



# An Intelligent Proofreading for Remote Skiing Actions Based on Variable Shape Basis

Tie Li<sup>1</sup> · Jun Wang<sup>1</sup> · Katarzyna Wiltos<sup>2</sup> · Marcin Woźniak<sup>2</sup>

Accepted: 2 September 2024  
© The Author(s) 2024

## Abstract

The current proofreading algorithms for action regulation mainly recover the 3D structure and action information of non-rigid objects from image sequences by factorization. Most of algorithms assume that the camera model is an affine model. This assumption only holds if the size and depth of the object change very little relative to the distance from the object to the camera, which is in the case of fixed-shape basis. When the object is very close to the camera, this assumption causes a large reconstruction error. This paper solves this problem by the intelligent proofreading algorithms for remote skiing teaching actions based on variable shape basis. Firstly, the improved Retinex algorithm is used to enhance the multi-frame video images of skiing actions to make the action details more prominent. Then, measurement matrix is calculated after eliminating the translation vector by coordinate transformation. Under the condition of rank constraint, the measurement matrix is decomposed by singular value decomposition algorithm, and the correct shape basis structure of 3D action features can be obtained by using the variable shape basis. Finally, by randomly initializing a parameter, the optimized parameter and the least square algorithm are used to optimize the randomly initialized parameter further. The iteration until the convergence of the objective function can be used to calculate the deformation degree of the actions. The test results show that this algorithm improves the proofreading accuracy of action regulation in skiing teaching, and the proofreading results of various uploaded sliding actions are correct, which can be applied to remote skiing teaching and community learning.

**Keywords** Remote skiing teaching · image proofreading · Variable shape basis · parameter initialization · iterative convergence · objective function

## 1 Introduction

As a highly professional sport with high technical requirements, skiing is very important for the majority of learners to have standardized action skills [1]. However, the traditional alpine skiing teaching is limited by fixed time, location and resources, which hinders the learning effect of enthusiasts to a certain extent. With the vigorous development of mobile education technology, online teaching programs combining mobile technology and rich teaching resources are emerging, enabling learners to obtain high-quality guidance and watch real-time action demonstrations at any time, which

undoubtedly greatly improves the flexibility and efficiency of learning skiing [2, 3]. The combination of alpine skiing with online education and community learning is expected to establish an innovative teaching model to provide alpine skiing enthusiasts with a personalized, keen learning experience and deepen their understanding and mastery of alpine skiing skills. In this context, the development of relevant proofreading algorithms for action regulation becomes particularly critical in alpine skiing teaching [4, 5].

Traditional video playback and manual analysis algorithms are inefficient and miss critical information. In recent years, there are few researches on the proofreading algorithms for skiing actions, but there are many researches on the recognition and proofreading for skiing actions, and some research results have been achieved. For example, in reference [6], Taima et al. study a human action recognition algorithm based on initial attention, which more accurately identifies and understands human behaviors and actions by focusing on specific regions or features in video or sensor

✉ Marcin Woźniak  
Marcin.Wozniak@polsl.pl

<sup>1</sup> Institute of Physical Education and Training, Harbin Sport University, Harbin 150008, China

<sup>2</sup> Faculty of Applied Mathematics, Silesian University of Technology, 44100 Gliwice, Poland

data. Traditional algorithms based on initial attention have limitations in capturing key features of skiing actions, and they may overlook subtle but critical features such as center of gravity shifts, speed changes, and posture adjustments. In reference [7], Hoang et al. study a human action classification algorithm based on deep Gabor network. Deep Gabor network combines the advantages of Gabor filter and deep learning network, which can automatically learn and extract multi-scale and multi-direction features in images or videos so as to analyze human actions more accurately. But skiing is usually done outdoors and is significantly affected by light and weather conditions. These changes may lead to a decline in the quality of the video images, which further affects the effectiveness of the deep Gabor network in recognizing skiing actions. In reference [8], Anindita et al. study the human action recognition algorithm based on machine learning and meta-heuristic algorithms. The human action recognition is completed by machine learning technology, and the machine learning model is optimized by simulating the natural evolution process to improve its generalization ability. However, there are significant differences among different skiers in action regulation, skill levels and habits. Machine learning can have difficulty adapting to these differences, leading to errors in recognizing the skiing actions of different individuals. In reference [9], Palak et al. study an improved human action recognition algorithm based on action flow and deep learning. Through the deep learning model, this algorithm can extract and learn the complex features of human actions, and combine the continuity of action flow to accurately identify various human actions, including different actions in skiing. However, the action flow information of this algorithm is easily affected by the quality of human action images. Therefore, more diversified skiing data sets, including data of different scenes, different skills and different lighting conditions, need to be collected to improve the generalization ability of the model so as to ensure the effect of action recognition.

The variable shape basis is extracted from a large number of shape samples and can represent the common features of these shapes and the variations between them. In the field of machine learning vision, variable shape bases can be used to construct a shape space where each point represents a specific shape. By interpolating or deforming in this shape space, new shapes can be generated or existing shapes can be expanded and adjusted. In the remote skiing teaching, we use the collected skier movement data to train the three-dimensional movement feature variable shape structure and identify the correct movement pattern. For example, if a skier's turn action deviates from the standard action, the correct action can be simulated by adjusting the variable shape base and instructing the skier how to make the adjustment. Therefore, in order to solve the problem in the above algorithms, and quickly and accurately identify the actions

that do not meet the regulation, so as to more effectively carry out action guidance and correction, and improve the teaching efficiency. On the basis of the above research, the combination of alpine skiing teaching and online distance education, the introduction of shape-changing basic idea, the design of intelligent image proofreading algorithm for action adjustment in alpine skiing teaching, to provide a convenient, flexible and personalized learning path for alpine skiing enthusiasts, and provide a full range of teaching resources and technical support for their learning. In this paper, an intelligent proofreading method for remote skiing teaching action is proposed based on variable shape basis and factorization. The improved Retinex algorithm is used to enhance the multi-frame skiing action video image, which effectively improves the image quality, especially in the case of poor lighting and weather conditions, and reduces the impact of image quality decline on the recognition effect. By calculating the measurement matrix after eliminating the translation vector by coordinate transformation and using the singular value decomposition method, the key features of ski action can be captured more accurately, which is helpful to overcome the limitation of traditional methods based on initial attention in capturing fine features. The correct shape base structure of 3D motion features in ski images can be obtained by using the shaper base idea, which can adapt to the differences of different skiers' movement norms, skill levels and habits, improve the generalization ability of machine learning models, and reduce the errors in recognizing different individual ski movements. By randomly initializing parameters and using least square method for optimization, and iterating until the objective function converges, the deformation degree of the action can be calculated, and the accuracy of the normative calibration of the action can be improved, especially when dealing with complex features and continuous motion.

## 2 Design of Proofreading Algorithm For Action Regulation in Skiing Actions

### 2.1 Enhancement of Video Image Sequences of Skiing Actions

Skiing actions often have high-speed, complex features such as glide trajectories, posture adjustments, jumps, and spins. If the video image quality is poor, such as low resolution, blurring, jittery or affected by illumination, these key action details may be ignored or misjudged, which result in reduced recognition accuracy of action regulation [10, 11]. Through image enhancement, the clarity and contrast of the image can be improved, so that the action details are more prominent so as to improve the accuracy of recognition. The Retinex algorithm is able to simulate the adaptability of the

human visual system to light changes, thus maintaining the consistency of image details and colors under different lighting conditions. By separating the reflection component and light component of the image, the algorithm can effectively improve the contrast and detail visibility of the image, which is crucial for capturing key features in the ski action. However, while enhancing image contrast, it can lead to color distortion, especially when dealing with complex scenes. In order to overcome this shortcoming, a color correction mechanism is introduced to reduce color distortion while enhancing contrast. Therefore, the improved Retinex algorithm is used to enhance the video images of skiing actions uploaded by users. The algorithm of Gaussian pyramid and convolution sampling is used to sample the video image  $T$  of skiing actions. The sampling result  $F_{\pi+1}(i, j)$  is as follows:

$$F_{\pi+1}(i, j) = \sum_{n=1}^2 \sum_{m=1}^2 A(n, m) F_{\pi}(n - i, m - j) \tag{1}$$

where  $n, m$  represent the number of rows and columns of pixels sampled from the video images of skiing actions,  $i \in n, j \in m$ ;  $A(n, m)$  represent the number of pyramid layers and the number of Gaussian convolution kernels, respectively. The Gaussian weights are obtained by using the pixel differences and spatial spacing of the video images of skiing actions, and the regions where there are large differences between neighboring pixels and where the light intensity are the minimum or the maximum are the noisy regions, which are then processed by reducing the center weights. The video images of skiing actions are filtered on the x axis and y axis to obtain the filter function:

$$\Gamma(T) = F_{\pi+1}(i, j) \varpi(T) (a(\varphi, T) + \rho(\varphi, T)) d\varphi \tag{2}$$

where  $\varphi$  and  $\varpi(T)$  denote the set of spatial pixel points, the sum of weights respectively;  $a, \rho$  denote the distance function, similarity function respectively;  $d$  denotes the derivation. After completing the bilateral filtering [12], the noise in the image can be reduced, making the subsequent contrast adjustment smoother and more natural, and avoiding image degradation caused by noise. After the bilateral filtering, the contrast compression of the image can be expanded to ensure that the overall contrast of the image is properly adjusted while maintaining the details, which provides a better basis for the subsequent image amplification and interpolation. The double cubic difference algorithm is used to obtain accurate interpolation graphics by using the pixel information of the four neighborhood of the image after contrast compression, which can enhance the magnification effect of the ski action video image while maintaining the clarity and detail of the image. Finally, the processed image sequence is re-encoded into a video format, ensuring smoother and more realistic ski action output. The results of bi-cubic difference processing are as follows:

$$\zeta(x, y) = \Gamma(T) \sum_{i=0}^3 \sum_{j=0}^3 \varpi_{ij} \varpi(x) \varpi(y) \tag{3}$$

where  $\varpi_{ij}$  denotes the weight coefficient of image pixels,  $\varpi(x), \varpi(y)$  denote the horizontal coordinate weight and vertical coordinate weight of the video image pixels of skiing actions, respectively.

After completing the formula (3), the processed video images of skiing actions are subtracted from the original images in logarithmic space, which aims to achieve color constancy and detail enhancement of the video images of skiing actions based on scale invariance. The color recovery of the enhanced images is achieved by using the S-type function, which compresses the range of regions on both sides of the brightness in the images based on the features of the S-type function. The overall visual effect of the images are improved by the use of hierarchical processing, after the completion of logarithmic domain subtraction process of the video images of skiing actions, which need select the sigmoid function to implement the S-curve function stretching [13]. The sigmoid function is as follows:

$$S(T) = \frac{1}{1 + e^{-\zeta(x,y)}} \tag{4}$$

The information in the ski action video image sequence has different importance in different spatial scales. Although the S-curve function stretching process can improve the contrast of the image, it can lead to the loss of some details or excessive enhancement on a single scale. Therefore, multi-scale processing technology is introduced. With multi-scale processing, contrast adjustments can be made separately at different scales, and the results can then be fused to achieve a more balanced contrast enhancement. Multi-scale processing can capture both detailed features in the image (such as the tiny movements of athletes) and macro structures (such as the layout of the entire ski scene), thus improving the adaptability and accuracy of the proofreading system. After completing the S-curve function stretching process, the contrast formula of video image sequences of skiing actions is as follows:

$$Q(T) = \frac{255}{1 + e^{-f_1 f_2 S(T)}} \tag{5}$$

where  $f_1, f_2$  denote the contrast coefficient, pixel value coefficient,  $Q(T)$  denotes the enhanced video images of skiing actions.

## 2.2 Extracting 3D Feature from Skiing Actions of Non-rigid Objects

"Non-rigid body" refers to the process of skiing, the human body as a system, its various components (such as joints,

limbs) are not fixed, but can change shape and position. This change is due to the flexibility of the body's joints and the contraction and relaxation of muscles. The joints of the human body (such as the shoulder, elbow, hip, knee, and ankle) allow the limbs and torso to move in multiple directions. For example, the bending and extension of the knee joint, the rotation and abduction of the hip joint, etc., these joint activities allow the skier's body to adapt to different skiing movements and terrain changes. The shape of the limbs changes during skiing, such as the swing of the arms, the bending and stretching of the legs. These deformations are not only to maintain balance, but also to adjust the contact between the ski and the snow surface to achieve turning, slowing down and other actions. The twisting of the torso is an important part of the skiing action, helping to transfer power and adjust the center of gravity. The non-rigid nature of the torso allows the skier to effectively use the rotation of the body to control the direction of the ski when turning. Because the shape and position of various parts of the human body are constantly changing, the extraction of three-dimensional features of skiing movements becomes more complicated. In non-rigid systems, dynamic characteristics are crucial to understanding the nature of ski action. These dynamic features can help identify the skier's skill level and the accuracy of his movements.

The current proofreading algorithms for action regulation mainly use factorization to recover the 3D structure and action information of non-rigid objects from image sequences, most of which assume that the camera model is an affine model. This is a zero-order approximation (weak perspective) or first-order approximation (parallel perspective) to a projection model of real perspective [14], this assumption that only holds if the size and depth of the object change very small relative to the distance from the object to the camera, which is also in the case of fixed-shape basis. When the object is very close to the camera, this assumption causes a large reconstruction error. This paper uses the variable shape basis to solve the above problem.

According to the content of image enhancement introduced in the last chapter, the measurement matrix  $\bar{W} = M_{2F \times 3K} B_{3K \times P}$  is got after eliminating the translation vector of the video sequence by coordinate transformation. Under the condition of rank constraint, the measurement matrix  $\bar{W}$  in the skiing image sequences can be decomposed into  $\bar{W} = \bar{M}_{2F \times 3K} \bar{B}_{3K \times P}$ . In order to obtain the correct 3D shape structure and action information in the skiing image sequences, two tasks are still needed.

First of all, this decomposition result is not unique, for any non-singular matrix  $Q$  of  $3K \times 3K$  order, there is  $\bar{W} = (\bar{M}Q)(Q^{-1}\bar{B})$ . It is necessary to find a transformation matrix  $Q$  to make the skiing action matrix  $M$  still have blocky structure, eliminate affine ambiguity and finally get the exact result. Secondly, it is necessary to decompose the

action matrix  $\bar{M}$  further, so as to obtain the rotation matrix  $\bar{R}_i$  and the weighting coefficient  $\omega_{il}$ .

### 2.2.1 Solving the Transformation Matrix

According to the rank constraint, the rotation matrix  $\bar{R}_i$  keeps the unit orthogonal, then:

or

$$\begin{cases} (\bar{M}Q)_{2i-1}(\bar{M}Q)_{2i-1}^T = (\bar{M}Q)_{2i}(\bar{M}Q)_{2i}^T = \sum_{l=1}^K \omega_{il} \\ (\bar{M}Q)_{2i-1}(\bar{M}Q)_{2i}^T = 0 \\ (\bar{M}Q)_{2i-1}(\bar{M}Q)_{2i-1}^T - (\bar{M}Q)_{2i}(\bar{M}Q)_{2i}^T = 0 \\ (\bar{M}Q)_{2i-1}(\bar{M}Q)_{2i}^T = 0 \end{cases} \quad (6)$$

Where  $(\bar{M}Q)_{2i-1}$  and  $(\bar{M}Q)_{2i}$  are the odd and even lines, respectively.

The transformation matrix  $Q$  can be obtained by using linear least square algorithm for formula (6). In real-time or near-real-time ski action video application scenarios, computational efficiency is an important consideration. The reason of using linear least square method to solve transformation matrix is that it can be efficiently completed by matrix operation and is suitable for large-scale data processing. However, linear least squares can encounter robustness and accuracy problems when dealing with data containing noise, outliers, or nonlinear deformation. To overcome the limitation of linear least square method, a singular value decomposition algorithm is introduced, which can capture the intrinsic structure of data, including nonlinear deformation. Through singular value decomposition, the main patterns of the data can be extracted, which contain the characteristics of nonlinear deformation, thus helping to describe and deal with nonlinear problems more accurately. Then the action matrix  $M = \bar{M}Q$  and shape basis matrix  $B = Q^{-1}\bar{B}$  are obtained.

### 2.2.2 Solving the Rotation Matrix and Weighting Coefficient

The rotation matrix  $M$  and the weighting coefficient  $\omega_{il}$  are obtained by further decomposition of the action matrix  $M$ . In order to perform singular value decomposition on the action matrix  $M$ , it is necessary to re-adjust the two rows of each matrix  $M$  to form a matrix block, that is:

$$M = [ M_1^T \ M_2^T \ \dots \ M_F^T ]^T$$

Then for each matrix block:

$$M_i = \left[ \omega_{i1} \bar{R}_i \cdot \cdot \cdot \omega_{iK} \bar{R}_i \right]_{2 \times 3K}$$

In the matrix block,  $\bar{R}_i = \begin{bmatrix} r_{i1} & r_{i2} & r_{i3} \\ r_{i4} & r_{i5} & r_{i6} \end{bmatrix}$ . The matrix block  $M_i$  is rearranged as follows:

$$\tilde{M}_i = \begin{bmatrix} \omega_{i1} \\ \omega_{i2} \\ \cdot \\ \cdot \\ \omega_{iK} \end{bmatrix}_{K \times 1} \left[ r_{i1} \ r_{i2} \ \cdot \ \cdot \ \cdot \ r_{i6} \right]_{1 \times 6} = \bar{\Omega}_i \bar{\mathfrak{R}}_i \quad (7)$$

Obviously, the rank of the matrix is maximized to 1, so the singular value decomposition is performed on  $\tilde{M}_i$  to get the deformed rotation matrix  $\bar{\mathfrak{R}}_i$  and the weighting coefficient matrix  $\bar{\Omega}_i$ . The singular value decomposition is performed on  $\tilde{M}_i$  which has the number of F, and the deformed rotation matrix  $\bar{\mathfrak{R}}_i$  and weighting coefficient matrix  $\bar{\Omega}_i$  of each skiing image frame can be obtained.

The decomposition result of formula (7) is still not unique, and for any non-zero constant  $C$ , then there is  $\tilde{M}_i = \bar{\Omega}_i C \frac{1}{C} \bar{\mathfrak{R}}_i$ . And the constant  $C$  can be obtained by minimizing formula (8):

$$f(C) = \min \|\Omega_i - \Omega_{i-1}\|_F \quad (8)$$

Therefore,  $\mathfrak{R}_i = \frac{1}{C} \bar{\mathfrak{R}}_i$ ,  $\Omega_i = \bar{\Omega}_i C$  can be further obtained.

The rotation matrix  $\bar{R}_i$  of  $2 \times 3$  order can be obtained by adjusting the row vector  $\mathfrak{R}_i$ . However, the rotation matrix  $\bar{R}_i$  obtained by singular value decomposition cannot guarantee its unit orthogonal, so the rotation matrix  $\bar{R}_i$  needs to keep unit orthogonal. Assuming that singular value of  $\bar{R}_i$  is decomposed into  $U \Sigma V^T$ , where  $U$  and  $V$  are respectively the orthonormal matrices of  $2 \times 2$  and  $3 \times 3$  order,  $\Sigma$  is  $2 \times 3$  order matrix composed of singular values of  $\bar{R}_i$ . If  $\bar{R}_i$  is unit orthogonal, then both singular values are 1. So, the rotation matrix:

$$\bar{R}_i = U \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} V^T \quad (9)$$

This formula guarantees that the rotation matrix satisfies the property of unit orthogonal. According to the results of rotation matrix and weighting coefficient, the shape matrix  $\tilde{S}_i$  can be obtained from the formula:  $\tilde{S}_i = \sum_{l=1}^K \omega_{il} S_l$ .

### 2.2.3 Solving the Number of Shape Bases

In the distance skiing teaching, the movements of different learners will be very different. In the scenario of distance learning, real-time or near-real-time proofreading of actions is usually required. A suitable number of shape

bases can ensure that the model has enough adaptability to adapt to the movement characteristics of different learners, and has good generalization ability, that is, it can maintain the accuracy of proofreading on new learners or new movements. The number of shape bases directly affects the ability of the model to express the skiing action. If the number of shape bases is too small, the model will not be able to accurately capture the subtle changes of the action, resulting in inaccurate calibration results. If the number of shape bases is too large, although it can capture more details, it will also lead to excessive complexity of the model, increase the computational burden, and may lose the generalization ability due to over-fitting in practical applications.

The selection of the number named  $K$  of shape bases is very important for the application of the algorithm. If the number is too large, although the accuracy of the algorithm can be guaranteed, it will increase the calculation time of the algorithm. Moreover, it is not the case that the more shape bases are selected, the higher the accuracy of the algorithm will be, so it is unnecessary to select too many shape bases [15, 16]. However, if the number is too small, although the real-time performance of the algorithm is enhanced, it will increase the errors of the algorithm and even make the algorithm invalid. Therefore, too small number of shape bases should not be selected. At present, the selection of shape bases is fixed, so selecting a proper number of shape bases is very important for the effectiveness of the algorithm. The rank of the measurement matrix reflects the number of intrinsic linear independent components of the data. In a shape-based model, these linearly independent components can be considered candidates for a shape basis. By solving the rank of the measurement matrix, it is possible to determine the number of basic patterns or features present in the data that can be used as shape bases to describe and reconstruct the ski action. The selection of  $K$  in this paper is obtained by calculating the rank of the measurement matrix, it is dynamic, and that is:

$$K = \left\lfloor \frac{rank(W)}{3} \right\rfloor \quad (10)$$

where  $rank(W)$  represents the rank of the measurement matrix  $W$ , which reflects the amount of linearly independent information in the data. By this method, the algorithm can adjust the number of shape bases adaptively according to the features of the input data, so as to improve the real-time performance while maintaining the effectiveness of the algorithm and realize the improvement of the fixed number of shape bases strategy in the traditional algorithm. Therefore, this paper greatly improves the real-time performance of the algorithm on the premise of ensuring the effectiveness of the algorithm, which is also the improvement of the algorithm compared with the traditional algorithm.

### 2.2.4 Optimizing the Weighting Coefficient

According to the 3D shape of the features of the skiing action images reconstructed by the traditional factorization method, the reconstructed image will rotate 180 degrees along the Z axis in the process of continuous changes between the front and back frames [17–19]. In addition, the changes between the front and back frames of the skiing images are also discontinuous. At the same time, in skiing, due to the complex and changeable environment, image acquisition will indeed face challenges such as noise, occlusion and light changes. Shape continuity constraints aim to maintain the smoothness and consistency of object shapes over time series. This constraint is usually based on the assumption that the shape of the object will not change dramatically in a short period of time. When a part of the skier is obscured, the shape continuity constraint can use the shape information before and after the occlusion to infer the shape of the obscured part. For example, if a skier's arm is obscured by a tree branch, the shape of the obscured part can be estimated using the motion trajectory and shape changes of the arm before and after the occlusion. In the case of drastic illumination changes, shape continuity constraints can identify and correct the shape errors caused by illumination changes by comparing the shape changes of adjacent frames. The shape information before and after the illumination changes can be used to adjust the shape estimation to reduce the influence of the change on the shape extraction. Therefore, in the process of further optimization of the weighting coefficient, this paper adds the shape continuity constraint, so that the object can change continuously between the front and back frames, which reduces the ambiguity of the object actions.

In theory, the rank of the adjusted skiing action matrix  $\tilde{M}_i$  of  $K \times 6$  order is 1, but in the actual calculation process, its rank is far greater than 1 or even close to the maximum 6, so there are large errors in the rotation matrix and weighting coefficient obtained by direct application of singular value decomposition. In this paper, on the basis of the rotation matrix and the weighting coefficient obtained by singular value decomposition of the adjusted action matrix  $\tilde{M}_i$  of  $K \times 6$  order, the weighting coefficient matrix  $\Omega_i$  is further optimized by using the least square algorithm to minimize the following formula.

$$f(\Omega_i) = \min(\|\tilde{M}_i - \Omega_i \mathfrak{R}_i\|_F + \|\tilde{S}_i - \tilde{S}_{i-1}\|_F) \quad (11)$$

where  $\Omega_i$  represents the weighted coefficient matrix of the  $i$  frame skiing image,  $\mathfrak{R}_i$  represents the row vector composed of the first two rows of the rotation matrix of the  $i$  frame image,  $\tilde{S}_i$  represents the shape matrix of the  $i$  frame image,  $i = 1, 2, \dots, F$ .

Finally, this paper also limits the symbol of the weighted coefficient matrix so that the symbol of the weighted

coefficient of all frames is consistent, which makes the shape of the key features of the skiing actions keep moving in the same direction without flipping during the change process, and further reduces the ambiguity of skiing actions of the non-rigid objects.

## 2.3 Intelligent Image Proofreading for Skiing Action Regulation

### 2.3.1 Process of Obtaining Parameter Matrix of Action Structure

At present, the problem of using nonlinear optimization algorithm to solve parameters is essentially solving the description of rotation matrix  $Q$  and shift matrix  $T$ . In the algorithm of using L-M optimization, these two parameters are regarded as the unknown in the objective function of formula 11. First of all, the initialization algorithm of rotation matrix  $Q$  is introduced.

In this paper, the algorithm of obtaining the rotation matrix by the factorization is combined to carry out the initialization of  $Q$ . The measurement matrix  $W_{2F \times P}$  composed of the feature points  $\begin{bmatrix} u_{ij} \\ v_{ij} \end{bmatrix}$   $i = 1, \dots, F$   $j = 1, \dots, P$  of each frame of the skiing images is known ( $F$  represents the number of image frames,  $P$  represents the feature points of a frame image), and the goal is to find the 3D structure  $\tilde{s}_{i3 \times P}$  and rotation matrix  $R_{i3 \times 3}$  of each frame image.

In a shape matrix, a complex three-dimensional shape is usually represented as a combination of a set of simple basic shapes (i.e. shape bases). These basic shapes can be complex shapes (such as templates for human body parts). By adjusting the parameters of these basic shapes, a variety of complex three-dimensional shapes can be generated. Linear combinations allow multiple basic elements to be combined by weighting coefficients to generate new complex elements. In three-dimensional shape representation, linear composition means that by adjusting the weights of the shape bases, different shape variations can be generated. Non-rigid body motion refers to the fact that an object changes its shape and size during motion. In the analysis of human motion, non-rigid body motion is particularly important because the joints and muscles of the human body allow complex shape changes in various parts of the body. In order to capture the dynamic change of the shape of a non-rigid body, it is necessary to consider the shape change on a time series by applying a weighted linear combination of shape bases in the time dimension to generate a continuous sequence of shape changes. Therefore, assuming that the 3D shape of non-rigid objects is a weighted linear combination of shape bases, then there is  $\tilde{S}_i = \sum_{l=1}^K \omega_{il} S_l$ , where  $\omega_{il}$  represents the weighting coefficient,  $S_l$  is the shape basis, and  $K$  is the number of

shape basis. When  $K = 1, \omega_{il} = 1$ , which reflects the case of rigid objects.

Under the weak perspective projection model, there is:

$$\begin{bmatrix} u_{i1}, \dots, u_{iP} \\ v_{i1}, \dots, v_{iP} \end{bmatrix} = \bar{R}_i \left( \sum_{l=1}^K \omega_{il} S_l \right) + \bar{T}_i e_n^T \tag{12}$$

where  $\bar{R}_i$  are the first two rows of the rotation matrix  $R_{i3 \times 3}$ ,  $\bar{T}_i$  are the first two elements of the translation vector  $T_{i3 \times 3}$ ,  $e_n^T = [1, \dots, 1]_{1 \times n}$ ,  $\bar{W} = M_{2F \times 3K} B_{3K \times P}$  is obtained through the transformation of the coordinate for formula (12) with origin as the centroid.

The matrix  $M_{2F \times 3K}$  contains information about the rotation matrix. The rotation matrix  $\bar{R}_i$  and weighting coefficient  $\omega_{il}$  are obtained by further decomposition of the skiing action matrix  $M$ . In order to perform singular value decomposition of the action matrix  $M$ , it is necessary to re-adjust the action matrix  $M$  and form a matrix block with every two rows for  $M$ , that is:

$$M = [ M_1^T \ M_2^T \ \dots \ M_F^T ]^T$$

Then for each matrix block:

$$M_i = [ \omega_{i1} \bar{R}_i \ \dots \ \omega_{iK} \bar{R}_i ]_{2 \times 3K}$$

In the matrix block,  $\bar{R}_i = \begin{bmatrix} r_{i1} & r_{i2} & r_{i3} \\ r_{i4} & r_{i5} & r_{i6} \end{bmatrix}$ . The matrix block  $M_i$  is rearranged as follows:

$$\tilde{M}_i = \begin{bmatrix} \omega_{i1} \\ \omega_{i2} \\ \vdots \\ \omega_{iK} \end{bmatrix}_{K \times 1} [ r_{i1} \ r_{i2} \ \dots \ r_{i6} ]_{1 \times 6} = \bar{\Omega}_i \bar{\mathfrak{R}}_i \tag{13}$$

Obviously, the rank of the matrix is maximized to 1, so the singular value decomposition is performed on  $\tilde{M}_i$  to get the deformed rotation matrix  $\bar{\mathfrak{R}}_i$  and the weighting coefficient matrix  $\bar{\Omega}_i$ . The singular value decomposition is performed on  $\tilde{M}_i$  which has the number of F, and the deformed rotation matrix  $\bar{\mathfrak{R}}_i$  and weighting coefficient matrix  $\bar{\Omega}_i$  of each skiing image frame can be obtained.

The decomposition result of formula (13) is still not unique, and for any non-zero constant  $C$ , then there is  $\tilde{M}_i = \bar{\Omega}_i C \frac{1}{C} \bar{\mathfrak{R}}_i$ . And the constant  $C$  can be obtained by minimizing formula (14):

$$f(C) = \min \| \bar{\Omega}_i - \bar{\Omega}_{i-1} \|_F \tag{14}$$

Therefore,  $\bar{\mathfrak{R}}_i = \frac{1}{C} \bar{\mathfrak{R}}_i$ ,  $\bar{\Omega}_i = \bar{\Omega}_i C$  can be further obtained.

The rotation matrix  $\bar{R}_i$  of  $2 \times 3$  order can be obtained by adjusting the row vector  $\bar{\mathfrak{R}}_i$ . However, the rotation matrix

$\bar{R}_i = \begin{bmatrix} r_{i1} & r_{i2} & r_{i3} \\ r_{i4} & r_{i5} & r_{i6} \end{bmatrix}$  is obtained from  $\tilde{M}_i$  by singular value decomposition. Since the quaternion algorithm is used to solve the rotation matrix. The rotation matrix of each frame partially uses three parameters to express  $R_i = [a_i, b_i, c_i]$ . The exponential mapping of the rotation matrix according to the quaternion algorithm displays as  $a_i = r_{i6}; b = -r_{i3}; c = r_{i2}$ . Then  $R_i = [a_i, b_i, c_i]$  is regard as the initial value of the rotation matrix  $Q_i$  in the  $i$  frame image. After getting  $Q_i$ ,  $Q_i$  could be brought it into the objective function. Then a set of values of  $T_i$  called  $T_i'$  are got by using a random initialization algorithm and a nonlinear optimization algorithm on the objective function. In this way, after two rounds of nonlinear optimization,  $T_i'$  and  $Q_i$  will be used as the initial value of the parameter matrix of the de-action structure.

### 2.3.2 Calculating the Deformation Degree of Actions

After the image feature coordinates are transformed so that the origin of the image coordinate system is located at the centroid of the object and the translation vector is eliminated, the measurement matrix  $\bar{W} = M_{2F \times 3K} B_{3K \times P}$  is got. Obviously, the rank of the measurement matrix is not greater than  $3K$  (assuming  $2m > 3K, n > 3K$ ). Therefore, the value of  $K$  is of great significance to the estimation of shape and action parameter of action image sequences, which is an important parameter reflecting the deformation degree of non-rigid object shape [20–22]. The estimation of the deformation degree of a non-rigid object is actually an estimation of the value of  $K$ . The presence of noise makes the position of the measured feature points change randomly, which affects the recognition of the value of  $K$ . In addition, when the feature points are lost, the lost feature points must be recovered before the value of  $K$  can be further estimated, otherwise, the value of  $K$  cannot be estimated. To solve these problems, the estimation algorithm of the value of  $K$  is proposed when considering both noise and data loss.

The coordinates representing the skiing actions can be viewed as a random process. The coordinates of all feature points in the  $i$  frame image are represented as a column vector  $\hat{W}_i = [ \bar{u}_{i1}, \dots, \bar{u}_{iP}, \bar{v}_{i1}, \dots, \bar{v}_{iP} ]^T$ . Then:

$$\hat{W}_i^T = [ \omega_{i1} \bar{R}_i^1, \dots, \omega_{iK} \bar{R}_i^1, \omega_{i1} \bar{R}_i^2, \dots, \omega_{iK} \bar{R}_i^2 ] \begin{bmatrix} S_1 \\ \vdots \\ 0 \\ S_K \\ S_1 \\ 0 \\ \vdots \\ S_K \end{bmatrix} + \xi^T \tag{15}$$

The above formula (15) can be abbreviated as:

$$\hat{W}_i = [M'_{1 \times 6K} B'_{6K \times 2P}]^T + \xi = B'^T M'^T + \xi \tag{16}$$

where  $\bar{R}_i^1$  and  $\bar{R}_i^2$  are the first and second line respectively. The noise  $\xi$  of the feature points can be regarded as a random process with the mean of zero. It is obvious that the shape basis  $B'$  of the skiing image sequences is unchanged. The correlation coefficient matrix  $\hat{W}$  can be calculated as follows:

$$R_{\hat{W}} = E[\hat{W}\hat{W}^T] = B'^T E[M^T M] B + C_\xi \tag{17}$$

$C_\xi$  is the noise covariance matrix, and how to calculate this matrix is discussed below. From the principle analysis, the noise covariance of the feature points is a function of tracking algorithm and its parameters and the luminance variation near the tracking skiing feature points. The errors of tracking point position introduced by the tracking algorithm are non-uniform and correlated, which depends on the structural features of the local image. For example, it has high reliability for tracking  $u_{ij}$  and  $v_{ij}$  of the corner point  $j$ ; For a point called  $j$  on a line, the tracking in its gradient direction has high reliability (normal flow), and the tracking reliability in its tangent direction is very low, which is the directional error. For  $x_{ij}$  of the given array image  $I_i$ :

$$x_{ij} = \begin{bmatrix} u_{ij} \\ v_{ij} \end{bmatrix}, i = 1, \dots, F, j = 1, \dots, P \tag{18}$$

The Hessian matrix describes the second derivative information of the local region of the image, reflecting the curvature and rate of change of the image at that point. In feature point detection, the determinant of the Hessian matrix (or its feature value) is often used to determine key points in the image, which usually correspond to edges, corners, or other significant features in the image. Since these feature points are relatively stable in the image, an approximate noise covariance matrix can be obtained by using the inverse of Hessian matrix, which reflects the sensitivity of these feature points to noise. The noise covariance of the  $j$  feature point in the  $i$  skiing image can be estimated by the inverse of the Hessian matrix, that is:

$$Q_{ij} = E[\Delta x_{ij} \Delta x_{ij}^T] = \begin{bmatrix} \frac{\partial^2 I(u_{ij}, v_{ij})}{\partial x^2} & \frac{\partial^2 I(u_{ij}, v_{ij})}{\partial x \partial y} \\ \frac{\partial^2 I(u_{ij}, v_{ij})}{\partial y \partial x} & \frac{\partial^2 I(u_{ij}, v_{ij})}{\partial y^2} \end{bmatrix}^{-1} = \begin{bmatrix} \sigma_{ij1}^2 & \sigma'_{ij} \\ \sigma'_{ij} & \sigma_{ij2}^2 \end{bmatrix} \tag{19}$$

The elements of Hessian matrix are the second derivative and partial differential of image brightness in the direction of the x axis and y axis. Formula (19) is approximately equal to the actual noise covariance after multiplying by a scaling factor, which does not have any influence on the application [23].

Therefore, the noise covariance matrix can be obtained by the following formula:

$$C_\xi = \frac{1}{F} \sum_{i=1}^F \begin{bmatrix} \sigma_{1i1}^2 & 0 & \sigma'_{1i} & 0 \\ & \ddots & & \ddots \\ 0 & \sigma_{Pi1}^2 & 0 & \sigma'_{Pi} \\ \sigma'_{1i} & 0 & \sigma_{1i2}^2 & 0 \\ & \ddots & & \ddots \\ 0 & \sigma'_{Pi} & 0 & \sigma_{Pi2}^2 \end{bmatrix}_{2P \times 2P} = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} \tag{20}$$

The inverse of the noise covariance matrix  $C_\xi^{-1}$  is calculated as follows:

$$C_\xi^{-1} = \begin{bmatrix} (c_1 - c_2 c_4^{-1} c_3)^{-1} & -c_1^{-1} c_2 (c_4 - c_3 c_1^{-1} c_2)^{-1} \\ -c_4^{-1} c_3 (c_1 - c_2 c_4^{-1} c_3)^{-1} & (c_4 - c_3 c_1^{-1} c_2)^{-1} \end{bmatrix} \tag{21}$$

Combined with the above formulas, the following formula can be got:

$$R_{\hat{W}} C_\xi^{-1} = B'^T E[M^T M] B C_\xi^{-1} + I_{2P \times 2P} \tag{22}$$

In the above formula (22):

$$R_{\hat{W}} = \frac{1}{F} \sum_{i=1}^F \hat{W}_i \hat{W}_i^T \tag{23}$$

In general, for the 3D features of skiing actions, the maximum of rank of  $B'^T E[M^T M] B$  is  $6K$ . The  $i$  eigenvalue of the matrix  $H$  is represented by using  $\mu_i(H)$ , then:

$$\mu_i(H) = \mu_i(B'^T E[M^T M] B C_\xi^{-1}) + 1, \quad i = 1, \dots, 6K \\ \mu_i(H) = 1, \quad i = 6K + 1, \dots, 2P \tag{24}$$

Therefore, the matrix  $H$  has  $6K$  eigenvalues greater than 1. The estimation of  $K$  is got when just counting the number of eigenvalues of the matrix greater than 1 and divide by 6.  $K$  indicates that the dimension of shape space to express the sequences of deformation feature points and the number of shape bases to instruct the sequence of simulated feature points. Therefore,  $K$  can be used to estimate the deformation degree of skiing feature shape sequences. The degree of deformation of skiing features after translation and rotation can be defined as follows:

$$Deformation Degree = \frac{Count(eigenvalues\ of\ H > 1)}{6} \tag{25}$$

### 3 Experiment and Analysis

In order to test the effect of this proofreading algorithm for action regulation in the remote skiing teaching, it is applied to the remote skiing teaching platform. After the remote



teaching platform is built, the function test is carried out to judge whether it meets the standard.

### 3.1 Basic Function Test

In order to ensure the accuracy of the functional test results of this algorithm in the experiment, it is necessary to set the basic functional parameters of the remote teaching platform embedded in it. Reasonable parameter setting is the basis to ensure the normal operation of the remote teaching platform, and is also an important step to verify whether the algorithm in this paper can be effectively implemented. In this experiment, the basic functions of the platform hardware and software are tested. In the hardware environment, the server project is published on the local Apache server, and the Postman software is used to send simulated network requests to the server for testing [24], so as to ensure that the server can correctly analyze the foreground network requests, execute the corresponding business logic processing and return the processing results. Then the database connectivity test is carried out, which includes whether the data can be stored, updated, deleted and obtained correctly and normally. In the software environment, the database and operating system of the platform are tested. The editing software is designed as Eclipse and ADT, the operating system is Window 10, and the database is MySQL 5.6. Set the actual parameters in the above software, select Apache Tomcat 9.0 server software, its port number is 8080, the database server port number is 3306, the maximum number of connections in the connection pool is 10, the initial number of connections in the connection pool is 5, and the screen resolution of the Android simulator is  $1440 \times 2960$ . The memory size is 4 GB, preliminary setting of simulation parameters and adding qumu options; With the help of existing tools, real-time acquisition of relevant data information; The above related test results are recorded and stored in a unified manner.

On this basis, the following experimental results are obtained: When entering the platform, alpine skiing skills can be learned directly through multiple interfaces. The login interface of the student side of this remote skiing teaching platform is shown in Fig. 1.

As shown in Fig. 1, the remote skiing teaching platform has the function of self-help action proofreading, which is mainly based on the embedding of the algorithm in this paper as the core technology.

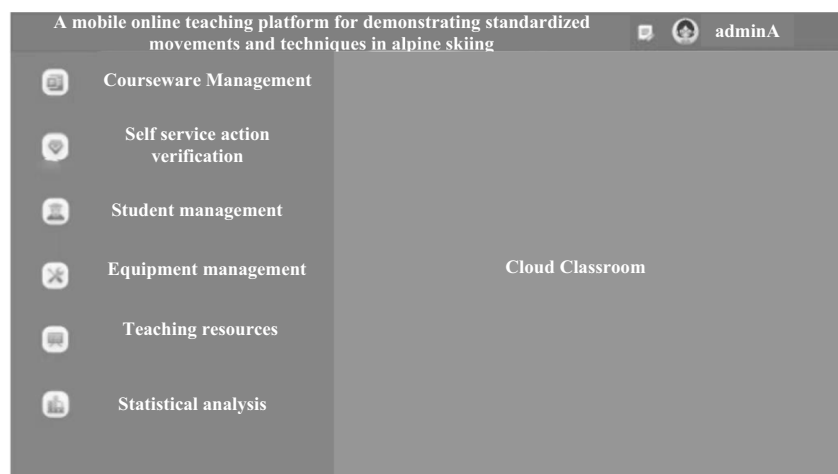
To train the model used in this algorithm presented in this paper, a video dataset containing a variety of skiing actions is constructed. This dataset contains skiing videos of different skiers, different skiing techniques, and different environmental conditions. Improper starting skills and improper center control are common problems for skiing beginners, and choosing these actions can represent the types of mistakes that need to be focused on in skiing teaching. The standard action represents the correct technical standard and is the reference for proofreading and correcting the wrong action. By labeling these specific types of movements, targeted feedback and guidance can be provided for ski instruction. This kind of comparative learning helps to improve the accuracy of the model when proofreading movements. Each video has been professionally labeled, which marks whether the skier's actions are regulated and the specific reasons why they are not regulated.

This data set provides rich samples for the training the model, which ensures that the model can learn enough features to represent and recognize the skiing actions.

### 3.2 Test and Analysis of Intelligent Proofreading Effect on Remote Skiing Teaching Actions

Figure 2 is the skiing action video images uploaded after the self-service action proofreading function is selected. As shown in Fig. 3(a), the skiing action video images uploaded by the student user have poor image quality and insufficient

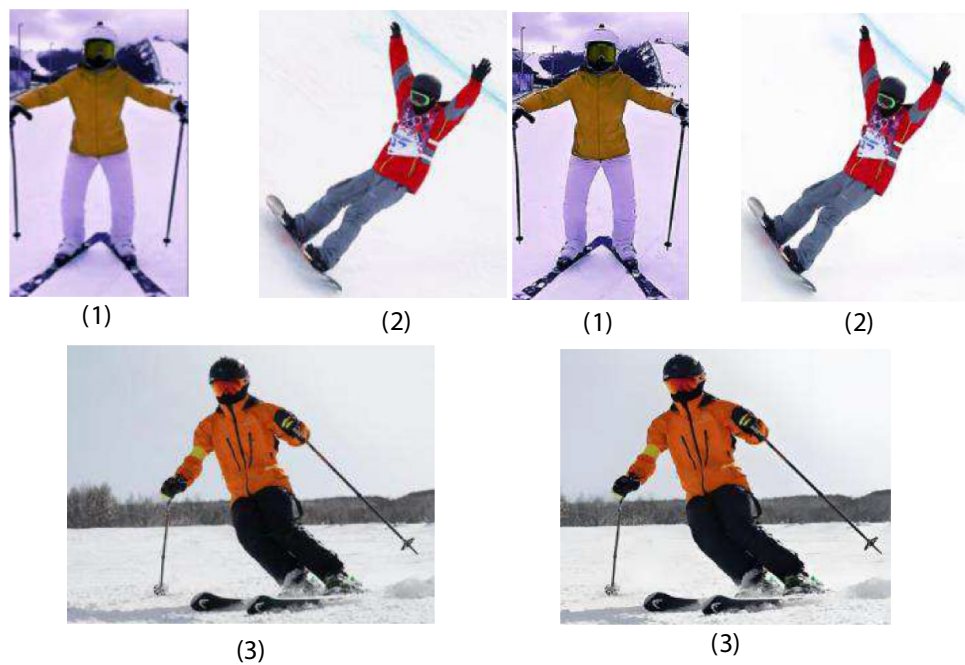
Fig. 1 Platform Interface



**Fig. 2** Video annotation example



**Fig. 3** Enhancement effect of skiing action video images



(a) Before enhancement (b) After enhancement

clarity due to poor weather conditions. In this paper, the video images of skiing actions enhancement algorithm based on improved Retinex is used to enhance the image processing, which significantly improves the visual effect of the image, as shown in Fig. 3(b). Such images are not only easy to watch and analyze, but also help the platform more accurately assess whether the student's skiing actions are regulated.

In the above Fig. 2, the enhanced (Fig. 1) is selected to detect the 3D key point features on the skeleton of the skiing action video images, and the experimental results are shown in Fig. 4.

The detection results of key points on the skeleton in Fig. 4 are taken as the input sample of the intelligent image proofreading algorithm for action regulation based on the adaptive multi-scale graph convolutional network. With the

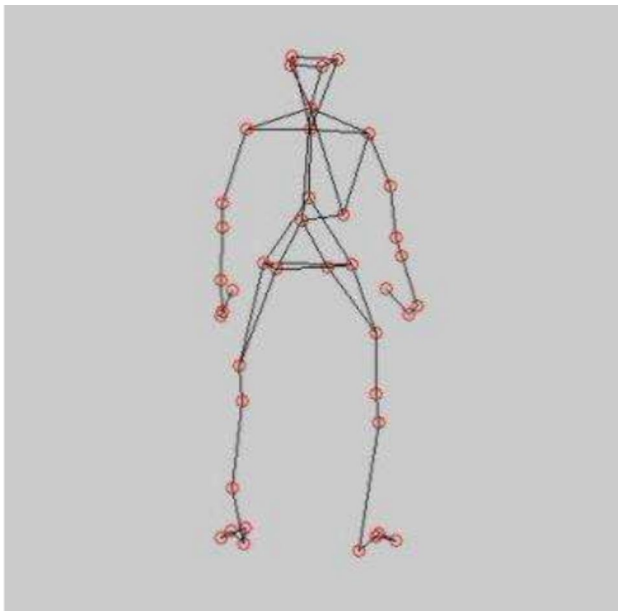


Fig. 4 Key point detection results on the skeleton of skiing actions

assistance of the visualization function of the remote skiing teaching platform, the action-normative intelligent proofreading results are displayed as shown in Fig. 5.

As shown in Fig. 5, in the skiing action video image uploaded by the student user, it is considered that there are errors in the action after proofreading by the algorithm in this paper, and the cause of the error is shown as keeping knee valgus. Keeping knees valgus when skiing is the wrong action. When skiing, learners should maintain a correct posture and avoid knees valgus. Proper skiing posture includes balancing the body, legs slightly bent, knees in line with

Table 1 The regulated proofreading effect of different algorithms on various skiing actions

Action Type	Actual data of skiing action images	
	Number of correct actions/piece	Number of incorrect actions/piece
Slide down the snow slope	50	10
Cross-country skiing	100	20
Ski-Jump	150	30
Freestyle	250	40

toes, weight in front of the body, etc. This is mainly because the improved Retinex algorithm introduced in the method in this paper can enhance the detail of the image, making small changes in the action and errors more obvious and thus easier to detect. At the same time, through the singular value decomposition and the shape changing basic idea, the three-dimensional motion features of the skiing movement can be accurately extracted, so the research method in this paper can judge whether the movement is correct.

In order to further test the proofreading function for action regulation of this algorithm presented in this paper for various skiing actions, 600 skiing video images are randomly selected, of which 500 are regulated images and 100 are non-regulated images. The human action recognition algorithm based on initial attention in reference [6] and the human action classification algorithm based on deep Gabor network in reference [7] are selected as the comparison algorithm respectively. Sliding down the snow slope, cross-country skiing, ski-jump and freestyle as examples, the proofreading error results of different algorithms are tested

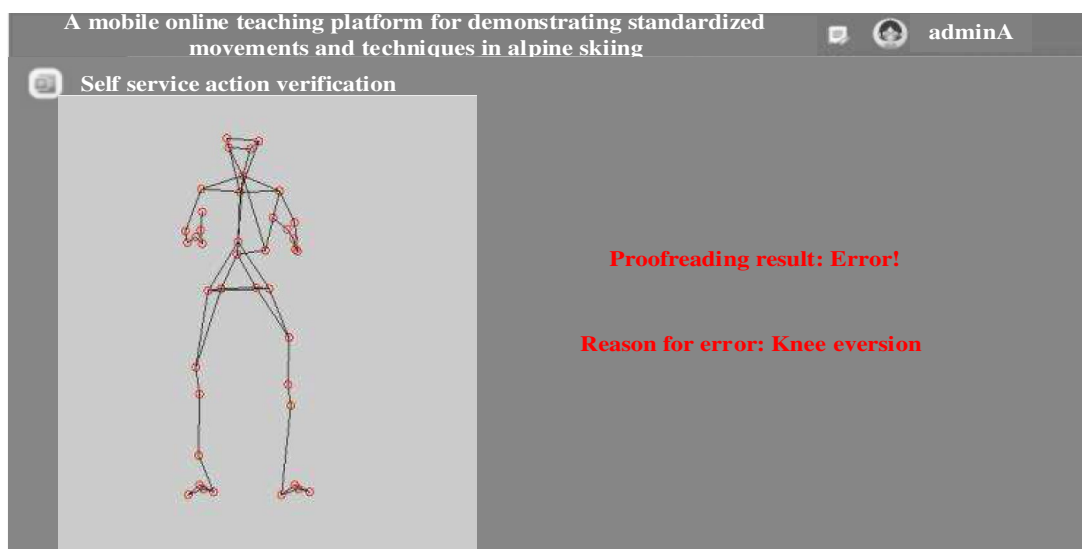


Fig. 5 Intelligent proofreading results for action regulation

**Table 2** The regulated proofreading effect of different algorithms on various skiing actions

Action Type	Method in this article		
	Number of correct actions/piece	Number of incorrect actions/piece	Number of method proofreading errors / piece
Slide down the snow slope	50	10	1
Cross-country skiing	100	20	2
Ski-Jump	150	30	1
Freestyle	250	40	3

**Table 3** The regulated proofreading effect of different algorithms on various skiing actions

Action Type	Reference [6] method		
	Number of correct actions/piece	Number of incorrect actions/piece	Number of method proofreading errors / piece
Slide down the snow slope	48	12	2
Cross-country skiing	99	21	3
Ski-Jump	149	31	3
Freestyle	201	39	5

**Table 4** The regulated proofreading effect of different algorithms on various skiing actions

Action Type	Reference [7] method		
	Number of correct actions/piece	Number of incorrect actions/piece	Number of method proofreading errors / piece
Slide down the snow slope	40	20	3
Cross-country skiing	105	15	4
Ski-Jump	145	35	7
Freestyle	198	42	6

under the prerequisite of different skiing actions. The results of proofreading are shown in Tables 1, 2, 3 4.

According to the experimental results, compared with the human action recognition algorithm based on initial attention in reference [6] and the human action classification algorithm based on deep Gabor network in reference [7], the proposed proofreading algorithm shows lower errors for the action regulation in four types of skiing actions, and also has higher accuracy in the recognition of different actions in the set of skiing actions. Because the algorithm proposed in this paper uses the shape-changing basis idea, it can get the correct three-dimensional motion feature shape basis structure in the ski image. By randomly initializing the parameters and using the least square method for optimization, the algorithm can gradually adjust the parameters to reach the optimal solution. This iterative process helps to reduce errors and improve the accuracy of action proofreading and recognition. This advantage is especially reflected in the complex technical actions such as ski-jump and freestyle skiing, which verifies the effectiveness and accuracy of

this algorithm in dealing with complex skiing actions. This algorithm can capture the features of skiing actions more accurately, provide more accurate proofreading results and provide a new solution for the field of remote skiing teaching, which is helpful to improve the teaching quality and user experience.

## 4 Conclusion

The intelligent proofreading algorithm for remote skiing teaching based on variable shape basis is proposed in this paper, which has brought a revolutionary change to the field of skiing teaching with its unique technical advantages. The algorithm makes full use of the latest progress of artificial intelligence and image processing technology, which can automatically identify and analyze the skier's action performance in the images, and realize the efficient and accurate proofreading accuracy of the skiing action regulation.

Compared with traditional teaching algorithms, this algorithm has obvious advantages. First of all, it breaks through the restrictions of time and space, making skiing teaching no longer subject to the constraints of time and space, and providing more people with the opportunity to contact and learn skiing. Secondly, through intelligent image technology, the algorithm can detect whether the skier's actions are regulated in real time and automatically, and give immediate feedback and guidance, which greatly improves the efficiency and accuracy of teaching. In addition, the algorithm is highly intelligent and personalized, which could carry out personalized teaching guidance according to different learners' level and needs to help learners master skiing skills faster and improve learning effects. At the same time, this algorithm also could carry out intelligent analysis and evaluation according to the learner's action performance, so as to provide powerful data support for teaching and promote the continuous improvement of teaching quality.

To sum up, the intelligent image proofreading method proposed in this paper, with its advantages of high efficiency, accuracy, intelligence and individuation, has injected new vitality into the development of remote skiing teaching. However, the coordinate transformation and measurement matrix calculation in the research depend on the camera Angle. If the position or Angle of the camera changes, the elimination effect of the translation vector may be affected, which in turn affects the accuracy of the measurement matrix. Therefore, in view of the impact of camera Angle changes on coordinate transformation and measurement matrix calculation, the following research will introduce dynamic Angle tracking technology to track camera motion in real time, and adjust coordinate transformation and measurement matrix calculation accordingly. This may require a combination of motion estimation and tracking algorithms. Through these targeted studies, the adaptability of the algorithm to camera Angle changes can be improved, so as to improve the accuracy and robustness of coordinate transformation and measurement matrix calculation.

**Acknowledgements** The paper was supported by key entrusted project of educational reform in Heilongjiang Province with No. SJGZ20220168. Authors acknowledge contribution to this project from the Rector of the Silesian University of Technology, Poland under a proquality grant no. 09/010/RGJ24/0031.

**Author Contribution** Tie Li wrote the main manuscript text, Jun Wang worked about algorithm, Katarzyna Wiltos provided helpful advice, Marcin Wozniak revised and reviewed the manuscript.

**Data Availability** No datasets were generated or analysed during the current study.

## Declarations

**Competing Interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Sletten HS, Eikevag SW, Silseth H, Grondahl H, Steinert M (2021) Force Orientation Measurement: Evaluating Ski Sport Dynamics. *IEEE Sens J* 21(24):28050–28056
2. Sollie O, Holmsen K, Steinbo C, Ommundsen Y, Losnegard T (2021) Observational vs coaching feedback on non-dominant whole-body motor skill performance - application to technique training. *Scand J Med Sci Sports* 31(11):2103–2114
3. Bekasiewicz A, Pankiewicz B, Wojcikowski M, Klosowski M, Koziel S (2021) Application of Open-Hardware-Based Solutions for Rapid Transition From Stationary to the Remote Teaching Model During Pandemic. *IEEE Trans Educ* 64(3):299–307
4. Luvizon DC, Habermann M, Golyanik V, Kortylewski A, Kortylewski C (2023) Scene-Aware 3D Multi-Human Motion Capture from a Single Camera. *Computer Graphics Forum* 42(2):371–383
5. Shujah IM, Khush B, Mansoor I, Rashid K, Zhong F, Ye M (2023) Representation for action recognition with motion vector termed as: SDQIO. *Expert Syst Appl* 212:118406
6. Taima R, Mim MA, Sadia A, Mohammad A, Yousuf SU, Salem AA, Khondokar FH, Mohammad AM (2023) GRU-INC: An inception-attention based approach using GRU for human action recognition. *Expert Syst Appl* 216:119419
7. Hoang TL, Son LP, Abdesselam B (2021) A Fast and Compact Deep Gabor Network for Micro-Doppler Signal Processing and Human Motion Classification. *IEEE Sens J* 21(20):23085–23097
8. Anindita S, Sajjan R, Jayita S, Chandreyee C (2024) A Survey of Machine Learning and Meta-heuristics Approaches for Sensor-based Human Action Recognition Systems. *J Ambient Intell Humaniz Comput* 15(1):29–56
9. Palak G, Prashant J, Deepali V (2022) An Improved Empirical Hyper-Parameter Tuned Supervised Model for Human Action Recognition based on Motion Flow and Deep Learning. *Int J Performability Eng* 18(11):808–816
10. Fu W, Luo Z, Liu S, Lloret J, de Albuquerque VHC, Saudagar AKJ, Muhammad K (2024) Spatiotemporal correlation based self-adaptive pose estimation in complex scenes. *Digital Communications and Networks*. <https://doi.org/10.1016/j.dcan.2024.03.007>
11. Jain R, Karsh R, Kumar B, Abul A (2022) Encoded motion image-based dynamic hand gesture recognition. *Vis Comput* 38(6):1957–1974
12. Sunar AS, Leon-Urrutia M, White S, Uhomoihi J (2021) Guest editorial: Distance learning, MOOCs and globalisation of engineering education. *Comput Appl Eng Educ* 29(4):645–647
13. Gabriele S, Michela L, Gabriele G, Alessandro R (2022) Survey of algorithms and evaluation of Retinex-inspired image enhancers. *J Electron Imaging* 31(6):63055

14. Ziaur R, Zafar A, Inayat K, Muhammad I, Uddin YG, Zhihua H (2022) Diverse image enhancer for complex underexposed image. *J Electron Imaging* 31(4):41213
15. Men Q, Ho ESL, Shum HPH, Leung H (2023) Focalized contrastive view-invariant learning for self-supervised skeleton-based action recognition. *Neurocomputing* 537:198–209
16. Rajappan RJ, Kandaswamy TK (2022) A composite framework of deep multiple view human joints feature extraction and selection strategy with hybrid adaptive sunflower optimization-whale optimization algorithm for human action recognition in video sequences. *Comput Intell* 38(2):366–396
17. Shi M, Aberman K, Aristidou A (2021) MotioNet:3D Human Motion Reconstruction from Monocular Video with Skeleton Consistency. *ACM Trans Graph* 40(1):1
18. Liu S, Luo Z, Li Y et al (2024) Solution of Wide and Micro Background Bias in Contrast Action Representation Learning. *Eng Appl Artif Intell* 133:108244
19. Liu S, Li Y, Fu W (2022) Human-centered attention-aware networks for action recognition. *Int J Intell Syst* 37(12):10968–10987
20. Cheng PYZ, Liu P (2023) Research on Moving Image Feature Recognition Based on Dynamic Contour Curve. *Comput Simul* 40(07):163–167
21. Wang Y, Xia Y, Liu S (2024) BCCLR: A Skeleton-Based Action Recognition with Graph Convolutional Network Combining Behavior Dependence and Context Clues. *Comput, Mater Continua* 78(3):4489–4507
22. Li Y, Qi X, Saudagar AKJ et al (2023) Student behavior recognition for interaction detection in the classroom environment. *Image Vis Comput* 136:104726
23. Choi Y, Lee Y, Kim B (2022) Group-based bi-directional recurrent wavelet neural network for efficient video super-resolution (VSR). *Pattern Recogn Lett* 164:246–253
24. Nuhoho R, Elimeli CW, Baffour AA (2022) A semi-supervised network based on feature embeddings for image classification. *Expert Syst* 39(4):12908

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.