimate facial poses of input face images. Due to facial poses, it can obtain corresponding face shape from pose-indexed shape search space for initialization in following iterative stages. It takes advantage of pose-indexed shape search space including a series of pose shape pairs for matching data from pose to face shape. Every angle in the space corresponds to a face shape. So the initial shape for face alignment consists of relevant shapes. (The author has presented a method to search a better initial shape from pose-index shape search space. This space consists of a series pose pairs used for matching face shape with poses. Every shape in space corresponds to angles used for pose indexed. So using pose indexed as the initial shape in following iterative stages.) This way to realize shape initialization can prevent the problems due to the initialization of the poor in the prediction. What's more, this method is able to perform better particularly when dealing with poses which are various, with occlusion, or under illumination. In a word, although this way for initialization is a little bit coarse, it performs better accuracy and robustness.

In [18], the author has proposed a new cascaded deep auto-encoder method for face alignment, including global exemplar-based deep auto-encoder networks (GEDAN) and a series of cascaded local deep auto-encoder networks (LDAN), as shown in Fig. 2. Firstly, in GEDAN, it uses low-resolution global facial image as an input to develop a primary facial landmark configuration. Secondly, in LDAN, it samples features in present landmark which develop index. Then, it uses increasingly improved image resolution to refine the landmark location. By this way, on the one hand, it can enhance robust for handling pose variation. On the other hand, it can improve the accuracy of landmark estimation. By using GEDAN, it can not only realize facial pose initiation better, but also

enlarge the capacity of deep auto-encoder networks in pose estimation to handle pose variation better to approach profile posture and posture under extreme illumination. By using LDAN can extract features related to posture information from local parts effectively. What's more, it can construct nonlinear model between local features and facial features.

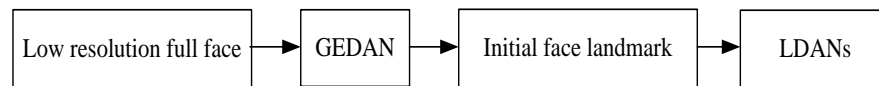


Fig. 2. Block diagram representing CDAN

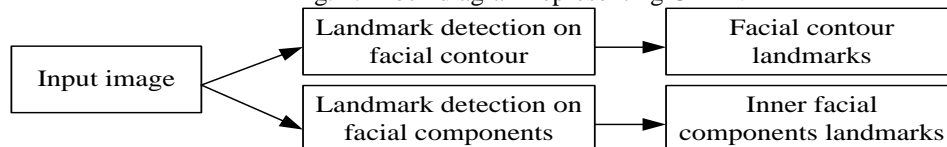


Fig. 2. Block diagram of two deep networks of DCNNs

In [19], the author raised a method combine production and resolution concept by weighted array depending on descent direction. This model adopts global parametric cascading based on descent direction. It allows the model to have better robust for very serious initiation and be able to converge to local minimization. In this way, it can realize the performance that resumes accurate landmark localization for good details of an article.

2.2 Face landmark detection

In [15], the author has raised robust cascaded regression framework which can deal with images with serve occlusion and big hand poses. Especially, the method is able to predict the occlusion of the landmark and the location of the landmark. Firstly, it uses mean face shape as initial landmark location and supposed that every dot in first iterative process is visible. Then, keeping updating the probability of the landmark visibility and landmark location iteratively substitutes for predicting binary landmark occlusion vector and the location of the landmark directly in order to realize better robustness. What's more, it definitely adds occlusion pattern for limitation to improve the performance of predicting occlusion. For landmark detection, it combines the visibility, local appearance and local shape to update their location iteratively. However, it can be refined in two directions in the future. On the one hand, it is able to be improved to realize real-time tracking detection. On the other hand, it can process more sophisticated conditions in reality by improving its algorithms.

In [20], the author has proposed a method for facial landmark detection based on facial contour and deep learning component constraints. The deep convolutional neural network in facial detection includes two deep networks, as shown in Fig. 3. One is used for detect the facial landmark detection of facial contour, and the other is utilized to detect landmark constraints of facial component. On the one hand, it divides deep convolutional neural network used for facial detection into two deep networks to improve the ability of detecting facial component landmark. On the other hand, it improves the ability of detecting facial contour by relating facial contour and component.

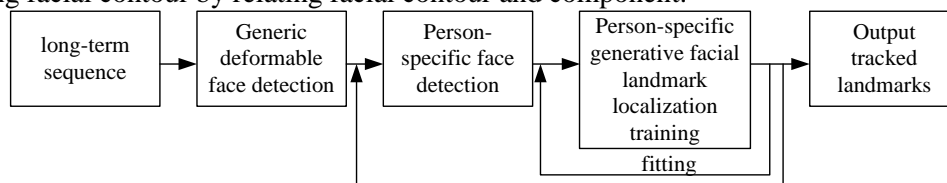


Fig. 3. The procedures of the method, FAME

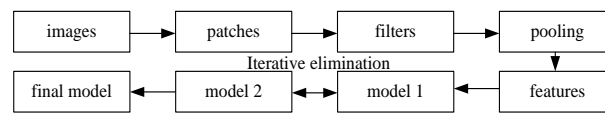


Fig. 4. Block diagram of the pipeline

In [21], the author has presented long-term facial landmark tracking standard, including over 110 recent videos with annotations firstly. And it has concluded the result of landmark tracking or detection in videos which nearly lasts for one minute firstly. What's more, this method performs better when there is only one person in the video with no occlusions or sever poses.

2.3 Classification

In [16], the author has raised a method, partitioning the input domain for classification to deal with complex questions. This method do not need utilize the number of partitioning as the prior by defining the biggest variance in every partitioning. In this way, classifiers are easier to control. Because constraining the input energy can limit the complexity of the questions, this method can control energy directly in every partitioning. What's more, it can improve the performance of linear support vector machine and classifiers in every partitioning to match more sophisticated neural networks' classifiers. However, due in partitioning process, a more interesting task in the future is to partition on the condition that packet data is supervised based on the type of label distribution.

In [22], the author has proposed a method, facial association through model evolution to reduce data to make the model related to name evolve, as shown in Fig. 4. Firstly, this method creates model by taking advantage of messy outcomes from name query, and this model evolves in continuous iterative processes so that name query is able to match with accurate face. Then, these models are used as faces with labels for database. What's more, the method, facial association through model evolution can remove outliers in the process of creating models on the condition that it can guarantee the diversity as much as possible.

2.4 Facial landmark localization

In [17], the author has put forward a new method, robust statistics face frontalization which only use a series of small range of frontal images to achieve the goal that associates subjective view reconstruction with landmark localization. The front view of an image has minimal rank in all poses so it can restore facial front view and design facial landmark. Hence this method is able to estimate facial front reconstruction, facial landmark localization, pose-invariant face recognition and face verification under unrestricted condition. What's more, robust statistics face frontalization is the first common and state-of-the-art method to realize landmark localization only using front image. In the future, it will be demonstrated that the method, robust statistics face frontalization is able to process all human faces, cat faces and other faces.

Table 1. Results of Selected Techniques in Face Recognition

No.	Method/Techniques	Results
1	CDAN	The CDAN is carried out on some different databases: HELEN, LFPW, and IBUG. The method can detect landmarks robustly, even there are various poses, expressions and partial occlusion.
2	PIMV	The PIMV has best accuracy compared with other algorithms for face alignment, the mean root squirt error (MRSE) on HELEN is 4.55, 3.23 on LFPW and 9.60 on IBUG.
3	Adaptive cascaded regression (ACR)	ACR was tested using the databases, LFPW, HELEN, AFW and IBUG. And its high AVC (the area under the curve) and low failure rate shows that ACR is highly accurate and robust.
4	DCNNs	The proposed method has less mean error, not only compared with conventional CNN, but also other state-of-the-art methods.
5	Robust cascade frame work	The proposed method performs well deal thing with images with seven occlusions or large head poses. The facial landmark detection errors is 5.49 on HELEN database and 3.93 on LFPW database.
6	The first benchmark for long-term facial landmark tracking	The method has presented the first comprehensive benchmark for assessing the performance of facial tracking methodologies in long-term videos.
7	Partition the input domain	There are significant improvement in classification results with the combination of both partitioning approaches compared with no partitioning or pixel-wise only partitioning.
8	FAME	The FAME is carries out on PubFig 83 database. With increasing iterative times, it performs more effective and improves accuracy.
9	A two-stage facial landmark localization method	The experiments were conducted on Ckf database. The proposed method has higher landmark wcalizatoin accuracy.
10	RSF	The RSF performs better than many other methods in reconstructing frontal view, landmark wcalizatoin, recongniting pose-invariant face and verifying face in in-the-wild situations
11	Offline deformable face tracking in arbitrary videos	The proposed method performs highly accurate landmark wcalizatoin and improves the true positive rate. Compared with other methods, it can still have a highly true positive rate in very challenging video category.

In [23], the author raised a common result of the face detection and landmark localization which can be used for recurrent train, accurate face detector varying from person to person and offline deformable face tracking landmark localization. This method proposes an accurate and effective channel (as shown in Fig. 5) which for long-term tracking offline sequence variability facial landmarks. This channel associate auto deformable face detection varying from person to person and the construction of landmark localization in offline mode so that it can track landmark captured from severely long-time sequences in any conditions. What's more, it is the first system which can eliminate false rate in all frames and drift problems in the videos of any length.

In [24], the author has proposed a two-stage facial landmark localization. Firstly, it locates facial features or landmarks by simulating facial detection, face alignment, and expression recognition at the same time. Then, it takes advantage of supervised descent method to refine the location of facial features.

3 Results

There are several methods have been discussed in this paper related to face alignment, facial landmark detection, classification and landmark localization. And the Table I will show the results of these methods.

4 Challenges

However, there are still some problems remained to be resolved. For example, the method partitioning the input domain for classification doesn't take class distribution in consideration directly. So in the future we can try to partition input domain under the conditions based on label distribution supervised data packet. In a word, the actual conditions are much more sophisticated just taking the images with occlusions and various poses for an example, so improving these algorithms constantly and realizing real-time tracking detection are quite essential in the future.

5 Conclusion

Recently, with the popularity on face recognition increasing, many excellent algorithms have been brought up in this particular domain. In this paper, some algorithms about face alignment, face landmark detection, classification and facial landmark localization are introduced in detail. From the

result provided by these algorithms, the ability of these techniques to deal with images has been improved dramatically and can realize higher accuracy and robustness. However, these algorithms still need to be perfected for being applied in more reality conditions.

References

- [1] Xiong X, Torre F D L. Supervised Descent Method and Its Applications to Face Alignment[C]// Computer Vision and Pattern Recognition. IEEE, 2013:532-539.
- [2] Cao X, Wei Y, Wen F, et al. Face Alignment by Explicit Shape Regression[J]. International Journal of Computer Vision, 2014, 107(2):177-190.
- [3] Ren S, Cao X, Wei Y, et al. Face Alignment at 3000 FPS via Regressing Local Binary Features[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2014:1685-1692.
- [4] Zhu S, Li C, Chen C L, et al. Face alignment by coarse-to-fine shape searching[C]// Computer Vision and Pattern Recognition. IEEE, 2015:4998-5006.
- [5] Burgos-Artizzu X P, Perona P, Dollar P. Robust Face Landmark Estimation under Occlusion[C]// IEEE International Conference on Computer Vision. IEEE Computer Society, 2013:1513-1520.
- [6] Qi H, Zhao Q, Wang X, et al. Pose-indexed based multi-view method for face alignment[C]//Image Processing (ICIP), 2016 IEEE International Conference on. IEEE, 2016: 1294-1298.
- [7] Yu X, Lin Z, Brandt J, et al. Consensus of regression for occlusion-robust facial feature localization[C]//European Conference on Computer Vision. Springer International Publishing, 2014: 105-118.
- [8] Zhu X, Ramanan D. Face detection, pose estimation, and landmark localization in the wild[C]//Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012: 2879-2886.
- [9] Ciresan D C, Meier U, Masci J, et al. Flexible, high performance convolutional neural networks for image classification[C]//IJCAI Proceedings-International Joint Conference on Artificial Intelligence. 2011, 22(1): 1237.
- [10] Karpathy A, Toderici G, Shetty S, et al. Large-scale video classification with convolutional neural networks[C]//Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. 2014: 1725-1732.
- [11] Asthana A, Zafeiriou S, Cheng S, et al. Robust discriminative response map fitting with constrained local models[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2013: 3444-3451.
- [12] Kazemi V, Sullivan J. One millisecond face alignment with an ensemble of regression trees[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014: 1867-1874.
- [13] Saragih J M, Lucey S, Cohn J F. Deformable model fitting by regularized landmark mean-shift[J]. International Journal of Computer Vision, 2011, 91(2): 200-215.
- [14] Tzimiropoulos G, Alabort-i-Medina J, Zafeiriou S, et al. Generic active appearance models revisited[C]//Asian Conference on Computer Vision. Springer Berlin Heidelberg, 2012: 650-663.
- [15] Wu Y, Ji Q. Robust facial landmark detection under significant head poses and occlusion[C]//Proceedings of the IEEE International Conference on Computer Vision. 2015: 3658-3666.
- [16] Romero A R, Jayawardena S, Cox M, et al. Partitioning the Input Domain for Classification[C]//Digital Image Computing: Techniques and Applications (DICTA), 2015 International Conference on. IEEE, 2015: 1-8.

- [17] Sagonas C, Panagakis Y, Zafeiriou S, et al. Robust statistical face frontalization[C]//Proceedings of the IEEE International Conference on Computer Vision. 2015: 3871-3879.
- [18] Weng R, Lu J, Tan Y P, et al. Learning Cascaded Deep Auto-Encoder Networks for Face Alignment [J]. IEEE Transactions on Multimedia, 2016, 18(10): 2066-2078.
- [19] Antonakos E, Snape P, Trigeorgis G, et al. Adaptive cascaded regression[C]//Image Processing (ICIP), 2016 IEEE International Conference on. IEEE, 2016: 1649-1653.
- [20] Baddar W J, Son J, Kim D H, et al. A deep facial landmarks detection with facial contour and facial components constraint[C]//Image Processing (ICIP), 2016 IEEE International Conference on. IEEE, 2016: 3209-3213.
- [21] Shen J, Zafeiriou S, Chrysos G G, et al. The first facial landmark tracking in-the-wild challenge: Benchmark and results[C]//Computer Vision Workshop (ICCVW), 2015 IEEE International Conference on. IEEE, 2015: 1003-1011.
- [22] Golge E, Duygulu-Sahin P. FAME: face association through model evolution[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2015: 43-49.
- [23] Chrysos G G, Antonakos E, Zafeiriou S, et al. Offline deformable face tracking in arbitrary videos[C]//Proceedings of the IEEE International Conference on Computer Vision Workshops. 2015: 1-9.
- [24] Wei W, Tian C, Zhang Y. A two-stage facial landmark localization method[C]//Orange Technologies (ICOT), 2015 International Conference on. IEEE, 2015: 157-160.