IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received May 26, 2020, accepted June 5, 2020, date of publication June 18, 2020, date of current version July 1, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3003332

Joint Alignment of Image Faces

GANG ZHANG [1], LIJIA PAN [1], JIANSHENG CHEN [2], YANMIN GONG [3], AND FUYUAN LIU [1]
[1]School of Software, Shenyang University of Technology, Shenyang 110023, China
[2]Department of Electronic Engineering, Tsinghua University, Beijing 100084, China
[3]Department of Electrical and Computer Engineering, The University of Texas at San Antonio, San Antonio, TX 78249, USA

Corresponding author: Gang Zhang (zhang_gang_1973@yahoo.com)

This work was supported by the National Natural Science Foundation of China under Grant 61372176.

ABSTRACT Researches on face alignment have made great progress, which benefits from the use of prior information and auxiliary models. However, that information lacks in a single face image has always affected the further development of these researches. The methods considering multiple face images provide a feasible way to solve the problem undoubtedly. Joint alignment where multiple face images are considered was presented in the paper. Face alignment was used for each face, and joint face alignment was used for optimizing the alignment results of all faces further. During joint alignment, both rigid variations of faces and non-rigid distortions were considered, however, they were regarded as two independent stages. Joint face alignment was a process where optimization was performed iteratively. In each iteration, both rigid variations and non-rigid distortions were performed sequentially, and moreover, the results of rigid variations were used as input of non-rigid distortions. At the stage of rigid variations, the key points of a face were divided into five groups to reduce the effect of global constraints which was imposed by face shape. After several iterations, the optimal solution of joint alignment can be obtained. The experimental results show that the joint alignment can obtain the optimal results than joint alignment using phased global rigid variations and non-rigid distortions and that using iterative global rigid variations and non-rigid distortions, and it can be used as a novel method for joint alignment.

INDEX TERMS Face analysis, face recognition, joint alignment, face alignment.

I. INTRODUCTION

As an important part in face analysis, face alignment can provide strong support for some applications related to faces, for examples, face analysis and face tracking. Its purpose is to determine the outer outline of a face and facial details. Due to the use of prior information and auxiliary models, performance of face alignment has been greatly improved. The representative methods contain active appearance model (AAM), constrained local model (CLM), supervised descent method (SDM), convolution neural network (CNN) and *etc*., in the field of face alignment [1]–[9]. Among them, the methods using deep learning have received the most attention currently, for example, CNN. These methods mainly consider a single face image. They pre-train a model and use it to make face alignment. During face alignment, only a single face is considered, and the prior information from the set of face images is used to train the model and constrain the alignment process. Joint alignment [10] is usually used as the optimization way to the alignment results of image faces.

The associate editor coordinating the review of this manuscript and approving it for publication was Mohammad Shorif Uddin.

It considers multiple faces at the same time, and optimizes these results by using the constraints among these faces.

The research of joint alignment can be traced back to Learned-Miller's congealing method [11], [12], where multiple images of the same object are considered. However, as it only considered in-plane rigid variations, the method cannot be used for face images. After all, faces do not only have global variations but also local non-rigid distortions. Besides, they also contain out-of-plane variations. The previous representative research is the method of Cootes *et al.* [13] in joint face alignment. The method considers global rigid variations of faces and local non-rigid distortions. It estimates global affine transformations by using the downhill simplex of Nelder-Mead [14], and determines local non-rigid distortions by using local transformations as well as gradient-based search. Besides, it uses the updatable mean face to impose the constrains among the faces. When the global affine transformations are made, all of the key points are considered in the faces.

Joint face alignment mainly focuses on face representation, optimization strategy and termination conditions. Sometimes, selection of the reference face is also considered.

Face representation can be divided into two kinds in joint face alignment, *i.e.* non-parametric and parametric. In the former, the set of the key points is used to describe the shape of a face and a matrix is used to store the shapes of multiple faces [15]. In the matrix, each row or each column corresponds to the shape of a face. In the latter, the shape of a face can be described by the parameters for global variations and local distortions [16]. The global variations describe the variations of face shape in the 3D space which contain translation, rotation and scale, and the local distortions describes the local variations of face shape caused by facial expressions.

Optimization relies on face representation. For non-parametric case, the positions of key points are considered. The Newton descent method and its variants are usually used to seek for the least square solution which can be obtained iteratively. During the iterations, initial alignment results which are obtained by using image face alignment are optimized further under the constraints of multiple faces. The effect is change of key points in positions. For parametric case, the parameter sets of multiple faces are updated iteratively by regression. The process for updating is affected by the constraints among multiple faces. The effect is change of the parameter values. The termination condition is a part of optimization and is used to determine when joint face alignment ends. For non-parametric case, the minimum rank of the matrix is usually used for the termination condition according to the characteristics of face representation. Besides, it can also be used that the difference between two consecutive changes in key points is less than a special threshold. Certainly, the maximum number of iterations can also be used. For parametric case, it can also be used that the difference between two consecutive changes in parameters is less than a special threshold. Besides, the maximum number of iterations can also be used. However, for the former, good effects are difficult to be obtained when the rank 1 of the matrix is used as the termination condition in the wild [17].

In principle, selection of the reference face will not affect the performance of joint alignment and any face can be used as a reference face, as long as the key points of the reference face do not deviate from the actual positions of the key points in each face too far. It is a feasible solution to use the mean face as an initial reference face. Besides, some methods for selecting the reference face are also presented. Anderson *et al.* [18] used the apparent similarity to cluster multiple faces, and used the graph structure to represent the intra-cluster and inter-cluster faces. A face was represented as a point in the graph and the relationship between the faces was represented by an edge. Then the minimum spanning tree was constructed, and its root was used as the reference face. Liao *et al.* [19] found that a large geodesic distance exists between a face and some faces on the Riemannian manifold represented by distortion transformations and it leads to alignment difficulties. So, the method used the apparent similarity to cluster faces, so that a small geodesic distance exists in a cluster and a large geodesic distance exists between clusters. Then it used tree structure to represent the clustering result and used the root of the tree as the reference face.

In current method using non-parametric representation for joint face alignment, the set of key points is usually regarded as a whole. Thus, it is easy to impose constraints by using facial shape models. Co-constraints among faces only act on the stage of global rigid variations. When non-rigid distortions are made, constraints imposed by using facial shape models can only be acted on individual face. The advantage to do this is that more strict constraints can be imposed, but the constraints also lead to local deviations of alignment result of individual face. Besides, constraints imposed at the stage of non-rigid distortions weakened.

Joint face alignment considers global variations of faces and non-rigid distortions in the paper. It uses non-parametric face representation. During the optimization, it uses the constraints among multiple faces to optimize the sets of key points gradually. It uses the maximum number of iterations as the termination condition. Besides, it uses the mean face as an initial reference face, and moreover, the reference face can be updated dynamically. The contributions of the paper are as follows. First, joint face alignment is an iterative process. Global variations of faces and non-rigid distortions are sequentially considered each iteration. If all of key points are considered in global transformations, that the constraints are too strict will lead to offsets of local alignment results. So the key points of each face are divided into five groups, and affine transformations are made for each group. Second, during joint alignment, if intermediate alignment results of the key points are close or up to the actual positions, it will lead to shrinking of the alignment results to continue iterating. Grouping key points of faces or using the maximum number of iterations as the termination condition will provide a solution to the problem. Third, although the faces from different people are only used, the method can also be used for multiple faces of the same person.

Joint face alignment considering global rigid and local non-rigid is described in section II. Joint face alignment considering point grouping is described in section III. The optimization process is described in section IV. Implementation details of joint face alignment is described in section V. The experiments and performance analysis are in section VI. Section VII concludes the paper.

II. JOINT FACE ALIGNMENT CONSIDERING GLOBAL RIGID AND LOCAL NON-RIGID

A face is different from a rigid object in that it does not only have rigid variations but also non-rigid distortions. Thus, during joint face alignment, constraints among faces not only need to be considered, but also rigid variations and non-rigid distortions of the individual face also need to be considered. Besides, a face contains shape and texture features. The former is used to represent faces in joint alignment, and the latter is used to constrain the optimization process of joint alignment. Suppose that $f_i(x, y)$ represent the i-th face image and $i \in [1,C]$ where C is the number of face images used in joint alignment, then joint face alignment can be represented as

$$\mathrm{JA(F)} = \mathrm{R(F)} + \mathrm{NR}(\mathrm{F}_j) + \mathrm{JC(F)} \tag{1}$$

where R( ) describes rigid transformations, F is the shapes of all faces used in joint alignment, NR( ) describes non-rigid transformations, F_j is the shape of any face used in joint alignment, JC( ) describes the joint constraints imposed among multiple faces. Besides, it needs to be noted that F is used to consider C faces. After all, rigid transformations and joint alignment optimization consider global shapes of C faces. It also needs to be noted that F_j is used to describe the shape of any face in C faces. After all, NR( ) considers individual variations of each face.

Here the joint face alignment which uses equation (1) is called as joint face alignment considering global rigid and local non-rigid. First, an alignment method for image faces, Supervised Descent Method, is used for each of the C faces to obtain the alignment result. The results from all faces are used as initial values of joint face alignment. Then a training set of faces which consists of front or near front faces is used to obtain the mean face shape, and the mean shape is used as the initial reference face. The global rigid transformations from the reference face to each of the C face shapes are determined by using RANSAC. Joint constraints are imposed among the C face shapes by using the reference face. Then non-rigid distortions are made for each of the C results after rigid transformations. Thus, non-rigid distortions of individual face are considered. The non-rigid distortion results of the C face shapes are superimposed by using General Procrustes Analysis, and the mean face shape is obtained. The mean face shape is closely related to the C face shapes, and is used to update the reference face. The above process is iterated until the joint alignment optimization of the C faces finishes. The optimization process of the joint face alignment considering global rigid and local non-rigid is described in Figure 1. Here {Face $i|i \in$ [1,C]} represents C faces and Face i is the i-th face, {Align $i|i \in$ [1,C]} represents alignment results by using SDM and Align i is the i-th result, {Rid $i|i \in$ [1,C]} is rigid transformation results of the C faces which are constrained by the reference face and Rid i is the i-th result, {Non-rigid $i|i \in$ [1,C]} represent non-rigid transformation results of the C faces and Non-rigid i is the i-th result.

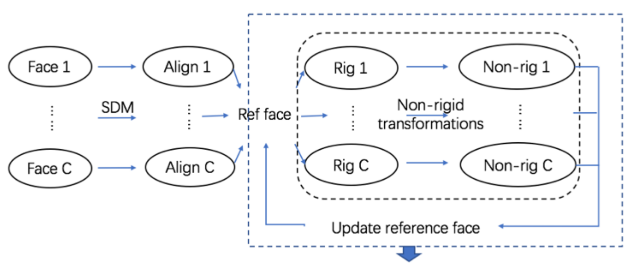


FIGURE 1. **Optimization process of joint face alignment considering global rigid and local non-rigid.**

III. JOINT FACE ALIGNMENT CONSIDERING POINT GROUPING

Joint alignment emphasizes simultaneous alignment of C face shapes. If the set of points which corresponds to each face shape is looked upon as a whole, the problem about overall offset takes place when joint face alignment considering global rigid and local non-rigid is made. This leads to the case that the alignment results in some of the C faces deviate from their actual positions locally. For the problem, a novel method is used. It divides the set of points in each face shape into five groups, outer contour, left eye, right eye, nose and mouth. Then rigid transformations are made for each group. Thus, the constraints are transferred from imposing on the point sets of face shapes to imposing on the point sets of each group, so that the problem about overall offset can be reduced. So the equation (1) can be represented as further

$$\mathrm{JA(F)} = \mathrm{R}(\mathrm{G}_i(\mathrm{F})) + \mathrm{JC}(\mathrm{G}_i(\mathrm{F})) + \mathrm{NR}(\mathrm{F}_j) + \mathrm{JC(F)} \tag{2}$$

where $\mathrm{G}_i(\mathrm{F})$ is the i-th group of face shape points, $\mathrm{JC}(\mathrm{G}_i(\mathrm{F}))$ is the constraints imposed among the i-th groups of C face shapes. Thus, the joint constraint in equation (1) can be divided into constraints among groups and those among face shapes.

During joint face alignment, each set of points are grouped for the reference face shape and the C face shapes. The way to grouping is shown in Figure 2, and these groups contain outer contour of a face, left eye, right eye, nose and mouth. First, the group (*a*) is considered for the reference face shape and the C face shapes, and the rigid transformations from the reference face shape to each of the C face shape are found by using RANSAC. The point group of the reference face shape is used to impose the joint constraints among the corresponding point groups of the C face shapes. Second, the groups (*b*), (*c*), (*d*) and (*e*) are considered sequentially for the reference face shape and the C face shapes. Then in the C face shapes, the five point groups of each face shape after rigid transformations are arranged according to the point order of the reference face shape. The rigid transformation results of these faces shapes can be obtained. Then non-rigid distortions are made for each of rigid transformation results of the C face shapes, so that non-rigid distortions of each face can be considered. Then the non-rigid distortion results of the C face shapes are superimposed by using General Procrustes Analysis, and the mean face shape can be obtained. The mean face shape is closely related to the C face shapes, and moreover, it is used to update the reference face shape. After the above process is iterated, the joint alignment optimization of the C faces finishes.

IV. OPTIMIZATION PROCESS

Joint face alignment is an iterative optimization process. As rigid transformations of faces and non-rigid distortions are two different stages, two different ways can be used to finish the iterative process. In the first way, iterations are made in each stage, rigid transformations and non-rigid distortions, and the result of the former is used as the input of the latter.

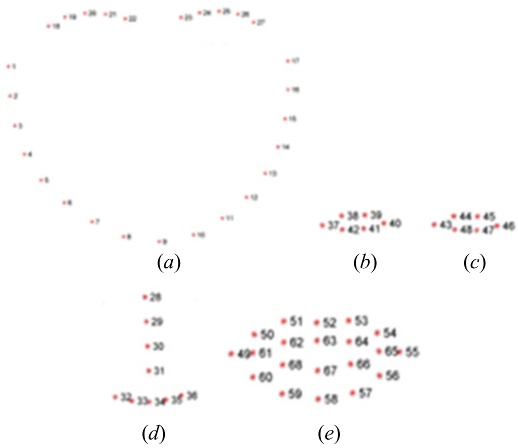


FIGURE 2. Five groups of shape points.

In the second way, iterations are made with two stages as a whole. The processes of two ways are shown in Figure 3. The results obtained by using image alignment methods for faces are {Align $j|j \in [1,C]$}. The results of the C face shapes after rigid transformations are {Rid $j|j \in [1,C]$}. OC, LE, RE, N and M correspond to the point groups of outer contour of a face, those of left eye, those of right eye, those of nose, those of mouth, respectively. The results of the C face shapes after non-rigid distortions are {Non-rig $j|j \in [1,C]$}. Here the second way is used.

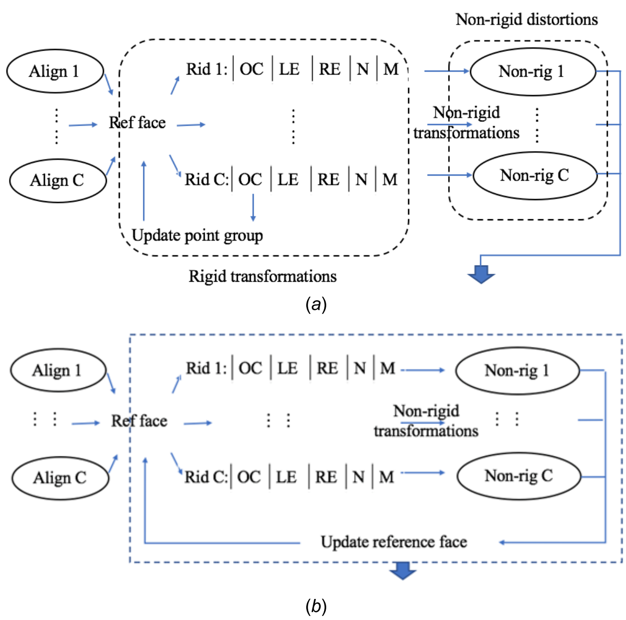


FIGURE 3. Iteration way (*a*) First way (*b*) Second way.

First, grouping is made for the reference face shape and the C face shapes. For any group i, RANSAC is used to find the rigid transformations from the reference face shape to any of the C face shapes. After the rigid transformations are made for five point groups, the results of these groups after rigid transformations are arranged according to the point order of the reference face shape. Thus, the results of the C face shapes after rigid transformations can be obtained. The advantage of grouping points is that joint constraints can be imposed among the same group of the C faces shapes. Besides, five groups of a single face are constrained by facial shape model. Then $JC(G_i(F))$ in equation (2) does not only contain the constraints among the same group of the C face shapes, but also implicitly contain the global constraints imposed by each face shape on five groups. So it can be represented as

$$JC(G_i(F)) = Ct(G_i(F)) + F_j(G) \tag{3}$$

where the constraint among the C point sets is $Ct(G_i(F))$ for the i-th group and the global constraint imposed by each face shape on five groups is $F_j(G)$.

To use the above iterative optimization process, non-parametric method can be used to describe the C face shapes. Here a matrix is constructed and each row of the matrix corresponds to the point set of a face shape. First, horizontal coordinates of points are stored and then vertical coordinates are done. Thus, the iterative optimization process can be regarded as the progressive process from the alignment results of image faces to actual positions, and variations of coordinate values can be seen.

As a termination condition of the iterative optimization process, the maximum number of iterations is used. Although the minimum rank of the matrix can be used as the termination condition for non-parametric representation and even the rank of 1, they are difficult to obtain good results for image faces in the wild.

V. IMPLEMENTATION DETAILS OF JOINT FACE ALIGNMENT

Here joint face alignment contains the contents as follows, obtaining initial alignment results of image faces, using non-parametric method to represent and store the C face shapes, iterative optimization of considering point grouping of face shapes, and termination condition for iterative optimization.

When initial alignment results of the C faces are computed, constrained local model (CLM), supervised descent method (SDM), convolution neural network (CNN) can be used. Here SDM is used, but the methods using deep learning can also be used. The purpose is only to obtain the alignment result of each face in multiple images.

Image face alignment is used for each of the C faces to obtain the 68 key points. The alignment results of the C faces will be used as an input of joint face alignment. Then the key points of the C face shapes are stored in a matrix where each row corresponds to a face shape, and moreover, the horizontal coordinates of these points first are stored and then the vertical coordinates. Then the 68 key points are divided into five groups. For each group, the rigid transformations from the reference face shape to each of the C face shapes are found by using RANSAC and the global constraints are imposed by using each face shape for its five group. For each of the C face shapes, the results of five groups after rigid transformations are arranged according to the order of points in the reference face shape. Afterwards, non-rigid distortions are made for the rigid transformation results of each of the C face shapes. Then General Procrustes Analysis is used to superimpose the non-rigid distortion results to obtain the mean face shape which will be used to update the reference face shape. The above process is iterated until the maximum number of iterations reaches. Thus, the optimization for joint alignment finishes. The whole optimization algorithm for joint alignment is described as followed.

Algorithm 1

Input: Alignment results by using image alignment for C faces

a) Using a matrix to store C face shapes
b) Grouping points for reference face shape and each of C face shapes
c) Finding rigid transformations from reference face shape to each of C face shapes for each point group
d) Transforming reference face by using rigid transformation parameters for each point group
e) Arranging five groups of points according to order of points in reference face shape for each of C faces.
f) Performing non-rigid distortions for rigid transformation results of each of C face shapes.
g) Using General Procrustes Analysis to superimpose non-rigid distortion results of C face shapes to obtain mean face shape.
h) Computing the 2-norm between reference face shape and mean face shape. If value is larger than a threshold, continue *i*), go to *k*).
i) Regarding mean face shape as a new reference face
j) Iterating steps *b*) to *h*) to maximum number of iterations.
k) Ending iteration.

Output: Joint alignment results of C faces

Here non-rigid distortions are to make the results of rigid transformations approach the actual positions of the corresponding points in the faces further, and moreover, are constrained by facial shape models. Here for each of the C faces, it selects the k points from the results of rigid transformations which are the close to their corresponding points in the initial alignment results. It needs to be noted that the set of the k points is the subset of the points obtained after rigid transformations. The parameters for mapping are determined from the results of rigid transformations to the initial alignment results by using the k points. Then the results of rigid transformations are mapped again by using these parameters.

VI. EXPERIMENTS AND PERFORMANCE ANALYSIS

Here non-parametric face representation is used, three methods for joint face alignment are considered. Of these methods, both global rigid and local non-rigid are considered. One is to consider rigid variations of the C face shapes used in joint alignment, then do non-rigid variations of each face shape. The method is named as that using phased global rigid and local non-rigid. Two is to regard global rigid and local non-rigid as two consequent steps and make iterations. The method is named as that using iterative global rigid and local non-rigid. These methods are mainly used in joint face alignment. The methods in the paper is named as that grouping points. It is different from the above methods in that points are grouped at the stage of rigid transformations. Thus, constraints can not only be imposed among the same group of the C face shapes, but also are imposed among five groups in each face shape. Besides, iterations are made with rigid transformations and non-rigid distortions as a unit. Here comparisons of performance will be made among these methods to joint face alignment.

The set of images for testing is the Helen face database [20]. From the database, the front faces and the faces which is close to front are selected to ensure that all the key points are visible in each face. Thus, the effect of large facial pose on three methods can be eliminated. In addition, the faces which are involved in joint alignment must meet two requirements. First, the local deviation of the alignment results only occurs in a small number of faces after the method for face alignment is used for the faces in multiple images. Second, for each of the key points, most of faces are aligned successfully or approximately. The faces which meet these conditions will be used to find the optimal method for joint face alignment.

The first testing is to analyze the performance of the joint face alignment using phased global rigid variations and local non-rigid distortions. Five face images are selected randomly from the testing set of faces each time. For each image, the rectangle surrounding the face is determined by using the AdaBoost classifier and then the initial alignment results can be obtained by using SDM. The alignment results of five face images is used as a group, and the joint face alignment using phased global rigid variations and local non-rigid distortions is made for the group. Some groups of visual results are shown in Figure 4 where each row corresponds to a group. Then the statistics are made. For each face of each group, the square sum of distances between the key points of the ground truth and those determined by join face alignment is computed, and then the values of all faces are accumulated in the group. The average of all groups is used as a measure.

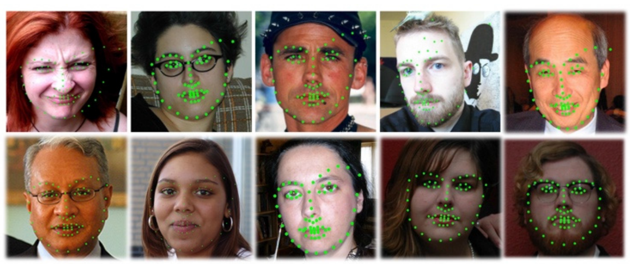


FIGURE 4. Results using phased global rigid variations and local non-rigid distortions.

The second testing is to analyze the performance of the joint face alignment using iterative global rigid and local non-rigid. It is different from the first testing. Five face images are randomly selected from the testing set of faces each time, and regarded as a group. For each image, the rectangle surrounding the face is determined and the initial alignment results are obtained. Then the joint face alignment using iterative global rigid and local non-rigid is made for the results of five faces. Some groups of visual results are shown Figure 5. Afterwards, the statistical method and the measure which is same as the first testing is used.

FIGURE 5. Results using iterative global rigid and local non-rigid.

The third testing is to analyze the performance of the joint face alignment considering point grouping in the paper. Five face images are randomly selected from the testing set of faces each time, and are used as a group. Each group of face images are considered one by one. For each image of any group, the rectangle surrounding the face is first determined, and then the initial alignment results are obtained. The joint face alignment considering point grouping is made for the results of five faces. Some groups of visual results are shown in Figure 6. Then the statistical method and the measure which is same as the first testing is used.

FIGURE 6. Results considering point grouping.

From figure 4, 5, 6, some results can be seen as follows. First, local alignment results of individual faces have larger offset when the method using phased global rigid variations and local non-rigid distortions is used. This is caused by weakening of global constraints imposed during the non-rigid distortion stage. When the method using iterative global rigid variations and local non-rigid distortions is used, better results can be obtained. The method can alleviate larger offset of local alignment results of individual faces. In contrast, the method considering point grouping in the paper obtains the best results, which benefits from two aspects. One is weakness of constraints during global rigid variations, which is due to use of point grouping. The other is strengthen of constraints during local non-rigid distortions, which is due to impose co-constraints among faces.

To observe joint alignment performance of different methods when the number of images participating in joint alignment is same and to testify the effect of varying the number of images on joint alignment performance, the number of images is set as 3, 5, 7, 9 and 11, respectively. The above process is used to determine statistics of performance for three methods, and all of the statistics are shown in Figure 7. The horizontal axe of Figure 7 is the number of images participating in joint alignment and the vertical axe is the measure value. According to the number of images, five groups are divided, three statistics of each group corresponds to the method using phased global rigid and non-rigid (GGP), that using iterative global rigid and non-rigid (GGS), and that grouping points in the paper (LGS), respectively. It can be seen that joint alignment performance of each method decreases with the number of images increased. In addition, the method grouping points obtains the best performance for each group.

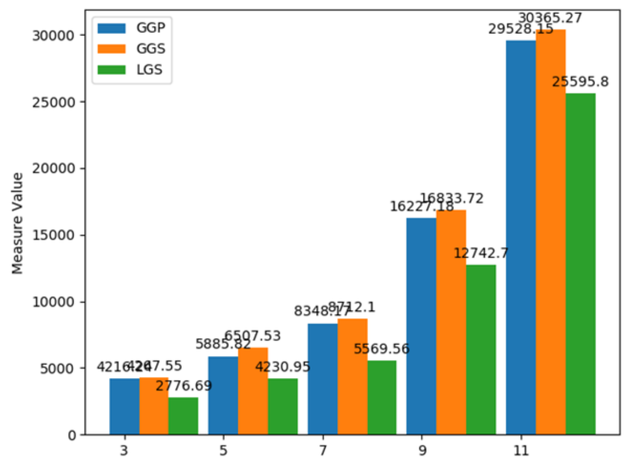


FIGURE 7. Statistics of performance for three methods using different number of images.

When the number of images participating in joint alignment is 5, the statistical results of different methods are shown in Table 1. It can be found that the joint face alignment using phased global rigid variations and non-rigid distortions is slightly better than that using iterative global rigid and non-rigid, and the joint face alignment in the paper obtains the optimal results.

The effect of three methods on local key points is shown in Figure 8. It can found that the key points which are close to the corner of left eye deviate slightly when the first method for joint face alignment is used. When the second method for joint face alignment is used, better results can be obtained. In contrast, the third method for joint face alignment obtains the best results. This means that it is valuable to consider rigid

FIGURE 8. Alignment results of local key points in three methods.

TABLE 1. Statistics of performance for three methods.

Method	Using phased global rigid and local non-rigid	Using iterative global rigid and local non-rigid	Using point grouping in the paper
Measure	5885.82	6507.53	4230.95

variations and non-rigid distortions. In addition, grouping points of faces are valuable.

From the perspective of computation efficiency, the method using phased global rigid variations and non-rigid distortions is close to that using iterative global rigid variations and non-rigid distortions. The method using point grouping in the paper is lower than these methods. Besides, some failures also occur when the method for joint face alignment in the paper is used. It was found that the problem is caused by the cases. For examples, the initial alignment results of faces are far away from their actual positions, and it are difficult to represent the non-rigid distortions of faces.

VII. CONCLUSION

An effective method for joint face alignment is proposed in the paper. The method does not only consider rigid variations of faces, but also non-rigid distortions. It looks upon both as two consecutive steps and makes iterations with both as a unit. In addition, points of each face are grouped at the stage of rigid transformations, so that the strict constraints imposed by facial shape models can be reduced. The performance of three method for joint face alignment is compared, using phased global rigid and local non-rigid, using iterative global rigid and local non-rigid, using point grouping in the paper. The experimental results show that joint face alignment considering point grouping obtain the best results, and joint face alignment using phased global rigid and local non-rigid is better than that using iterative global rigid and local non-rigid. Certainly, when joint face alignment in the paper is used, local deviation also arises. In the latter works, we will focus on that issue.

ACKNOWLEDGMENT

The authors would like to thank good suggestions from anonymous reviewers. The authors are also thankful to other researches for providing open source algorithms and the image sets.

REFERENCES

[1] I. Matthews and S. Baker, "Active appearance models revisited," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 135–164, Nov. 2004.

[2] D. Cristinacce and T. F. Cootes, "Feature detection and tracking with constrained local models," in *Proc. Brit. Mach. Vis. Conf.*, 2006, pp. 929–938.

[3] X. Xiong and F. De la Torre, "Supervised descent method and its applications to face alignment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 1–8.

[4] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

[5] A. Bulat and G. Tzimiropoulos, "How far are we from solving the 2D & 3D face alignment problem? (and a dataset of 230,000 3D facial landmarks)," in *Proc. ICCV*, Venice, Italy, Oct. 2017, pp. 1021–1030.

[6] H. Han, A. K. Jain, F. Wang, S. Shan, and X. Chen, "Heterogeneous face attribute estimation: A deep multi-task learning approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 11, pp. 2597–2609, Nov. 2018.

[7] F.-J. Chang, A. T. Tran, T. Hassner, I. Masi, R. Nevatia, and G. Medioni, "Deep, landmark-free FAME: Face alignment, modeling, and expression estimation," *Int. J. Comput. Vis.*, vol. 127, nos. 6–7, pp. 930–956, Jun. 2019.

[8] S. Tulyakov, L. A. Jeni, J. F. Cohn, and N. Sebe, "Viewpoint-consistent 3D face alignment," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 9, pp. 2250–2264, Sep. 2018.

[9] Y. Zhong, J. Chen, and B. Huang, "Toward end-to-end face recognition through alignment learning," *IEEE Signal Process. Lett.*, vol. 24, no. 8, pp. 1213–1217, Aug. 2017.

[10] S. Maqsood, U. Javed, M. M. Riaz, M. Muzammil, F. Muhammad, and S. Kim, "Multiscale image matting based multi-focus image fusion technique," *Electronics*, vol. 9, no. 3, p. 472, Mar. 2020.

[11] E. G. Learned-Miller, "Data driven image models through continuous joint alignment," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 2, pp. 236–250, Feb. 2006.

[12] R. Annunziata, C. Sagonas, and J. Cali, "Jointly aligning millions of images with deep penalised reconstruction congealing," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 81–90.

[13] T. F. Cootes, C. J. Twining, V. S. Petrovic, K. O. Babalola, and C. J. Taylor, "Computing accurate correspondences across groups of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 11, pp. 1994–2005, Nov. 2010.

[14] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.*, vol. 7, no. 4, pp. 308–313, Jan. 1965.

[15] X. Peng, S. Zhang, Y. Yu, and D. N. Metaxas, "Toward personalized modeling: Incremental and ensemble alignment for sequential faces in the wild," *Int. J. Comput. Vis.*, vol. 126, nos. 2–4, pp. 184–197, Apr. 2018.

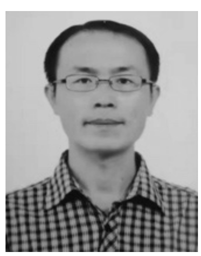
[16] A. Jourabloo and X. Liu, "Pose-invariant face alignment via CNN-based dense 3D model fitting," *Int. J. Comput. Vis.*, vol. 124, no. 2, pp. 187–203, Sep. 2017.

[17] Z. Q. Wei, Y. Yang, S. Zhang, and K. Yang, "Registration based on dual-feature Gaussian mixture model and dual-constraint spatial transformation," *J. Softw.*, vol. 29, no. 11, pp. 3575–3593, Nov. 2018.

[18] R. Anderson, B. Stenger, and R. Cipolla, "Using bounded diameter minimum spanning trees to build dense active appearance models," *Int. J. Comput. Vis.*, vol. 110, no. 1, pp. 48–57, Oct. 2014.

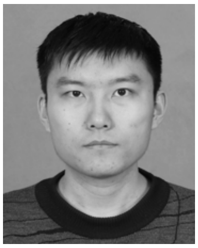
[19] S. Liao, D. Shen, and A. C. S. Chung, "A Markov random field groupwise registration framework for face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 4, pp. 657–669, Apr. 2014.

[20] V. Le, J. Brandt, Z. Lin, L. Bourdev, and T. S. Huang, "Interactive facial feature localization," in *Proc. ECCV*, Florence, Italy, Oct. 2012, pp. 679–692.



GANG ZHANG received the M.S. and Ph.D. degrees from the College of Information Science and Engineering, Northeastern University, China, in 2004 and 2009, respectively.

From 2011 to 2013, he held a postdoctoral research position with Tsinghua University, China. Since 2013, he has been a Teacher and an Associate Professor with the Shenyang University of Technology, China. His research interests include machine learning, face recognition, image retrieval, and artificial intelligence.

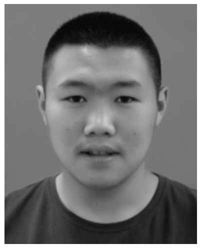


LIJIA PAN is currently pursuing the master's degree with the Shenyang University of Technology, China. His research interests include machine learning and face recognition.

JIANSHENG CHEN received the M.S. degree from the School of Information Science and Technology, Tsinghua University, China, in 2002, and the Ph.D. degree from the Faculty of Engineering, The Chinese University of Hong Kong, in 2006.

From 2007 to 2009, he held a postdoctoral research position with The Chinese University of Hong Kong. Since 2009, he has been a Teacher and an Associate Professor with Tsinghua University. His research interests include computer vision and recognition and understanding of images and videos.

YANMIN GONG received the M.S. degree in information and communication engineering from Tsinghua University, China, in 2012, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Florida, USA, in 2016. She is currently an Assistant Professor with the Department of Electrical and Computer Engineering, The University of Texas at San Antonio, San Antonio, TX, USA. Her research interests include machine learning, big data, the Internet of Things, cyber-physical systems, mobile computing, and wireless networks.



FUYUAN LIU is currently pursuing the master's degree with the Shenyang University of Technology, China. His research interests include machine learning and face recognition.