

Retraction

Retracted: Fast Modelling Algorithm for Realistic Three-Dimensional Human Face for Film and Television Animation

Complexity

Received 17 October 2023; Accepted 17 October 2023; Published 18 October 2023

Copyright © 2023 Complexity. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] L. Xu, "Fast Modelling Algorithm for Realistic Three-Dimensional Human Face for Film and Television Animation," *Complexity*, vol. 2021, Article ID 3346136, 10 pages, 2021.

Research Article

Fast Modelling Algorithm for Realistic Three-Dimensional Human Face for Film and Television Animation

Limin Xu 

Department of Software, Pingdingshan University, Pingdingshan, Henan 467000, China

Correspondence should be addressed to Limin Xu; 4352@pdsu.edu.cn

Received 29 April 2021; Revised 22 May 2021; Accepted 31 May 2021; Published 14 June 2021

Academic Editor: Zhihan Lv

Copyright © 2021 Limin Xu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Aiming at the face photos of film and television animation, this paper proposes a new fast three-dimensional (3D) face modelling algorithm. First of all, based on the LBF algorithm, this paper proposes a multifeature selection idea to automatically detect multiple features of the face. Secondly, in order to solve the shortcomings of slow training speed while achieving large pose face alignment, the regression-based CNN is selected as the algorithm to achieve fast convergence. Then, due to the influence of various factors, the extracted feature points are not completely correct, and Gabor features are used to screen the matching of feature points. Finally, by analysing the principle of 3DMM 3D face reconstruction, a single-view 3D face reconstruction method based on CNN is proposed. The experimental results show that the algorithm in this paper has good reconstruction performance and real-time performance and can realize the rapid modelling of human face.

1. Introduction

Human face is a very important communication channel for human beings, and it is the carrier of human's complex expressions and language such as happiness, anger, sorrow, and happiness. However, three-dimensional (3D) modelling is a basic problem in the field of computer vision and computer graphics [1, 2]. Realistic 3D face modelling has been a research hotspot in the field of computer vision and computer graphics for nearly 30 years, and it has a wide range of applications in film and television animation, human-computer interaction, video games, and communications [3, 4].

With the development of technologies such as computer graphics and virtual reality, especially the emergence of virtual announcers, research based on the design of virtual characters has been widely valued. The human face contains a large amount of characteristic information, which is the most important organ for distinguishing human beings. The computer simulation study of a specific human face is of great significance. 3D face models play an important role in applications such as film and television production, computer games, distance education, medical beauty, and intelligent recognition [5–7].

At present, the resolution of image collection is lower than that of digital cameras, the scanning effect of parts

with low reflectivity such as hair is slightly worse, and effective data cannot be obtained. Therefore, we propose a new algorithm that can effectively use these bad data. This paper mainly focuses on the field of film and television animation. First, based on the LBF algorithm, a multifeature selection idea is proposed to automatically detect multiple features of the face. Secondly, in order to solve the shortcomings of slow training speed while achieving large pose face alignment, the regression-based CNN is selected as the algorithm to achieve fast convergence. Then, due to the influence of various factors, the extracted feature points are not completely correct, and Gabor features are used to screen the matching of feature points. Finally, by analysing the principle of 3DMM 3D face reconstruction, a single-view 3D face reconstruction method based on CNN is proposed. This method can be directly applied to actual film and television animation production, which greatly improves the production efficiency of virtual actors.

The rest of our paper is organized as follows. Related work is introduced in Section 2. Section 3 describes the algorithm proposed in this paper. Experimental results and analysis are discussed in detail in Section 4. Finally, Section 5 concludes the whole paper.

2. Related Work

Realistic 3D face reconstruction can be divided into two ways [8]. One is to use more complex hardware devices such as a 3D scanner, supplemented by some simpler algorithms to obtain face geometry and texture data. The second is to obtain face images through ordinary cameras and use more complex graphics processing combined with computer vision algorithms to obtain face data.

2.1. 3D Face Modelling Using 3D Laser Scanner. The 3D laser scanner directly obtains face data based on the principle of triangulation. At present, the most famous 3D scanner is developed by Cyberwar, and many research groups use this equipment to conduct research [9, 10]. Cai et al. [11] used Laplace transform to preprocess the 3D face data collected by the scanner. After that, facial feature points are extracted in Laplace space, and the general face mesh model is adjusted to obtain a better target 3D face model. For nonuniform 3D face data, Jiang et al. [12] further improved the adaptive multiprecision fitting method, which can effectively solve the problems of holes and serious unevenness of data density. Vogiatzis et al. [13] used a 3D mapping scanner to obtain face information and used computer-based common geometry to transform between face surfaces and texture conformal mapping to construct a realistic 3D face.

2.2. 3D Face Modelling Based on Structured Light Scanner. Because 3D laser scanners are expensive and require high computer hardware, researchers have developed a structured light scanner composed of a projection light source and a camera. According to the principle of structured light ranging, it effectively solves the matching of corresponding points in binocular vision problem. Knoops et al. [14] established a rapid face acquisition system based on structural light and used the system to construct a 3D face data set including 120 people. On this basis, the 3D face recognition algorithm based on contour line and mesh surface matching is improved. Leng et al. [15] improved the registration method of the face image and the geometric model, modelled the texture of the target face organs, and conducted animation research on the basis of the obtained 3D face model. Liu et al. [16] proposed a Kinect-based 3D face modelling system. Hsieh et al. [17] used the colour sensor of the Kinect camera to extract facial landmarks. On this basis, the general face model is adjusted to register a single depth image, and then the target face model is established.

2.3. 3D Face Modelling Based on Computer Vision. The two-dimensional (2D) image contains depth information. The machine vision-based 3D face modelling method does not require prior knowledge of face shape. It can directly use geometry and triangulation knowledge to obtain the 3D face data on the 2D image sequence and reconstruct the reality sense of 3D human face. Hernandez et al. [18] proposed a method of reconstructing the 3D structure of the human face by using two frontal face images and using SFM. This

method uses the geometric symmetry and regularity of the common features of human faces to quickly and accurately find the matching points required by the SFM algorithm. Liao et al. [19] realized 3D face modelling based on a single video sequence based on computer vision spline fitting, affine transformation, and other technologies. Jiang et al. [20] proposed a 3D face modelling method based on binocular passive vision. This method uses the weak feature detection method in the image to extract the key points of facial features and estimate the disparity to realize the reconstruction of the 3D face.

2.4. Based on General Face Model. Islam et al. [21] used orthogonal images to reconstruct a 3D model of the target face and used a structured snake model to extract facial feature points. Zhong et al. [22] used the MPEG4 standard to define the facial feature points in the front and side images and used the radial basis function interpolation method to modify the general face model. Joo et al. [23] proposed a face image deformation method based on multiple images taken under multiple viewing angles, using computer vision methods to estimate camera parameters and at the same time recover the face pose and the 3D coordinates of 13 facial feature points. Sharma et al. [24] studied video-based 3D face modelling. This method uses the motion shape restoration method to obtain the 3D information of the face and uses the optical flow method to restore the 3D shape of the face from the video sequence. The whole process adopts Markov chain Monte Carlo method to optimize and approach the target face.

2.5. Based on 3D Deformation Model. Dai et al. [25] proposed a linear shape texture matching algorithm and a reverse composite image matching algorithm for the high computational complexity of the random Newton optimization algorithm. Fayaz et al. [26] proposed a method to restore the deformed model from three stereo images. The pixel point features on one image are mapped to another image through deformation parameters, and the feature difference between the projected point and the actual image is calculated. Hu et al. [27] studied the restoration of the deformation coefficient from the 3D image in the presence of occlusion. Manju et al. [28] used Ada-Boost to identify the face in the video, actively analysed the model to calibrate the key points of facial features, and restored the camera parameters according to the gold standard algorithm to obtain the deformation coefficient to reconstruct the face.

2.6. Methods Based on Statistical Learning. Kim et al. [29] connected the image greyness and the corresponding depth value as a vector and used a multivariate normal distribution to describe the statistical characteristics of the vector. Ruan et al. [30] used a regression method based on standard correlation analysis to estimate the depth image from the RGB image. The principle of this method is to find out two projection directions so that the original input and output data are projected to these two directions to maximize the

correlation. Zhou et al. [31] integrated face image grayscale and 3D surface shape feature space into a hybrid model and used partial least squares regression as a learning algorithm for depth estimation.

3. Fast 3D Face Modelling Algorithm

3D face models established in film and television animation often require fine details [32, 33]. A commonly used method is to map a 2D image to the surface of an object, which reflects the illumination properties of the face surface and shows detailed information such as skin tone, wrinkles, and hair colour. In this paper, we can obtain high-resolution facial texture images compared with the method based on the scanner hardware to obtain the texture directly by using the film and television special effects face photos.

The model established in this paper fully expresses the rich detailed information on the face surface and meets the requirements of film and television animation. The specific process of the algorithm is as follows:

- (1) Face alignment
- (2) Extract facial features
- (3) Filter the feature points
- (4) Face reconstruction

3.1. Face Alignment Network Model. In order to solve the above shortcomings, a visual block structure is introduced into the structure of the CNN model and several such structures are connected in series to form a shallow CNN cascade when the large attitude face uniformity is realized. This network structure uses the feature extraction structure of the previous visualization block to carry out the next visualization block, avoiding the need to construct multiple CNNs for multistep extraction. At the same time, this method of using only one CNN can also achieve end-to-end training and reduce the training time. Compared with the traditional CNN cascade network, it can converge faster in the training stage of the model.

General neural networks are based on a fixed topology for weight and threshold training. The cascaded neural network starts from a small network, adds hidden units in the training process, and finally forms a multilayer structure of waterfall structure [34–36]. The CNN network structure used in this paper is similar to the cascaded CNN, and each “shallow CNN” in the network structure is defined as a visualization block. In each block, there is a “visualization layer” based on the latest estimated parameters, which acts as a bridge between the visualization blocks. This design can avoid the shortcomings of the traditional cascaded CNN mentioned earlier and provides end-to-end training for model matching. The proposed network structure is composed of several visualization blocks connected as shown in Figure 1.

The input of the network is a single 2D face image, its corresponding initialized parameter $\text{Obj} = \{M, P\}$, and the output is the final estimated value of the target parameter. All the visualization blocks are based on the interconnected

backpropagation form to realize the optimization iteration of the network. Therefore, compared with the typical CNNs cascade, in the network training process, the network convergence can be achieved with fewer iterations.

A visualization block is mainly composed of a visualization layer, two convolutional layers, and two fully connected layers. The lengths of the two fully connected layers in each block are 800 and 236, respectively. Among them, the visualization layer generates a feature map according to the latest target parameter Obj . Each convolutional layer is followed by a batch normalization layer (BN layer) and a ReLU layer, which extract deeper features based on the output features of the previously connected visualization block and visualization layer. Between two fully connected layers, a ReLU layer and a dropout layer are connected behind the previous fully connected layer. The latter fully connected layer is used to estimate the updated target parameters, defined as ΔObj , and the output of the visualization block is deeper features and new target parameters ($\Delta\text{Obj} + \text{Obj}$).

The choice of loss function has an important influence on the effect of the model. Based on the particularity of the structure of the CNN network model in the paper, a total of two loss functions are used.

3.1.1. Weighted Parameter Distance Loss Function. The basic idea is to assign a corresponding weight to each parameter.

$$E_{\text{WPD L}} = \frac{\text{Obj}'^T \cdot \alpha \cdot \text{Obj}'}{1 + \alpha}, \quad (1)$$

$$\text{Obj}' = \text{Obj} - \Delta\text{Obj}. \quad (2)$$

3.1.2. Euclidean Distance Loss Function. The second loss function used in the network model is based on the synthetic 2D feature point error of the face, which is defined as follows:

$$E_O = \frac{f(\text{Obj} + \Delta\text{Obj}) + \text{Pre}}{1 + \alpha}. \quad (3)$$

Among them, the variable Pre represents the expected value of the 2D feature point position, and the function of the function f is to use the currently estimated 3D parameter value to obtain the corresponding 2D feature point output value.

3.2. Extraction of Facial Feature Points. The human face is composed of facial tissues such as eyes, nose, mouth, cheek contours, and the relative positions of these organs and tissues are fixed, so the corresponding feature point positions are also kept consistent. Therefore, in the geometric description of the human face, the geometric relationship of these organizations can be used as an important feature. Geometric feature description is to use a set of geometric feature vectors to represent the tissues and organs of the human face. The feature point extraction of the face is based on the contour of the face.

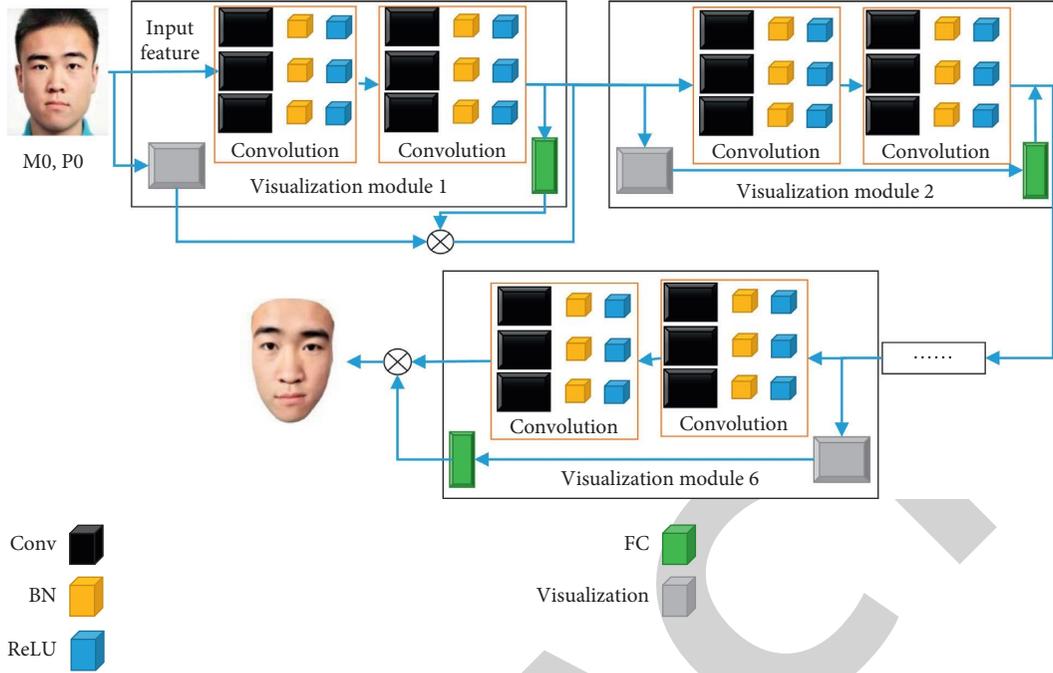


FIGURE 1: The neural network structure diagram proposed in this paper.

The core idea of LBF is to learn LBF features by randomly picking points near the feature points and performing residual operations on these points. Each feature point will learn several random trees, and the random trees will be combined to form a random forest. Then use the 0 or 1 feature vector generated by the random forest to represent the information of each feature point. When the random forests of these feature points are combined together, a global feature is formed. Finally, use this global feature to perform a global linear regression operation.

Based on the LBF algorithm, this paper proposes a multifeature selection idea. The improved LBF feature extraction process is shown in Figure 2. Applying the LBF algorithm to the operation of feature point extraction is to automatically find out the feature areas of the face such as eyes, nose, and mouth on the basis of face detection.

Given a face photo with n pixels, the variable $Q(a)$ represents a feature point in the picture, and the variable $E(a)$ represents a nonlinear feature extraction function. Expressed mathematically, find the smallest Δa of $P(a + \Delta a)$.

$$P(a + \Delta a) = \frac{E(Q(a + \Delta a)) - E(Q(a))}{E(Q(a + \Delta a)) + E(Q(a))} \quad (4)$$

It can be seen from formula (4) that when solving the regression key points of each step, it is necessary to know the output key point a of the previous level and the feature vector λk obtained from the surrounding positions.

$$a_{k+1} = E(a_k, \lambda_k). \quad (5)$$

At this time, formula (5) is the calculation formula for each step. In traditional learning, each learning task is learned together to improve the performance of classification. In our

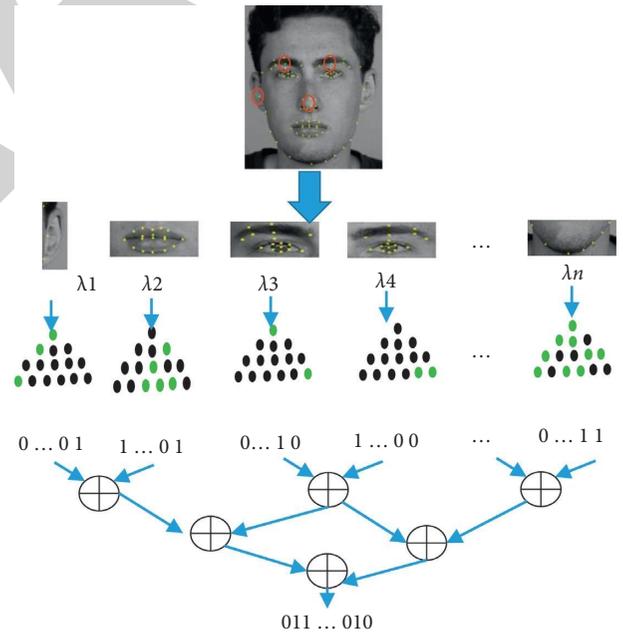


FIGURE 2: LBF feature extraction process.

method, the learning of a feature is a task. At this time, the objective function of equation (6) can be obtained according to the multifeature learning framework.

$$\arg \min \sum \sum \text{loss}(b_k, a_k) + P(a_k + \Delta a_{k-1}). \quad (6)$$

3.3. Screening of Facial Feature Points. The reason for feature point screening is that when extracting facial feature points from an image sequence, it will be affected by various factors

and the extracted feature points are not completely correct. For example, the self-occlusion of the face, the occlusion of the external environment, and the errors in the extraction of feature points using the LBF algorithm in this paper will cause the extracted feature points to be unsatisfactory.

The Gabor wavelet transform has similar characteristics to the stimulus response system of mammalian visual cortex cells. The features extracted by the Gabor filter have good direction selectivity and scale selectivity. And the influence of different illumination can be overcome by designing a suitable Gabor filter. However, if the high-dimensional information of these feature points is directly calculated, the accuracy of face reconstruction will be affected. In order to reduce the impact of wrong feature points on high-dimensional face modelling, this paper proposes a feature point screening method, which uses the Gabor feature method to judge the matching of feature points. First, the Gabor filter is used to extract the texture features of the facial area and the Gabor features of the image block where the candidate feature points are extracted. After that, the extracted feature and the texture feature of the reference image block are used as correlation coefficients to determine the matching of feature points.

The definition of the Gabor function is as follows:

$$G(x, y, z) = \frac{(\pi/2f^y)e^{(i\pi x/8)}}{(\beta-1)(\beta+1)} e^{(\pi^2 z^2/4f^y)} \left(e^{(\pi/2f^y)} e^{(i\pi x/8)} - e^{(-\beta^2/2)} \right). \quad (7)$$

Among them, the variable $z = (x, y)$ represents the pixel value at the coordinate (x, y) . The variable f is the frequency domain nuclear spacing factor, and the variable β is the ratio of the width of the window to the wavelength.

The facial area of the face in the i -th detection image is area (z) ; then its Gabor feature is

$$G'(x, y, z) = \text{area}(z) * G(x, y, z). \quad (8)$$

Select a face image from the image sequence, extract feature points from it, manually modify the feature points with errors, and use them as reference feature points. Calculate the correlation coefficient between the Gabor feature of the image block where the candidate feature point is located and the Gabor feature of the image block where the reference point is located, and average it.

3.4. 3D Face Reconstruction Based on CNN. The reconstruction of a 3D face is actually a matching process for the input image. That is, the 3D model is matched with the input image to obtain relevant parameters, and satisfactory matching results are obtained by adjusting the parameters.

By analysing the principle of 3DMM's 3D face reconstruction, a single-view 3D face reconstruction method based on CNN is proposed. Fusion of the features of the 3D face from different images of the same individual is used as the expected output of the CNN model to train a network model that can output a single 2D face image corresponding to the individual's 3D face shape. The network implementation block diagram constructed in this paper is shown in Figure 3.

The traditional 3D deformation model is based on the global vector to find the fitting coefficients, and the solution method adopts the least square method. Because the face has a complex physical structure and the face details are relatively concentrated, if the global function is used, the parameter will be ignored when fitting. The facial impact brought by the detail area prevents the face details from being fully restored. Therefore, this paper adopts the regionalized face reconstruction method to reconstruct each feature area separately. By solving the fitting parameters, the 3D feature points of the input 2D face image are obtained, the average model is selected to construct a 3D equation, the points of a certain partition of the average model are input as a function, the corresponding mapping point location information is solved, and the interpolation process is performed. Perform the above operations on each block to form a complete 3D model.

The main idea of the model is to construct a relationship equation between the feature points obtained from the input 2D image and the sample feature points of the 3D face data set. The model is divided into regions according to the feature points, fitting and solving are performed for each block, and finally, a realistic 3D model is formed. The following describes the model building process.

3.4.1. The Feature Point Calibration of the Input Image.

The feature point extraction is based on the input face image to be restored, using the LBF feature point detection algorithm to perform feature point extraction and calibration on the reconstructed 2D image. The calibration position and the calibration position of the sample in the 3D face data set realize a one-to-one correspondence.

3.4.2. Construct a Face Model Based on Feature Points.

According to the feature point information calibrated in the 3D face sample in the face alignment stage, a 3D face model is constructed. The 3D shape vector composed of the feature point set can be regarded as a subset of the 3D face feature vector.

3.4.3. Regional Block. The key to face reconstruction is to construct a realistic 3D model, which is mainly reflected in the details of the face, such as eyes, nose, and mouth. Traditional face reconstruction is to construct an overall fitting function for some feature points. This method cannot recover the face details very well, resulting in a lack of realism. Therefore, this paper proposes a block method to achieve regional reconstruction. Since there is a corresponding relationship between the feature points between the 2D input face image and the 3D face sample, the data samples are partitioned.

3.4.4. Reconstruction of the Face Model. According to the deformation model theory, b_{pq} represents the q -th block of the p -th 3D face model in the face sample, which can be expressed by the following formula:

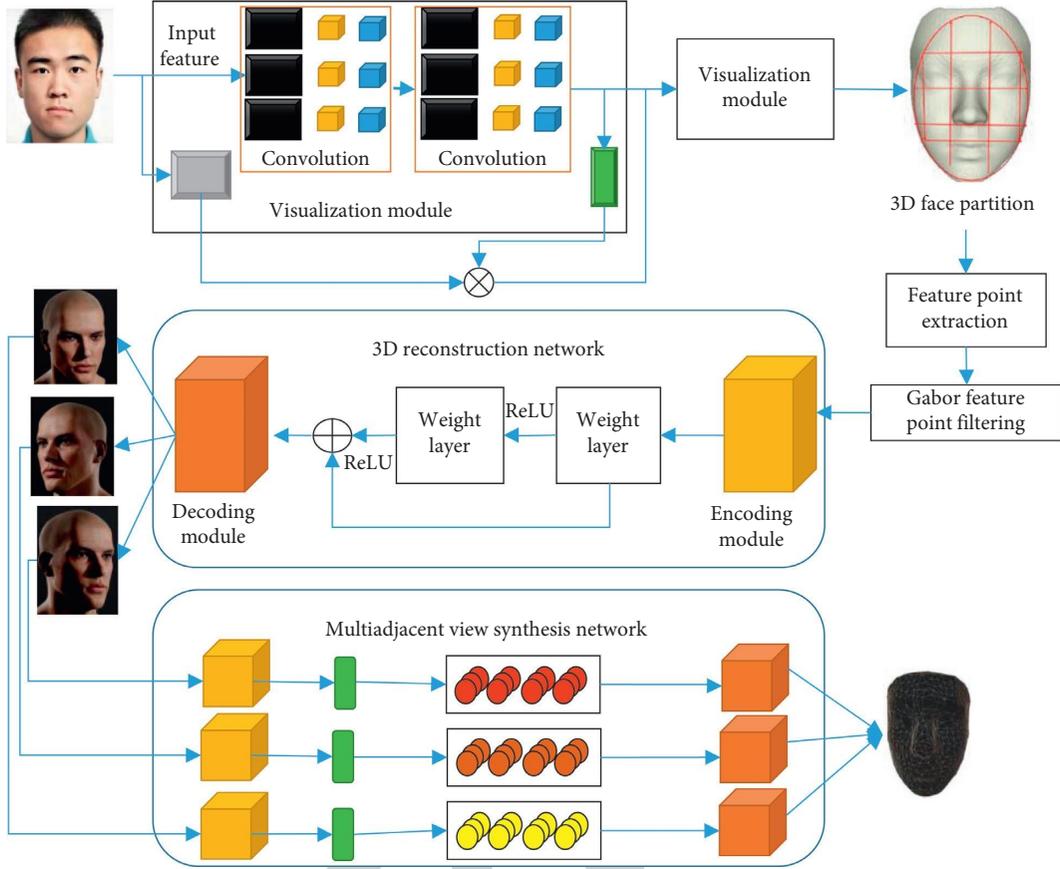


FIGURE 3: The implementation framework of the network constructed in this paper.

$$b_{pq} = ((x_i, y_i, z_i))^T, \quad i = 1, \dots, n. \quad (9)$$

In formula (9), n represents the total number of feature blocks of the 3D face model in the 3D face data set, p represents the p -th 3D face model, q represents the numbered block of each p -th model, and x, y, z , respectively, represent the 3D coordinate value in 3D Cartesian coordinates.

4. Results and Discussion

4.1. Experiment Description and Model Training. We use the server with 4 GPU cards as the experimental hardware equipment. Because Caffe's open-source framework contains a large number of pretrained models for use, this framework is chosen to implement the algorithm.

The LFW data set is currently a common test set for face recognition. The face pictures provided are all from natural scenes in life. Therefore, the difficulty of recognition will increase, especially due to factors such as multiple poses, lighting, expressions, age, and occlusion, even if the same person's photos are very different. And in some photos, more than one face may appear. For these multiface images, only the face with the centre coordinates is selected as the target, and other areas are regarded as background interference. The LFW data set has a total of 13,233 face images, and each image has a corresponding name. There are 5749

people in total, and most people only have one picture. The size of each picture is 250×250 , most of which are colour images, but there are also a few black and white face pictures. We use accuracy, error, and convergence for evaluation.

In the training phase, the weight attenuation of the network model is set to 0.005, the momentum factor is 0.99, and the initial value of the learning rate is $1e-6$. As the number of iterations increases, the learning rate continues to decrease. At the 20th iteration, the value of the learning rate is $5e-6$. At the 29th iteration, its value is $1e-7$. The training of the final model has completed a total of 40 iterations.

4.2. Performance Analysis of the Alignment Network. Based on the LFW data set, experiments are carried out on the proposed alignment network and compared with LPFA, PIFA, CCL, and RCPR algorithms. The NME indicator is used to quantitatively evaluate the final output of the algorithm. The experimental results are shown in Figure 4.

The experimental results show that the alignment error of the alignment network in this paper is smaller than other methods, which quantitatively shows that the alignment effect of the network in this paper is better. The error of each visualization block obtained based on the LFW data set is shown in Figure 5.

As the visualization block increases, the error value decreases more and more slowly, at the cost of increased running time. Therefore, the determination of the number of

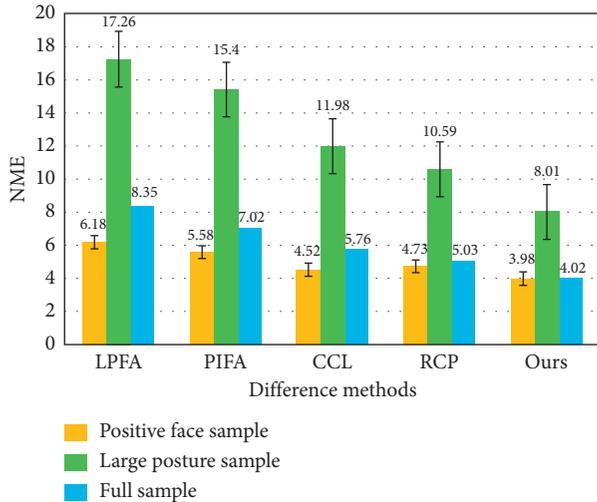


FIGURE 4: Experimental results of different alignment algorithms.

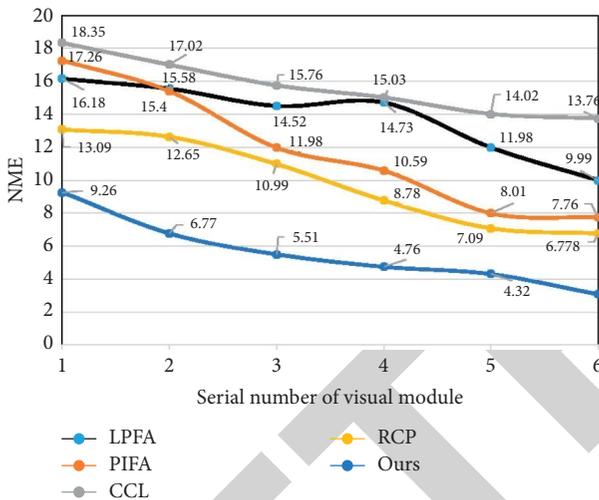


FIGURE 5: The error of each visualization block based on the LFW data set.

visualization blocks in the algorithm model is also based on the experimental results, while taking into account the running time of the algorithm and the effect of the algorithm.

In order to verify the performance of the proposed loss function, the effects of network models based on different loss functions, the effects of network models based on different loss functions were compared and analysed, including PDL, VDL, and WPD. The network models based on these three loss functions have been trained until they converge. The test error of the network model during each iteration is shown in Figure 6.

The experimental results show that the network model using PDL has a large error and cannot converge to a satisfactory result. The effect of VDL is slightly better than PDL, but the problem of morbid curvature makes it only focus on some parameters during the optimization process, which limits its performance. In comparison, WPD better fits the priority of each parameter and adaptively optimizes the weight of the parameter to achieve the best effect.

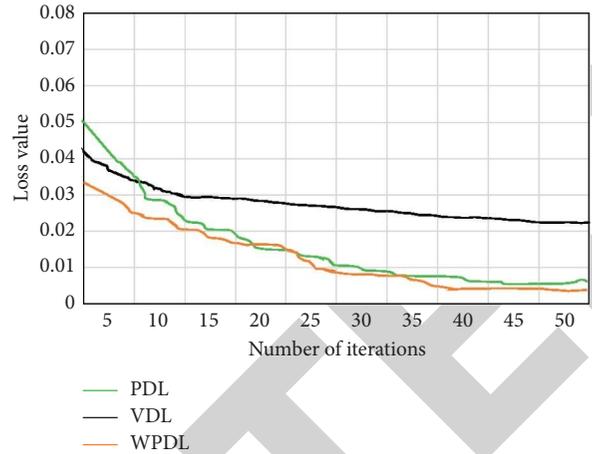


FIGURE 6: Test errors corresponding to different loss functions.

4.3. Performance Analysis of Multiple Feature Selection Algorithms. The traditional LBF uses a single feature to complete the positioning of the feature points. Such a result may cause the detection result to be inaccurate. A multi-feature selection LBF algorithm is proposed here. For this reason, the initialization models of the four algorithms of ESR, SIFI, LBF, and multifeatured LBF are compared in the accuracy of the algorithms. A comparison was made in LFW data set. The reason for using the above three data sets is that it was discovered during the experiment that there are many special cases of human faces in the LFW data set, with faces of various angles and resolutions. The performance of the algorithm is tested on the LFW dataset, and the result is shown in Figure 7. Experimental results show that the multifeature selection LBF model is significantly better than other models.

4.4. Performance Analysis of Gabor Filter. In order to prove the feasibility of using the Gabor filter to extract texture features, we can reconstruct the face model by judging the matching degree of feature points.

Select 20 people on the LFW data set, each with 6 face images. LBF is used to extract 13 contour feature points of each face image as test points, and Gabor features and GLCM are used to extract texture features of 5×5 image blocks where the test points and reference points are located. Among them, the reference point is a point selected from the image sequence after manual correction. According to the method introduced above, obtain the texture feature correlation coefficient between the feature point and the reference point. For the correlation coefficient greater than 0.9, it is considered to be the correct matching point. Table 1 lists the experimental results of these two methods. It can be seen that the Gabor feature discriminates feature point matching more reliably.

Figure 8 describes the 3D information calculated by directly using these feature points and all the correct matching points without judging the matching of the feature points. It can be seen that the wrong matching points have an impact on the accuracy of 3D face structure restoration.

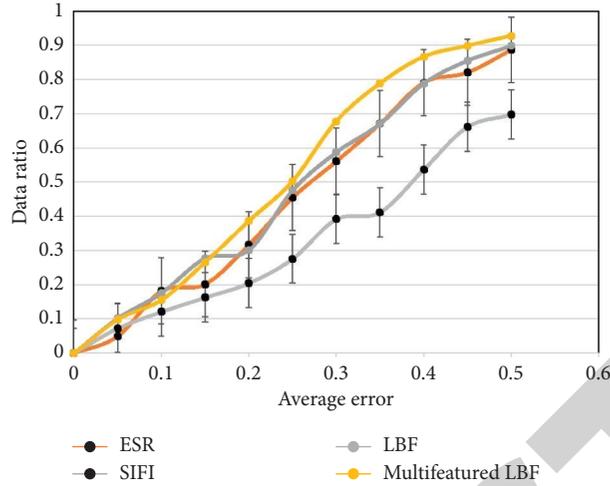


FIGURE 7: Comparison of initialization experiments of different algorithms.

TABLE 1: Accuracy judgment of contour feature points.

Methods	GLCM	Gabor
Number of test points	1866	1866
Number of judged exactly points	1321	1337
Accuracy rate	0.971	0.985

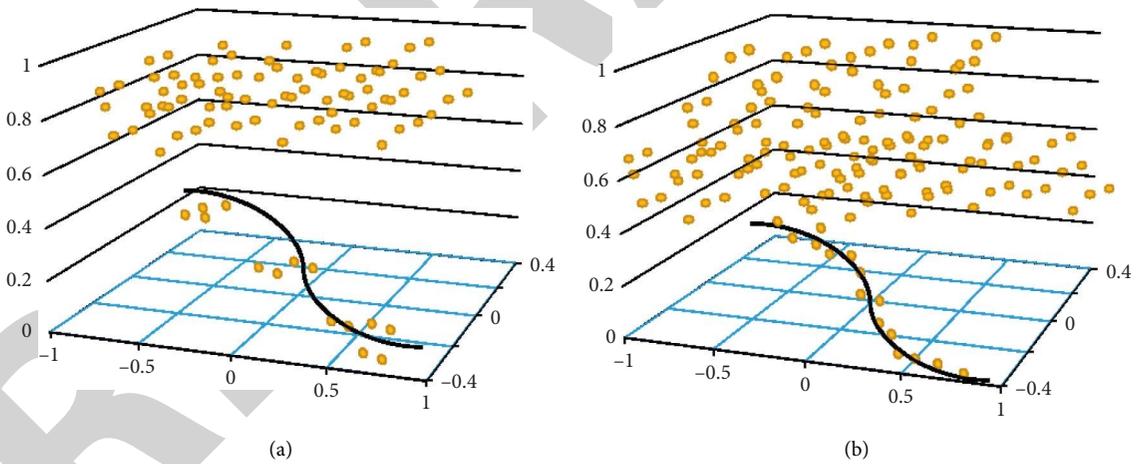


FIGURE 8: 3D sparse data. (a) 3D data obtained from wrong feature points and (b) 3D data obtained from correct feature points.

The above experimental results show that, after using the LBF algorithm to extract the feature points, it is an indispensable step to screen the correctness of the feature points, because it has a great impact on the accuracy of 3D face modelling. This paper uses the Gabor feature of the image block to effectively filter the feature points and remove the error generated when the LBF algorithm extracts the feature points.

4.5. Performance Analysis of 3D Face Reconstruction. Figure 9 shows the experimental results obtained by various reconstruction methods, including the literature [28],

3DMM, optical flow method, and R3DMM algorithm. The experimental results are shown in Figure 9.

The results show that, compared with 3DMM and R3DMM, the accuracy of 3D face reconstruction is improved by combining traditional 3DMM with deep learning. Compared with the same sampling deep learning and 3DMM combined with the R3DMM algorithm, the reconstruction accuracy is also improved. It proves that the reconstruction network proposed in this paper can significantly improve the effect of face reconstruction.

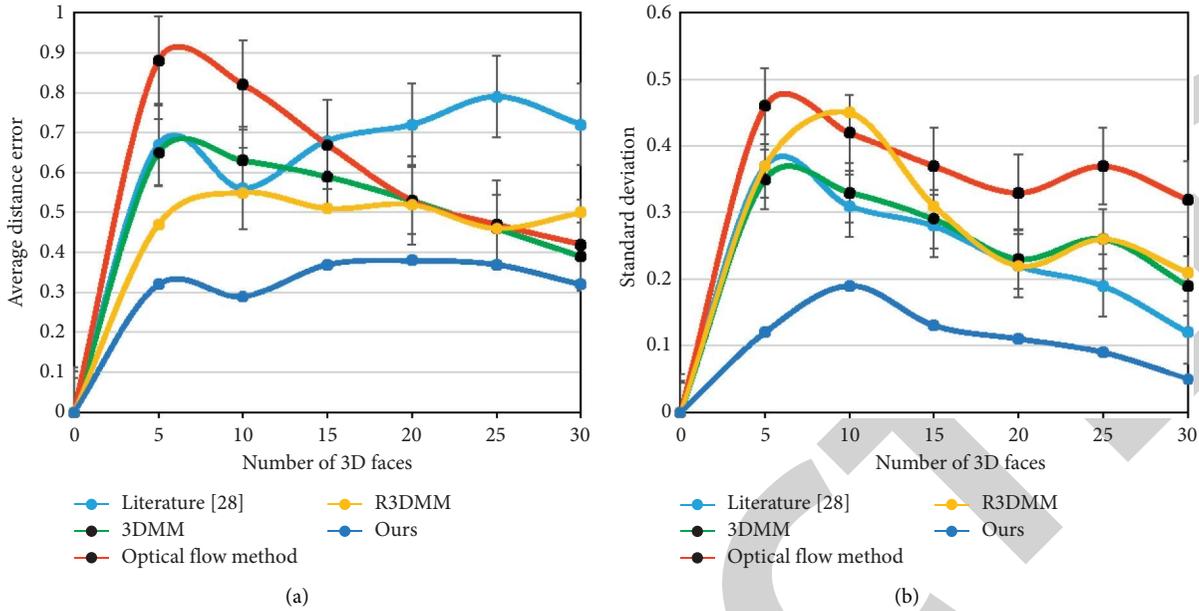


FIGURE 9: Errors of different reconstruction algorithms.

5. Conclusion

Realistic 3D face models have broad application prospects in many fields such as face recognition, film and television production, and medicine. Therefore, in the field of computer vision, 3D face modelling has always been one of the hotspots of research. This paper proposes a new fast modelling algorithm for 3D faces by using face photos of film and television special effects. First of all, based on the LBF algorithm, this paper proposes a multifeature selection idea to automatically detect multiple features of the face. Secondly, in order to solve the shortcomings of slow training speed while achieving large pose face alignment, the regression-based CNN is selected as the algorithm to achieve fast convergence. Then, due to the influence of various factors, the extracted feature points are not completely correct, and Gabor features are used to screen the matching of feature points. Finally, by analysing the principle of 3DMM 3D face reconstruction, a single-view 3D face reconstruction method based on CNN is proposed. The proposed improved method is tested separately and compared with the previous algorithm to verify the effectiveness of the improved method in this paper. And the method is applied to 3D face modelling, and the simulation experiment shows that it has better reconstruction performance.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known conflicts of interest or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank the Pingdingshan University Teaching Reform Research Foundation: College-level online open course “3D animation production”; Pingdingshan University Internet+ Education and Teaching Reform Foundation; based on “integral practice program approaching” online teaching exploration, according to the 3D animation production course as an example (Foundation code: PX-920434).

References

- [1] G. Song, J. Zheng, J. Cai, and T.-J. Cham, “Recovering facial reflectance and geometry from multi-view images,” *Image and Vision Computing*, vol. 96, p. 103897, 2020.
- [2] M. Chen, S. Lu, and Q. Liu, “Uniqueness of weak solutions to a Keller-Segel-Navier-Stokes system,” *Applied Mathematics Letters*, vol. 121, p. 107417, 2021.
- [3] I. Kemelmacher-Shlizerman and R. Basri, “3D face reconstruction from a single image using a single reference face shape,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 2, pp. 394–405, 2010.
- [4] R. Bodur, B. Bhattarai, and T. K. Kim, “3D dense geometry-guided facial expression synthesis by adversarial learning,” in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 2392–2401, Waikola, HI, USA, January 2021.
- [5] M. Chen, S. Lu, and Q. Liu, “Uniqueness of weak solutions to a Keller-Segel-Navier-Stokes model with a logistic source,” *Applications of Mathematics*, pp. 1–9, 2021.
- [6] J. Wen, J. Yang, B. Jiang et al., “Big data driven marine environment information forecasting: a time series prediction network,” *IEEE Transactions on Fuzzy Systems*, vol. 29, p. 1, 2020.
- [7] J. Yang, J. Zhang, and H. Wang, “Urban traffic control in software defined internet of things via a multi-agent deep reinforcement learning approach,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, pp. 1–13, 2020.

- [8] Y. Chen, F. Wu, Z. Wang, Y. Song, Y. Ling, and L. Bao, "Self-supervised learning of detailed 3D face reconstruction," *IEEE Transactions on Image Processing*, vol. 29, pp. 8696–8705, 2020.
- [9] P. B. Shah and Y. Luximon, "Review on 3D scanners for head and face modeling," in *Lecture Notes in Computer Science*, pp. 47–56, Springer, Cham, Switzerland, 2017.
- [10] S. Sharma and V. Kumar, "3D landmark-based face restoration for recognition using variational autoencoder and triplet loss," *IET Biometrics*, vol. 10, no. 1, pp. 87–98, 2021.
- [11] L. Cai and F. Da, "Estimating inter-personal deformation with multi-scale modelling between expression for three-dimensional face recognition," *IET Computer Vision*, vol. 6, no. 5, pp. 468–479, 2012.
- [12] L. L. Jiang, D. L. Maskell, and J. C. Patra, "Parameter estimation of solar cells and modules using an improved adaptive differential evolution algorithm," *Applied Energy*, vol. 112, pp. 185–193, 2013.
- [13] P. Vogiatzis, M. Ma, S. Chen, and X. D. Gu, "Computational design and additive manufacturing of periodic conformal metasurfaces by synthesizing topology optimization with conformal mapping," *Computer Methods in Applied Mechanics and Engineering*, vol. 328, pp. 477–497, 2018.
- [14] P. G. M. Knoop, C. A. A. Beaumont, A. Borghi et al., "Comparison of three-dimensional scanner systems for craniomaxillofacial imaging," *Journal of Plastic, Reconstructive & Aesthetic Surgery*, vol. 70, no. 4, pp. 441–449, 2017.
- [15] C. Leng, H. Zhang, B. Li et al., "Local feature descriptor for image matching: a survey," *IEEE Access*, vol. 7, pp. 6424–6434, 2018.
- [16] J. Liu, T. Tateyama, Y. Iwamoto et al., "A preliminary study of Kinect-based real-time hand gesture interaction systems for touchless visualizations of hepatic structures in surgery," *Medical Imaging and Information Sciences*, vol. 36, no. 3, pp. 128–135, 2019.
- [17] C. H. Hsieh, J. D. Lee, and C. T. Wu, "A Kinect-based medical augmented reality system for craniofacial applications using image-to-patient registration," *Neuropsychiatry*, vol. 7, no. 6, pp. 927–939, 2017.
- [18] M. Hernandez, T. Hassner, J. Choi, and G. Medioni, "Accurate 3D face reconstruction via prior constrained structure from motion," *Computers & Graphics*, vol. 66, pp. 14–22, 2017.
- [19] Y. Liao, P. Xiong, W. Min, W. Min, and J. Lu, "Dynamic sign language recognition based on video sequence with BLSTM-3D residual networks," *IEEE Access*, vol. 7, pp. 38044–38054, 2019.
- [20] D. Jiang, Z. Zheng, G. Li et al., "Gesture recognition based on binocular vision," *Cluster Computing*, vol. 22, no. 6, pp. 13261–13271, 2019.
- [21] M. N. Islam, M. Seera, and C. K. Loo, "A robust incremental clustering-based facial feature tracking," *Applied Soft Computing*, vol. 53, pp. 34–44, 2017.
- [22] Z. Zhong, M. A. Rong, H. E. Lei et al., "An efficient large-scale mesh deformation method based on MPI/OpenMP hybrid parallel radial basis function interpolation," *Chinese Journal of Aeronautics*, vol. 33, no. 5, pp. 1392–1404, 2020.
- [23] H. Joo, T. Simon, and Y. Sheikh, "Total capture: a 3D deformation model for tracking faces, hands, and bodies," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8320–8329, Lake City, UT, USA, June 2018.
- [24] S. Sharma and V. Kumar, "Voxel-based 3D face reconstruction and its application to face recognition using sequential deep learning," *Multimedia Tools and Applications*, vol. 79, pp. 1–28, 2020.
- [25] H. Dai, N. Pears, W. Smith, and C. Duncan, "Statistical modeling of craniofacial shape and texture," *International Journal of Computer Vision*, vol. 128, no. 2, pp. 547–571, 2020.
- [26] S. Fayaz, S. A. Parah, G. J. Qureshi, and V. Kumar, "Underwater image restoration: a state-of-the-art review," *IET Image Processing*, vol. 15, no. 2, pp. 269–285, 2021.
- [27] Y. Hu, Q. Chen, S. Feng, T. Tao, A. Asundi, and C. Zuo, "A new microscopic telecentric stereo vision system-calibration, rectification, and three-dimensional reconstruction," *Optics and Lasers in Engineering*, vol. 113, pp. 14–22, 2019.
- [28] D. Manju and V. Radha, "A novel approach for pose invariant face recognition in surveillance videos," *Procedia Computer Science*, vol. 167, pp. 890–899, 2020.
- [29] Y. Kim and S. B. Kim, "Optimal false alarm controlled support vector data description for multivariate process monitoring," *Journal of Process Control*, vol. 65, pp. 1–14, 2018.
- [30] Q. Ruan, M. Zhang, D. Lv, and H. Yang, "SAD and stock returns revisited: nonlinear analysis based on MF-DCCA and Granger test," *Physica A: Statistical Mechanics and Its Applications*, vol. 509, pp. 1009–1022, 2018.
- [31] Z. Zhou, J. Morel, D. Parsons, S. V. Kucheryavskiy, and A.-M. Gustavsson, "Estimation of yield and quality of legume and grass mixtures using partial least squares and support vector machine analysis of spectral data," *Computers and Electronics in Agriculture*, vol. 162, pp. 246–253, 2019.
- [32] C. Zhang and H. Wang, "Swing vibration control of suspended structures using the active rotary inertia driver system: theoretical modeling and experimental verification," *Structural Control and Health Monitoring*, vol. 27, no. 6, p. e2543, 2020.
- [33] Z. Niu, B. Zhang, J. Wang et al., "The research on 220 GHz multicarrier high-speed communication system," *China Communications*, vol. 17, no. 3, pp. 131–139, 2020.
- [34] Q. Jiang, F. Shao, W. Lin et al., "Optimizing multistage discriminative dictionaries for blind image quality assessment," *IEEE Transactions on Multimedia*, vol. 20, no. 8, pp. 2035–2048, 2017.
- [35] J. Yang, C. Wang, B. Jiang et al., "Visual perception enabled industry intelligence: state of the art, challenges and prospects," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 2204–2219, 2020.
- [36] K. Sim, J. Yang, W. Lu, and X. Gao, "MaD-DLS: mean and deviation of deep and local similarity for image quality assessment," *IEEE Transactions on Multimedia*, p. 1, 2020.