Article - Engineering, Technology and Techniques

Estimation of Human Motion Posture Using Multi-labeling Transfer Learning

Yang Wang[1]
https://orcid.org/0000-0002-4301-5910

Jie Ren[2]
https://orcid.org/0000-0002-4375-7068

Shangbin Li[1]*
https://orcid.org/0000-0002-8556-3632

Zhijun Hu[3]
https://orcid.org/0000-0002-9965-4594

Raja Soosaimarian Peter Raj[4]
https://orcid.org/0000-0002-7216-2207

[1]Harbin Engineering University, Physical Education Department, Harbin, China; [2]Harbin Sport University, College of Physical Education and Training, Harbin, China; [3]Guangxi Normal University, College of Mathematics and Statistics, Guilin, China;[4]Vellore Institute of Technology, School of Computer Science and Engineering, Vellore, India.

*Correspondence: lishangbin@hrbeu.edu.cn (S.L.).

HIGHLIGHTS

- This paper addresses the problem that it is difficult to mine and extract the potential features of human motion posture feature vector.

- The posture compensation function is used to compensate the error part to ensure the accuracy of posture estimation results

- The posture compensation function is used to correct and compensate the estimation results and improve the reliability of the estimation results.

Abstract: Human posture estimation is the basis of many computer vision tasks, such as motion recognition, violence detection and behavior understanding. Therefore, it is of great significance to study the estimation algorithm of human motion posture (HMP). To address the problem of poor estimation effect of traditional HMP estimation algorithm, in this paper, an estimation algorithm for HMP using multi-labeling transfer learning is proposed. First, the original human motion image is labeled by using the multi-label transfer learning, the HMP features are extracted, and the original image classification is completed. Second, a regulator is constructed based on the classification results of the original image, and the regulator is used to adjust the estimation results of HMP based on a convolutional neural networks. Finally, the posture compensation function is used to compensate for the error part to realize the estimation of HMP. In the experiment, the Human3.6m data set and MPII data set were used as the basis for testing. The results show

that the proposed algorithm has high correct recognition rate of HMP. The similarity between the posture estimation results, and the target image is 92%-97%. The accuracy of posture estimation is 98.1%. The proposed algorithm can be widely used in many fields, such as human-computer interaction, recognition authentication and intelligent monitoring.

Keywords: Human motion posture (HMP); Posture estimation; Multi-labeling transfer learning; Image label.

INTRODUCTION

Posture estimation refers to the estimation of posture parameters of various parts of the human from the input image sequence, such as the position of various parts of the body in three-dimensional (3D) space or the included angle between various joints of the body [1]. Through these pose parameters, the motion of human can be reconstructed in 3D space. Posture estimation is one of the challenging aspects of human behavior analysis, and its main task is to enable computers to automatically perceive people in a scene and determine what they are "doing". At present, more and more digital products are integrated into people's daily life, constantly producing a variety of pictures and video data every day. Society is a collection of people, in which humans will inevitably want to extract the main content from these picture and video data, and then hope to effectively understand and process human activities in these data. Therefore, it needs to be processed with the help of current tools. Therefore, human motion posture (HMP) estimation has become an important part of computer vision, and this function has practical application value in real life, specifically in human-computer interaction, film capture and animation production, automatic driving, track tracking, video indexing and retrieval, body identification and intelligent monitoring and other fields. Therefore, HMP estimation is an important task in computer vision, and is also an indispensable part of computer understanding human motion and behavior. Posture estimation is an important issue in the field of computer vision [2-3]. According to the analysis of existing relevant research, due to the complexity and diversity of HMP, the traditional posture estimation algorithms are prone to generate calculation errors, and it is difficult to mine some potential HMP features, which brings certain difficulties to accurately realize the posture estimation. Therefore, it is very necessary to take a more reasonable and accurate way to estimate the HMP based on advanced computer tools [4]. Among existing researches on accurate estimation algorithms of HMP, Leibovich M and coauthors [5] proposed a tensor method and a human motion estimation algorithm for aperture imaging. The reflection of the moving target is separated from the reflection from the stationary background, and the moving and stationary reflections are imaged respectively. To this end, the data is expressed as a third-order tensor, which is formed by partially overlapping sub-aperture data. The tensor robust principal component analysis is applied to tensor data, and human motion estimation is realized according to the results of principal component analysis. However, the algorithm has the problem of low similarity between postures. Layton O [6] proposed a human motion post-estimation algorithm based on convolutional neural networks (CNN) and optical flow template. A biologically inspired neural network is built, so the patterns in optical flow are learned, and the self-motion of the observer can be encoded. The network combines the unsupervised learning algorithm of fuzzy art with the hierarchical structure based on primate visual system. This design provides fast local feature learning across parallel modules in each network layer, and the trained neural network is used to estimate HMP.

The above algorithms have achieved the estimation and calculation process of HMP. However, the estimation and calculation result error is relatively large, which seriously affects the use effect of the above methods. Therefore, this paper proposes an estimation algorithm of HMP based on multi-marker transfer learning. The main contributions of this paper are as follows: (1) The multi-labeling transfer learning is applied to the process of HMP features, through the labeling and feature transfer of the image. This addresses the problem that it is difficult to mine and extract the potential features of HMP feature vector.(2)Because there are still some errors in the preliminary estimation result of HMP of CNN, the posture compensation function is used to compensate the error part to ensure the accuracy of the posture estimation result.(3)After the HMP estimation is completed, the filter built in this paper is used to reduce the error of the estimation result, and the posture compensation function is used to correct and compensate the estimation result, to improve the reliability of the estimation result.

Related works

Wu R and coauthors [7] proposed a motion estimation algorithm for the whole video using improved K-means clustering and super-pixel technology. A simple linear iterative clustering pre-segmentation method is used to obtain the super-pixels of the video frame, and clustering is performed according to the motion vector

of the super pixel centroid to eliminate the large value clustering center. The feature points of the remaining super-pixels are matched between two adjacent frames, the motion vector space of the feature points is established, and the improved K-means clustering is used for clustering. Finally, the most abundant clusters are reserved, and the global motion is obtained through homography transformation, to achieve the goal of motion estimation. However, the algorithm has the problem of a low correct recognition rate, and the practical application effect is not good. Li X and coauthors [8] proposed a post-estimation algorithm of human motion based on Backpropagation (BP) neural network. BP neural network is used to estimate the torque of human elbow joints. Its physiological and physical input elements include shoulder posture, elbow-related muscle activation, elbow position and angular velocity. By controlling the elbow exoskeleton, the joint torque of the human is estimated, to determine the HMP. However, the algorithm has a high over-recognition rate and poor application effect for practical problems. Li J and coauthors [9] designed a human posture estimation algorithm based on radar and cross source point cloud fusion technology. The unified simplified expression of geometric elements in conformal geometric algebra is used, the traditional point-to-point correspondence is broken, and the matching relationship between points and balls is constructed. For the fused point cloud, the plane clustering method based on CGA is used to eliminate the point cloud diffusion, and then the 3D contour model is reconstructed. With the twister and Clohessy Wiltshire equations used, the posture and other motion parameters of the non-cooperative target are obtained through the traceless Kalman filter, to estimate the HMP. However, the algorithm has the problem of poor effect of posture image annotation, which is difficult to achieve the desired application effect. Lauer J and coauthors [10] has established a learning architecture based on DeepLabCut, which provides the motion posture tracking function required by different scenes. However, the diversity of HMP is not considered, resulting in insufficient feature analysis; Wang C and coauthors [11] analyzed the low-resolution image data of human posture estimation, and proposed a new confidence perception learning method, which captures the statistical importance of model output in a small batch of learning. It is an efficient low-resolution HMP estimation method. However, the algorithm runs for a long time and is not efficient.

Therefore, this paper proposes an estimation algorithm of HMP based on multi-labeling transfer learning, which extracts feature information by labels HMP, and classifies HMP images; The CNN is used for preliminary estimation of HMP, and a regulator is constructed to adjust the estimation results to compensate for errors, to achieve estimation of HMP. The results show that the proposed algorithm has a good overall attitude estimation effect, the similarity between the attitude estimation results and the target image is between 92% and 97%. The estimation accuracy is 98.1%, and the proposed algorithm has high efficiency and high correct recognition rate.

METHODOLOGY

Calculation process of estimation of HMP

To address the problem of poor estimation effect of traditional HMP estimation algorithms, an HMP estimation algorithm based on multi-labeling transfer learning is proposed. The proposed algorithm framework is shown in Figure 1.
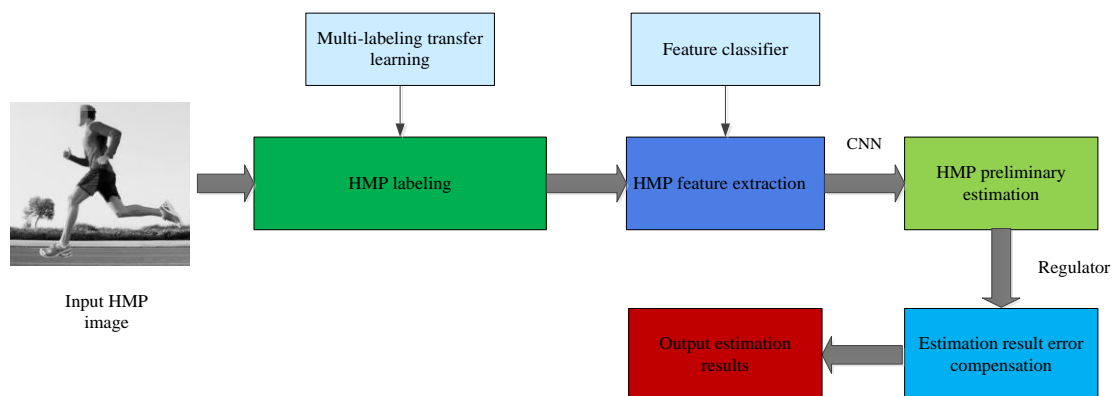

Figure 1. Framework of the proposed algorithm

It can be seen from Figure 1 that we first obtain the HMP labeling results based on multi-labeling transfer learning, and complete the HMP feature extraction by original image identification, feature vector acquisition and feature classifier construction. Then we construct a CNN to initially estimate the HMP, and use a regulator

to adjust the HMP estimation results based on the CNN, and the posture compensation function is used to compensate for the error part to output the final HMP estimation result.

HMP labeling and feature extraction

In this study, multi-labeling transfer learning is the core technology. Through several case studies, the original image is labeled using multi-labeling transfer learning, and the HMP features are extracted [12-13]. The multi-labeling transfer learning framework using feature transfer is shown in Figure 2.
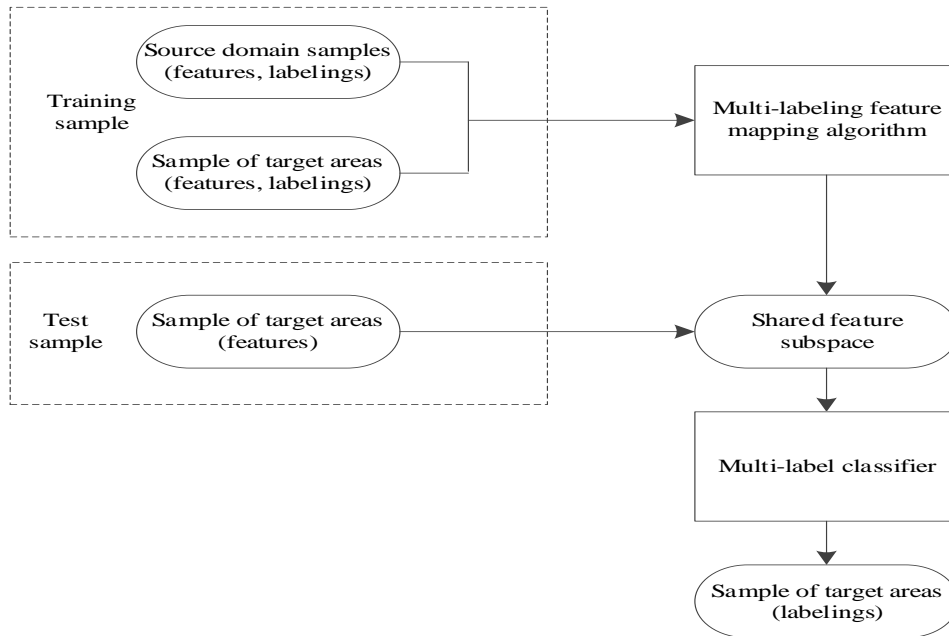
Figure 2. Multi-labeling transfer learning architecture based on feature transfer

According to Figure2. In the multi-labeling transfer learning architecture based on feature transfer, the multi-labeling feature mapping algorithm is used to process the training samples and mark the sample features of the source domain and the target domain, respectively. On this basis, the sample features of the target domain of the test samples are marked. Combined with the processing results of training samples and test samples, a shared feature subspace is constructed to construct a multi-labeling classifier to obtain the sample labels of the target domain [14], to realize the HMP labeling. Combined with the result of feature labeling, the feature extraction of HMP is completed.

HMP labeling

Supposed that the training sample of HMP is represented by $X = \left[X^{(s)}, X^{(t)} \right]$, which, $X^{(s)}$ represents a training sample. $X^{(t)}$ is the test sample, X is regarded as a dictionary, and X is tri-factorized by a non-negative matrix. Therefore, the objective function of HMP labeling based on multi-labeling transfer learning is as follows:

$$f = \left\| X - XWU \right\|_F^2 + \lambda \bar{D}\left(X^{(s)}, X^{(t)} \right) + \gamma_1 \left\| W \right\|_F^2 + \gamma_2 \left\| U \right\|_F^2 \tag{1}$$

where $U \in R^{k \times n}$ represents the low dimension of HMP features sharing the feature subspace. $W \in R^{n \times k}$ represents the relationship matrix between X and U. $\left\| W \right\|_F^2, \left\| U \right\|_F^2$ represents the complexity control function of W and U. F represents complexity. γ_1 and γ_2 represent different matrix factorization coefficients, respectively. $\bar{D}\left(X^{(s)}, X^{(t)} \right)$ is the fitting regular term of the source domain sample and the target domain sample. λ is a non-negative constant.

To further improve the efficiency of HMP labeling, $\bar{D}\left(X^{(s)}, X^{(t)}\right)$ is converted into the following Equation

$$\bar{D}\left(X^{(s)}, X^{(t)}\right) = \left\| \frac{1}{n}\sum_{i=1}^{n}\varphi(x_i) - \frac{1}{m}\sum_{j=1}^{m}\varphi(x_j) \right\|^2 +$$

$$\sum_{c=1}^{C}\theta_c \left\| \frac{1}{n}\sum_{i=1}^{n}\varphi\left(x_i^{(c)}\right) - \frac{1}{m}\sum_{j=1}^{m}\varphi\left(x_j^{(c)}\right) \right\|^2$$

(2)

where $\left(x_i, x_j\right)$ and $\left(x_i^{(c)}, x_j^{(c)}\right)$ represent source domain samples and target domain samples, respectively. θ is the weight coefficient of multi-labeling feature. φ is the degree of adaptation. m, n is the data amounts of test samples and training samples, respectively. M represents an adaptation matrix. The matrix is expressed by the following Equation

$$M = M_0 + \theta M_c$$

(3)

where M_0 and M_c represent the initial matrix and the multi-labeling feature mapping matrix, respectively, and their calculation Equations are as follows:

$$\left(M_0\right)_{ij} = \begin{cases} \dfrac{1}{n^2}, x_i, x_j \in X^{(s)} \\ \dfrac{1}{m^2}, x_i, x_j \in X^{(t)} \\ -\dfrac{1}{mn}, \ \ others \end{cases}$$

(4)

$$\left(M_c\right)_{ij} = \begin{cases} \dfrac{1}{n^{(c)2}}, x_i, x_j \in X_c^{(s)} \\ \dfrac{1}{m^{(c)2}}, x_i, x_j \in X_c^{(t)} \\ -\dfrac{1}{n^{(c)}m^{(c)}}, \begin{cases} x_i \in X_c^{(s)}, x_j \in X_c^{(t)} \\ x_i \in X_c^{(t)}, x_j \in X_c^{(s)} \end{cases} \\ 0, others \end{cases}$$

(5)

According to the above analysis, the HMP labeling objective function based on multi- labeling transfer learning can be rewritten into the following form:

$$f = \left\| X - XWU \right\|_F^2 + \lambda tr\left(UMU^T\right) + \gamma_1 \left\| W \right\|_F^2 + \gamma_2 \left\| U \right\|_F^2$$

(6)

To clarify the relationship between features in the feature space of HMP, and optimize the quality and efficiency of feature labeling, hypergraph regularization terms are introduced into the above objective function to clarify the relationship between HMP features, avoid the destruction of geometric structure information in feature space, and avoid the loss of HMP after labeling. Hypergraph regularization terms can be expressed by the following Equation

$$\bar{H}\left[X^{(s)}, X^{(t)}\right] = \sum_{i,j=1}^{m+n}\left(\phi(u_i) - \phi(u_j)\right)^2 W_{E_{ij}} =$$

$$\sum_{i,j=1}^{m+n}\phi(u_i)W_{E_{ij}}\phi(u_j)$$

(7)

where ϕ represents a manifold learning parameter. u_i and u_j represent different hypergraph regular term parameters. $W_{E_{ij}}$ is hypergraph parameter.

If the regular term of hypergraph is introduced into the objective function, the improved objective function of human posture labeling can be expressed by the following Equation

$$f = \left\| X - XWU \right\|_F^2 + \lambda tr\left(UMU^T\right) + \kappa tr\bar{H}\left[X^{(s)}, X^{(t)}\right]$$
$$+ \gamma_1 \left\| W \right\|_F^2 + \gamma_2 \left\| U \right\|_F^2 \tag{8}$$

where κ denotes the control coefficient of the regular term of hypergraph.

Feature extraction of HMP

According to the results of HMP labeling, a multi-labeling classifier of HMP features is constructed, and the classifier is used to extract HMP features.

First, the collected original image is constructed into the form of training sample database. The basic labeling is used to obtain the labeling matrix of the image and set it as Q. According to the Laplace matrix J of this training sample library, determine the basic structure of the image set and map it to the eigenvector of the matrix Q

$$\min(\sqrt{\left| T^i S - Q \right|} + \sqrt{\left| L - Q_i \right|}) \tag{9}$$

where T is the multi-labeling feature mapping matrix of image [15]; L refers to the image labeling matrix after preprocessing. The calculation result of Equation (9) is combined with the elastic network algorithm to obtain the feature selection calculation Equation of the test sample set

$$\min(\sqrt{\left| T^i S - Q \right|} + \alpha \sum_{i=1} \left| L - Q_i \right|) \tag{10}$$

where α denotes the motion feature calculation parameter, and construct the shared feature subspace[16] according to α. To improve the use effect of the classifier designed in this paper, the support vector machine is integrated with some theories of multi-labeling transfer learning, and a multi-labeling classifier is constructed, which regards the labels of motion posture features [17] as $U = (u_1, u_2, ..., u_n)$. If there is a certain order relationship in this sequence, the feature relationship between the two images can be expressed as:

$$\begin{cases} m_i \rightarrow m_j \\ w(m_i) \rightarrow w(m_j) \end{cases} \tag{11}$$

Assuming that there is a linear relationship in Equation (11), then according to this relationship and the labeling result of the image, the image is paired to obtain the sample vector and the labeling value associated between the two images

$$\begin{cases} m_{(1)} \rightarrow m_{(2)} \\ r = \begin{cases} 1, \chi_{(1)} \rightarrow \chi_{(2)} \\ 0, \chi_{(2)} \rightarrow \chi_{(1)} \end{cases} \end{cases} \tag{12}$$

where $m_{(1)}$ is the motion posture feature of image 1; $m_{(2)}$ refers to the motion posture feature of image 2; $\chi_{(1)}$ stands for the labeling of the motion posture feature of image 1; $\chi_{(2)}$ is the labeling of the motion posture feature of image 2. According to this Equation, the images in the set can be marked and sorted[18]. To set image feature statistical rules, there are

$$K = \min m_{(1)} \rightarrow m_{(2)} + F \sum_{i=1}^{n} \eta_i$$

$$s.t.(m_{(1)} \rightarrow m_{(2)}) \geq 1$$

$$\eta_i \geq 1$$

(13)

Equation (13) is used as the multi-labeling classifier of human posture features in this study, and the multi-labeling classifier [19-20] is used to extract HMP features, which lays a solid foundation for the subsequent HMP estimation.

HMP estimation algorithm

Based on the above multi-labeling transfer learning to label the original human motion image, extract the HMP features and classification results, a regulator is constructed in combination with the classification results of the original image, and the regulator is used to adjust the HMP estimation results based on CNN, and the attitude compensation function is used to compensate the error part to achieve estimation of HMP. The process of HMP estimation algorithm is as follows:

Input: human motion image

Output: results of HMP estimation

CNN is constructed to improve the overall image feature space expression ability. At the same time, according to the pre-obtained image feature values [21-22], the image is re-allocated. CNN activation function is set according to the image classification results, and the specific contents are as follows:

$$\sigma = f \left\| (z_i, o_i) \right\| = |z_i| \oplus |o_i|$$

(14)

where o is the excitation function of neural network; z refers to the real number sequence in neural network. The image processing results and the characteristic values of HMP are input into the neural network, and the calculation Equation of two-dimensional (2D) HMP estimation is obtained through convolution calculation

$$\wp_{i,j} = f \left\| (X_{i,j}) \right\| * \overline{X_{i,j}}$$

(15)

where $X_{i,j}$ refers to the 2D image processing result; $\overline{X_{i,j}}$ is the 2D motion posture feature [23]. To ensure that this design method can be applied to the processing of 3D images or video images. A regulator is added to the calculation of Equation (15) to control the calculation accuracy of HMP estimation. The specific regulator calculation process is set as follows:

$$y(k) = \frac{R'c'(t)}{R'\sum c'(t)\theta}$$

(16)

where $c'(t)$ is the time error during the estimation and calculation process; $y(k)$ denotes the output result of regulator; R' refers to the integral adjustment process of estimation calculation, which is used to eliminate the errors of calculation results. After the above basic operation is completed, a low-pass filter and a high pass filter are added. At the same time of calculation, the filtering process is completed. The filter transfer function [24] is set as follows:

$$\Re = \frac{(A'+E')}{|A'|}$$

(17)

where A' refers to the low frequency image data in image; E' refers to high frequency data in image. Equation (17) is used to filter the noise generated in the calculation and control the accuracy of the estimation result to a certain extent. After the above operations are completed, the posture compensation [25] is performed on the obtained results. The calculation process is as follows:

$$\sum \Delta e' = \sum \aleph + \Delta \varpi$$

(18)

where $\sum \Delta e'$ refers to posture compensation; $\sum \aleph$ is the estimation result error vector; $\Delta \varpi$ is spatial error vector obtained from historical data. At the same time, the estimation results are processed to obtain the final output of HMP estimation results. The content set in above is sorted out and connected with the current

algorithm in an orderly manner. So far, the design of the estimation algorithm of HMP based on multi-labeling transfer learning is completed.

EXPERIMENTAL ANALYSIS AND RESULTS

Data set

The experimental environment is based on a Windows 10 64-bit system, Memory is 32GB.GPU version is NVIDIA GeForce RTX 4060 8GB. The TensorFlow framework is used to train the model. The experimental data used in this experiment are all from public datasets. After the analysis of a large number of datasets, Human3.6M data set (http://vision.imar.ro/human3.6m/) and MPII data set(http://human-pose.mpi-inf.mpg.de/#download) are selected as experimental data sets. Human3.6M data set contains 10 volunteers, including 5 men and 5 women. Each volunteer makes 10 different groups of actions without organization. In the process of data collection, there is no restriction on the clothes of volunteers, but the shooting scenes are the same. After the image acquisition is completed, the image set is labeled according to the human bone structure to obtain the final data set used in the experiment.

MPII data set is a large single human posture estimation data set. Most of the images in this data set come from video websites, including 300 kinds of human motion, which covers all human actions in daily life. The MPII data set used in this experiment contains a total of 5000 images from 500 volunteers. This part of the image is labeled according to the human bone structure, and the experimental data are obtained.

After the experimental data set is set, it is divided into training set and experimental set. Among them, 20 images in human3.6M data set are training sets, and the other 80 images are training sets; in MPII data set, 1000 images are training sets, and the rest 4000 images are training sets. After the above operations are completed, this part of data is taken as the data basis of the experiment, and the experimental operation process is completed based on these two data sets.

Experimental scheme

In the process of the experiment, the MPII data set is used as the main data set, and the HMP estimation algorithm selected in the experiment is trained according to the preset training group, and then the experiment process is completed using the MPII data set and the experimental set in human3.6M data set. Through the analysis, it can be seen that the flexibility of human legs is high, and it may be integrated with the background, which is easy to affect the posture estimation of human motions. This study will mainly estimate the leg motion estimation in the data set. To better complete the experimental process, the experimental set in the data set is divided into four groups, as shown in Table 1.

The experimental group was treated with HMP estimation algorithm, and the initial learning rate was set to 0.001 during the calculation process, and the learning rate attenuation coefficient was 10000. With the continuous increase in the number of experiments, the learning rate gradually declined. After the completion of the overall experiment, the sample loss in the calculation, and fine-tune the experimental results were removed and analyzed. To better obtain the application effect of the proposed algorithm, there are relatively many comparative indicators in this experiment, and the specific content will be set separately.

Table 1. Results of experimental data set (images/piece)

Experimental group	Number of MPII data set	Number of leg motion	Number of Human 3.6M data set	Number of leg motion
1	1000	200	20	5
2	1000	300	20	10
3	1000	250	20	7
4	1000	250	20	13

To ensure the controllability of the experiment, it is necessary to combine human motion detection with limb detection in the experiment. At the same time, the maximum calculation inhibition is set during the experiment, and the image filtering processing is completed in the calculation process, to ensure the accuracy of the experimental results and not affect the experimental analysis process. In this study, the algorithms of RGME[7], HJTEM[8], PENST[9], MAPE[10], and LHPE[11] are selected for comparison and analysis with the proposed algorithm. Therefore, the differences between different algorithms are determined, and the advantages and disadvantages of the proposed algorithm are analyzed more comprehensively.

Evaluation criteria

According to the previous research results, the evaluation indicators of the evaluation algorithm of HMP are set as the following aspects in this experiment.

The correct recognition rate of single HMP

$$y(k) = |A| * \frac{1}{k(i)}$$

(19)

where $|A|$ is the number of images of specified actions in the experimental set; $k(i)$ denotes the number of correctly recognized images of the specified action. This index analyzes the HMP recognition ability of the proposed algorithm and other algorithms. However, only through this index can the single image recognition ability be analyzed, and it is impossible to comprehensively analyze the overall recognition ability of HMP of the algorithm selected in the experiment. Therefore, the over-recognition rate index is set to obtain the global image analysis ability of the algorithm in the experiment.

Over recognition rate of HMP

$$y'(k) = |\overline{A}| * \frac{1}{k(i)}$$

(20)

where $|\overline{A}|$ refers to the number of images of non-specified actions in the experimental data set. Through the above Equation, the negative proportion of HMP error recognition results is analyzed, and the integrity analysis of the algorithm in the experiment is realized.

Image annotation measurement of HMP

In the calculation process of this index, to improve the pertinence of the experiment, an image in the experimental set is selected as the experimental object to complete the analysis process. The specific experimental image is shown in Figure 3.

Figure 3. Experimental subjects

In the process of HMP estimation, after the HMP is completed, the image needs to be labeled to complete the subsequent motion estimation process. Therefore, labeling measurement is also one of the important indicators to measure the application effect of HMP estimation algorithm. The specific calculation process is as follows:

$$\partial = \frac{g \cap h}{h}$$

(21)

where g is the information points in the image that can be labelled without analysis; h is standard information points can be completed only after manual analysis. According to this ratio, the analysis ability of key points of HMP of different algorithms is determined.

Similarity between HMP

Take Figure 3 as an example, according to the annotation points of HMP image, the image set is divided into multiple blocks. The human actions are sorted, and the action similarity between images is analyzed, so that the estimation algorithm can ensure the correlation between motion postures for control.

$$a_{ij} = \exp\left(\frac{-d^2_{ij}}{\sqrt{2\pi}}\right)$$

(22)

where d_{ij} is Euclidean distance after the same human action moves in two cycles. The smaller the value of d_{ij} is, the closer the motion is, and the more similar the HMP is.

HMP estimation accuracy

After the calculation of the above indicators, the overall process of HMP estimation is completed. The calculation results of HMP estimation are sorted out, and the accuracy of HMP estimation is analyzed with different algorithms. The specific calculation process is as follows:

$$D(b,c) = \frac{\left|f(b)\right| - \left|f(c)\right|}{n}$$

(23)

where $\left|f(b)\right|$ is the joint Euclidean distance of target image; $\left|f(c)\right|$ refers to the Euclidean distance estimation of joint of target image; n stands for the number of human joints in the image.

Estimation efficiency of HMP

The shorter the estimation time of HMP, the higher the estimation efficiency. The specific calculation Equation of HMP estimation time is as follows:

$$T = \sum_{i=1}^{N} t_i$$

(24)

where t_i represents the time consumption of the i th HMP estimation item. N represents the total amount of HMP estimation items.

The above six indicators are the indicators to evaluate the performance of HMP estimation algorithm in this study. The application effects of different algorithms are analyzed according to the experimental scheme and this part of the indicators.

RESULTS AND DISCUSSION

According to the analysis of the experimental results in Table 2. The correct recognition rate of the proposed algorithm is roughly the same as that of other algorithms. However, it is still better than the other algorithms. For human3.6M data set, the average correct recognition rate of single HMP of the proposed algorithm is 97.8%, which is 0.8%,0.99%,0.9%,0.5% and 0.6% higher than RGME[7], HJTEM[8], PENST[9], MAPE[10] and LHPE[11], respectively. For MPII data set, the average correct recognition rate of single HMP of proposed algorithm is 97.8%, which is 1.3%,2.4%,2.1%,1.9% and 2.8% higher than RGME[7], HJTEM[8], PENST[9], MAPE[10] and LHPE[11], respectively. The correct recognition rate of single HMP of the proposed algorithm in this expository text is high, and the fluctuation of the correct recognition rate of single HMP of the proposed algorithm is small, indicating that the recognition process is relatively stable.

Table 2. The comparison results of correct recognition rate of single HMP (unit: %)

Data sets	Algorithm	Group 1	Group 2	Group 3	Group 4	Average
Human3.6M data set	**Proposed**	**97.8**	**97.8**	**97.5**	**98.0**	**97.8**
	RGME[7]	97.0	97.1	97.2	96.5	97.0
	HJTEM[8]	97.0	97.0	97.4	96.1	96.9
	PENST[9]	96.7	97.1	96.5	97.2	96.9
	MAPE[10]	97.6	96.5	97.7	97.4	97.3
	LHPE[11]	97.6	97.5	96.6	97.0	97.2
MPII data set	**Proposed**	**98.0**	**97.8**	**97.9**	**97.5**	**97.8**
	RGME[7]	96.5	96.1	96.3	96.9	96.5
	HJTEM[8]	95.9	96.6	95.0	94.0	95.4
	PENST[9]	95.5	95.6	96.2	95.4	95.7
	MAPE[10]	97.1	96.0	95.5	94.8	95.9
	LHPE[11]	96.2	95.2	94.4	94.0	95.0

From the above results in Table2, it can be seen that the proposed algorithm is relatively less affected by the background image. Therefore, the proposed algorithm has a high correct recognition rate.

Human3.6M data set is used as the basis of the HMP over recognition rate experiment. To reduce the complexity of the analysis of the experimental results, the two experimental data sets are integrated, and only the overall over-recognition rate of the experimental group is calculated. The experimental results are shown in Table 3.

Table 3. The comparison results of over-recognition rate of HMP (unit:%)

| Algorithms	Group 1	Group 2	Group 3	Group 4	Average
Proposed	2.17	3.15	3.27	3.65	3.06
RGME[7]	4.24	4.53	4.67	5.01	4.61
HJTEM[8]	4.62	4.51	4.74	5.01	4.72
PENST[9]	4.55	4.42	4.78	5.07	4.71
MAPE[10]	4.15	4.25	4.65	5.06	4.53
LHPE[11]	4.26	4.40	4.68	5.17	4.63

According to the data in Table3. The average over-recognition rate of HMP in the proposed algorithm is 3.06%, which is 1.55%, 1.66%, 1.65%, 1.47% and 1.57% lower than RGME [7], HJTEM [8], PENST [9], MAPE [10] and LHPE [11], respectively. The over-recognition rate of the proposed algorithm is generally low, so it can be determined that the proposed algorithm has a high ability for motion posture integrity analysis. Compared with the proposed algorithm, the use effect of other algorithms is poor, which will affect the subsequent HMP estimation effect.

By analyzing the contents in Figure 4, it can be seen that the number of labeling points of HMP image in the proposed algorithm is 7, which is relatively large and distributed in some joints of the legs, which are 2, 2, 4, 4 and 2 higher than RGME [7], HJTEM [8], PENST [9], MAPE [10] and LHPE [11], respectively. This shows that the experimental comparison of these algorithms has relatively few marker points, and there is a problem with missing joint marker points. At the same time, Equation (21) is used for analysis, and it can be seen that the proposed algorithm has a relatively low labeling measurement value for HMP image, which is controlled at about 0.0561. The other algorithms have a relatively high measurement value of HMP image annotation, which can prove that the proposed algorithm has a strong ability to identify key points of motion posture.

（a）Proposed   （b）RGME[7]

（c）HJTEM[8]   （d）PENST[9]

(e) MAPE[10]   (f) LHPE[11]
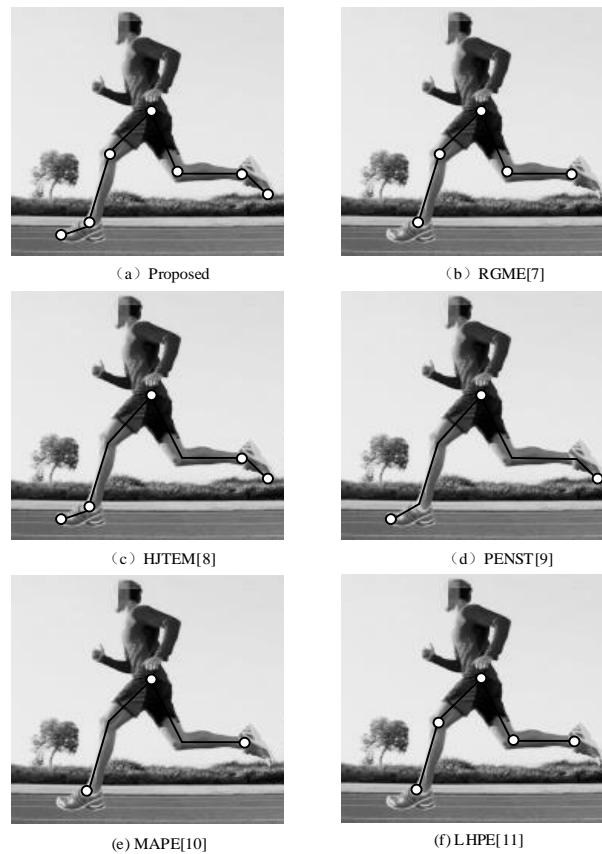
Figure 4. Comparison results of HMP image annotation measurement

According to the results in Figure 5, it can be seen that the similarity between HMP of the proposed algorithm is 92-97%, which is the highest among all algorithms. Particularly when the number of experiments reaches 30, the similarity of the proposed algorithm reaches 97%. Among other compared algorithms, the algorithm of RGME [7] has a relatively high similarity of HMP, with a maximum of about 90%. However, still lower than the proposed algorithm. The similarity between the algorithms of HJTEM [8] and PENST [9] is around 80%. The similarity between the algorithms of MAPE [10] and LHPE [11] does not exceed 70%. In this index experiment, the similarity between HMP obtained by the proposed algorithm is high, and there is a certain correlation between visible images. After the use of other algorithms, the similarity between the estimation result and the target image is low. According to this result, it can be determined that the correlation between the estimation result and the target image can be controlled after the use of the proposed algorithm.
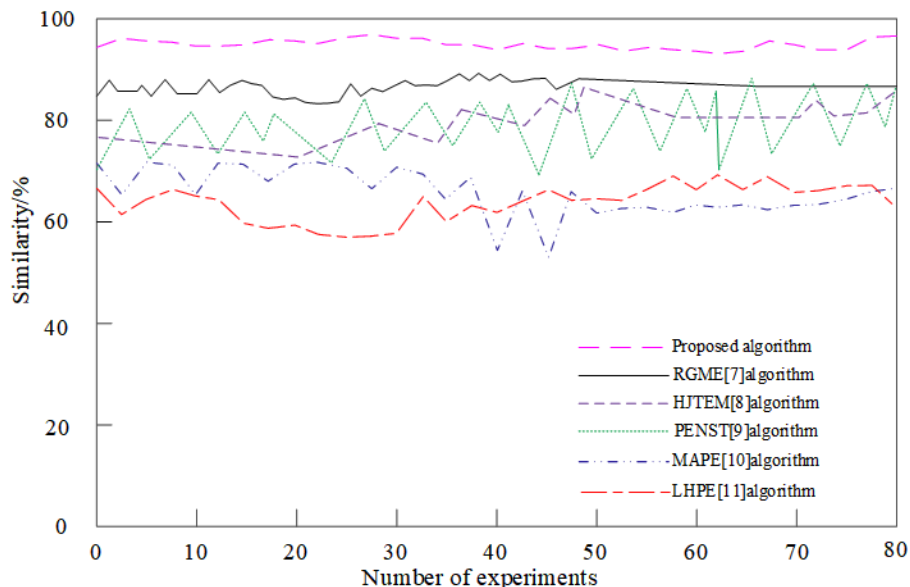
Figure 5. Comparison results of similarity between HMP

The accuracy of HMP estimation of different algorithms is compared, and the experimental results are shown in Figure 6.
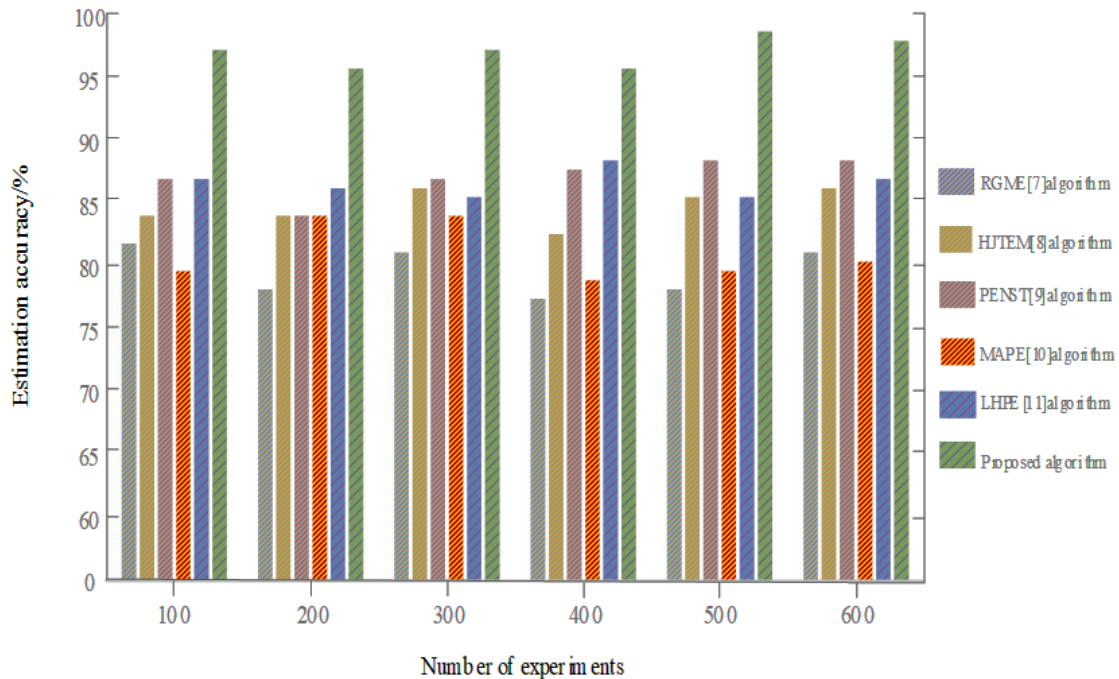
Figure 6. Comparison results of HMP estimation accuracy

The results in Figure 6 shows that the HMP estimation accuracy of the proposed algorithm is as high as 98.1%, among the other algorithms, the estimation accuracy of the algorithms in RGME [7] and MAPE [10] is around 80%, the estimation accuracy of HJTEM [8], PENST[9] and LHPE[11] is around 85%, and the HMP estimation accuracy of the five compared algorithms does not exceed 90%. Compared with the proposed algorithm, the accuracy of HMP estimation of other algorithms is relatively low. Although other algorithms have met the accuracy requirements of current motion posture estimation, the overall level is low. When the type of image background structure increases or the HMP is more complex, the use effect of these algorithms will be affected. In the experiment, the overall computational performance of the proposed algorithm is relatively stable and the application effect is high.

The results of HMP estimation efficiency of different algorithms are shown in Table 4.

Table 4. Comparison results of HMP estimation efficiency (unit: s)

Algorithms	Data sets	
	Human3.6M data set	MPII data set
Proposed	**0.69**	**0.74**
RGME[7]	1.25	1.34
HJTEM[8]	1.32	1.68
PENST[9]	1.57	1.47
MAPE[10]	2.66	2.37
LHPE[11]	2.71	2.57

According to the results in Table 4, it can be seen that for the Human3.6M data set, the estimation time of HMP of the proposed algorithm is 0.69s, which is the lowest among the six algorithms, which is 0.56s,0.63s,0.88s,1.97s and 2.02s lower than RGME [7], HJTEM[8], PENST[9], MAPE[10] and LHPE[11], respectively; For the MPII data set, the estimation time of HMP of the proposed algorithm is 0.74s, which is the lowest among the six algorithms, which is 0.60s,0.94s,0.73s,1.63s and 1.83s lower than RGME [7], HJTEM[8], PENST[9], MAPE[10] and LHPE[11], respectively; Compared with the proposed algorithm, other algorithms have poor estimation time and overall efficiency. Therefore, the proposed algorithm can quickly obtain the estimation results of HMP.

CONCLUSIONS

For the current algorithm of estimation of HMP cannot be well explored potential features in the application process, and there is a certain calculation error. An estimation algorithm of HMP based on multi-labeling transfer learning is proposed in this study, and good research results have been achieved. The experimental results show that the average correct recognition rate of single HMP is 97.8%, the average over recognition rate of HMP is 3.06%, the number of annotation points of HMP image is 7, the similarity between HMP is 92-97%, and the accuracy of HMP estimation is more than 97.3%. The estimated time of HMP is less than 0.74s. This proves that this algorithm has high application value in many fields, such as human-computer interaction and motion, film capture and animation production, pedestrian capture of automatic driving, trace tracking, video indexing and retrieval, identity identification and intelligent monitoring. However, the current study is insufficient for the analysis of actual scenes of human motion and video structure. In future research, further consideration of the background in actual video capture and other complex situations is needed to continuously improve and enhance the effect of HMP analysis.

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

1. Hui Y, Liang Y, Hu X, Wu X, Liu H. Person re-identification combined with style transfer and pose generation, Int J Pattern Recogn.2022;36(2):2256003. doi:10.1142/S0218001422560031
2. Hu T, Xiao C, Min G, Najjari N. An adaptive stacked hourglass network with Kalman filter for estimating 2D human pose in video. Expert Syst. 2020; 38(5):e12552.doi:10.1111/exsy.12552
3. Liu Q. Aerobics posture recognition based on neural network and sensors, Neural Comput Appl.2022; 34 (5):3337-48. doi:10.1007/s00521-020-05632-w
4. Bai G, Luo Y, Pan X, Wang Y,Wang J, Guo J. Double chain networks for monocular 3D human pose estimation, Image Vision Comput.2022; 123:104452. doi: 10.1016/j.imavis.2022.104452
5. Leibovich M, Papanicolaou G, Tsogka C. Synthetic aperture imaging and motion estimation using tensor methods. Siam J Imaging Sci. 2020; 13(4): 2213-49
6. Layton O. ARTFLOW: A fast, biologically inspired neural network that learns optic flow templates for self-Motion estimation. Sensors-basel. 2021; 21(24):8217. doi:10.3390/s21248217
7. Wu R , Xu Z , Zhang J,Zhang L. Robust global motion estimation for video stabilization based on improved k-means clustering and superpixel. Sensors-basel.2021; 21(7):2505. doi:10.3390/s21072505
8. Li X, Liu S, Chang Y, Li S, Fan Y, Yu H. A human joint torque estimation method for elbow exoskeleton control. Int J Hum Robot.2020;17(3):1950039. doi:10.1142/S0219843619500397
9. Li J, Zhuang Y, Peng Q, Zhao L. Pose estimation of non-cooperative space targets based on cross-source point cloud fusion. Remote Sens-basel.2021;13(21):4239. doi:10.3390/rs13214239
10. Lauer J, Zhou M, Ye S, Menegas W, Schneider S, Nath T.et al. Multi-animal pose estimation, identification and tracking with DeepLabCut. Nat Methods, 2022, 19(4):496-504. doi:10.1038/s41592-022-01443-0
11. Wang C, Zhang F, Zhu X, Ge S. Low-resolution human pose estimation. Pattern Recogn, 2022, 126:108579. doi:10.1016/j.patcog.2022.108579
12. Chen W, Liu L, Lin G, Chen Y, Wang J. Class structure-aware adversarial loss for cross-domain human action recognition. Let Image Process. 2021; 15(14):3425-32. doi:10.1049/ipr2.12309
13. Zhi Y, Tom D, Nicola B. Online learning for 3D liDAR-based human detection: experimental analysis of point cloud clustering and classification methods. Auton Robot. 2020; 44(2): 147-64. doi:10.1007/s10514-019-09883-y
14. Li J, Xu S, Qin X, A hierarchical model for learning to understand head gesture videos, Pattern Recogn. 2022; 121:108256. doi:10.1016/j.patcog.2021.108256
15. Zheng Z, Wang Y, Zhang X, Wang J. Multi-scale adaptive aggregate graph convolutional network for skeleton-based action recognition. Appl Sci-basel. 2022;12(3):1402. doi:10.3390/app12031402
16. Qian G, Zhang L, Wang Y. Single-label and multi-label conceptor classifiers in pre-trained neural networks. Neural Comput Appl. 2019;31(10):6179-6188. doi:10.1007/s00521-018-3432-2
17. Park M, Tran D, Lee S, Park S. Multilabel image classification with deep transfer learning for decision support on wildfire response. Remote Sens-basel.2021;13(19)：3985. doi:10.3390/rs13193985
18. Yang K, She W, Zhang W, Yao J, Long S. Multi-label learning based on transfer learning and label correlation. Cmc-Comput Mater Con. 2019;61(1):155-69. doi:10.32604/cmc.2019.05901
19. Xia Z, Xing J, Li X, Gesture tracking and recognition algorithm for dynamic human motion using multimodal deep learning. Secur Commun Netw.2022; 2022:4387337, doi:10.1155/2022/4387337
20. Zhao Y, Yarovoy A, Fioranelli F. Angle-insensitive human motion and posture recognition based on 4D imaging radar and deep learning classifiers, Ieee Sens J.2022;22(12):12173-12182.doi: 10.1109/JSEN.2022.3175618.
21. Li Y, Li K, Wang X, Xu R. Exploring temporal consistency for human pose estimation in videos. Pattern Recogn. 2020;103:107258. doi:10.1016/j.patcog.2020.107258

22. Jongh W, Jordaan H, Daalen C. Experiment for pose estimation of uncooperative space debris using stereo vision. Acta Astronaut. 2020;168:164-73. doi:10.1016/j.actaastro.2019.12.006

23. Huang X, Zhang Y, Chen L, Wang J. U-net-based deformation vector field estimation for motion-compensated 4D-CBCT reconstruction. Med Phys. 2020;47(7):3000-12.doi:10.1002/mp.14150

24. Martín-Doas J, Peinado A, López-Espejo I, Gomez A. Dual-channel speech enhancement based on extended kalman filter relative transfer function estimation. Applied Sciences. 2019; 9(12):2520. doi:10.3390/app9122520

25. Chung J, Ong L, Leow M. Comparative analysis of skeleton-based human pose estimation. Future Internet, 2022, 14(12): 380. doi:10.3390/fi14120380