Training Data Selection and Optimal Sensor Placement for Deep-Learning-Based Sparse Inertial Sensor Human Posture Reconstruction

Zhaolong Zheng [1,2], Hao Ma [1,2], Weichao Yan [1,2], Haoyang Liu [3,*] and Zaiyue Yang [1,2]

[1] Shenzhen Key Laboratory of Biomimetic Robotics and Intelligent Systems, Department of Mechanical and Energy Engineering, Southern University of Science and Technology, Shenzhen 518055, China; 11930587@mail.sustech.edu.cn (Z.Z.); mah@sustech.edu.cn (H.M.); 12031235@mail.sustech.edu.cn (W.Y.); yangzy3@sustech.edu.cn (Z.Y.)

[2] Guangdong Provincial Key Laboratory of Human-Augmentation and Rehabilitation Robotics in Universities, Southern University of Science and Technology, Shenzhen 518055, China

[3] School of Sports Engineering, Beijing Sport University, Beijing 100084, China

[*] Correspondence: liuhaoyang@bsu.edu.cn

Abstract: Although commercial motion-capture systems have been widely used in various applications, the complex setup limits their application scenarios for ordinary consumers. To overcome the drawbacks of wearability, human posture reconstruction based on a few wearable sensors have been actively studied in recent years. In this paper, we propose a deep-learning-based sparse inertial sensor human posture reconstruction method. This method uses bidirectional recurrent neural network (Bi-RNN) to build an a priori model from a large motion dataset to build human motion, thereby the low-dimensional motion measurements are mapped to whole-body posture. To improve the motion reconstruction performance for specific application scenarios, two fundamental problems in the model construction are investigated: training data selection and sparse sensor placement. The problem of deep-learning training data selection is to select independent and identically distributed (IID) data for a certain scenario from the accumulated imbalanced motion dataset with sufficient information. We formulate the data selection into an optimization problem to obtain continuous and IID data segments, which comply with a small reference dataset collected from the target scenario. A two-step heuristic algorithm is proposed to solve the data selection problem. On the other hand, the optimal sensor placement problem is studied to exploit most information from partial observation of human movement. A method for evaluating the motion information amount of any group of wearable inertial sensors based on mutual information is proposed, and a greedy searching method is adopted to obtain the approximate optimal sensor placement of a given sensor number, so that the maximum motion information and minimum redundancy is achieved. Finally, the human posture reconstruction performance is evaluated with different training data and sensor placement selection methods, and experimental results show that the proposed method takes advantages in both posture reconstruction accuracy and model training time. In the 6 sensors configuration, the posture reconstruction errors of our model for walking, running, and playing basketball are $7.25°$, $8.84°$, and $14.13°$, respectively.

Keywords: Bi-RNN; Max-Relevance and Min-Redundancy; training data selection; pose estimation; optimal sensor placement

1. Introduction

Motion capture has a wide range of applications in medical rehabilitation, virtual reality, sports training and other fields [1–3]. However, the current inertial motion-capture devices require subjects to wear 17 or more inertial sensors [4]. It can be intrusive, time-consuming, and prone to sensor misplacement during mounting. Tracking with a minimal, lightweight configuration of sensors is therefore desirable. Many studies have shown

that human motion includes a lot of redundant information and can be described by dimensions lower than the degree of freedom of human motion [5–7]. This opens the door to the study of using sparse inertial sensors human motion capture. To improve the ease of use of inertial motion-capture technology, in recent years researchers have begun to pay attention to motion-capture technology based on fewer wearable inertial sensors [8–10]. Existing research mainly focuses on motion reconstruction methods based on motion prior data and models. For example, a priori model of human motion including local posture linear regression models based on multilayer neural networks and deep-learning recurrent neural networks has been constructed, using 5–6 inertial sensor measurements as input to the model to predict the whole-body posture. Based on the attitude estimation of the prior model, von Marcard et al. merged the prior model with inertial measurements and proposed a motion reconstruction method with offline optimization.

At present, the research of motion-capture technology based on sparse inertial measurement is in the preliminary stage. Several motion prior models have been proposed for general motion capture, but the posture reconstruction accuracy is still limited. In addition, the existing research has not studied the influence of the position of the sparse inertial sensor on the reconstruction accuracy. The motivation of the work is to improve the accuracy of the sparse inertial sensor pose reconstruction, especially for certain target application scenarios. We believe that selecting appropriate training data and the optimal sensor placement on the human body are promising means to improve the accuracy of pose reconstruction under specific application scenarios.

Current methods mainly work in general purpose motion capture, where a large human motion dataset with various movements is used for model training. The basic assumption of training data for machine learning models is that the collected data should be independent and identically distributed (IID) in the target application scenario [11]. The model trained for general purpose may not be optimal for a specific application. Because the dataset is usually accumulated through many tests, it maybe imbalanced in different types of movements. The posture reconstruction accuracy has potential to be improved for certain applications where the movements are more regular and predictable. However, the practical challenge is that the existing motion-capture dataset may not fulfill the IID condition of a specific application scenario, and collecting large amount of IID data from each target application scenario can be time-consuming. To fully use existing dataset, we propose a new data selection method prior to training machine learning models according to the target application scenario. The general idea is that first we collect a small amount of IID reference data from the target scenario. Then, continuous data with small redundancy are selected from the existing dataset, so that they are complied with the a priori distribution of the reference data.

In addition to selecting appropriate training data, the placement of the sparse inertial sensors also plays an important role in the posture reconstruction. It is also worth studied for the optimal performance in different applications. For example, in normal gait, the thigh and shank sensor measurements have strong relevance, and it is possible to infer thigh posture according to the shank posture. In this case, these two sensors have large information redundancy. However, when playing football, the leg movement is highly random in a long period of time, where the independence between the thigh and shank measurements is larger. Generally, the relevance and independence among groups of sensors on the human body vary under different scenarios. The choice of sparse sensor placement will have an obvious impact on the body posture reconstruction accuracy. It is necessary to establish a model of the information volume related to the sensor configuration. With this model, we can determine the optimum number and placement positions of the sensors, and in this way, we can improve the accuracy of motion capture while ensuring the ease of use of the system.

In this paper, we first introduce a method to choose appropriate training data from the accumulated dataset, so that the selected data have similar distribution to data collected from the specific application scenario and has small redundancy. Then, the optimum

number and placement position of the sparse inertial sensors are determined through information volume evaluation. Finally, the selected training data are used to train a deep-learning neural network, where the inputs are the motion measurements of the selected sensors and the output is the whole-body posture. The body posture reconstruction performance under different training data and sensor placements are evaluated through experiments. In summary, the novelty and contributions of this paper include:

- A novel optimization-based efficient data selection scheme is proposed to select continuous and IID data with small redundancy from a large dataset for a specific application scenario.
- The inertial sensor placement problem in sparse inertial sensor posture reconstruction is first formulated into an optimization problem, and a heuristic optimization method is proposed based on a metrics of the amount of information and the redundancy of a group of sensors.

The rest of this paper is organized as follows. Section 2 introduces related work. Section 3 provides overview of the proposed methods. Section 5 introduces the theory and algorithm of training data selection. Section 6 describes the correlation and the redundancy evaluation of the motion information collected in different sensors and the corresponding sensor position selection algorithm. Section 7 presents the experimental results, and we draw some conclusions in Section 8.

2. Related Work

2.1. Inertial Motion Capture

The inertial motion-capture systems include multiple inertial measurement units (IMUs) installed on various segments of the subject's body, which can obtain motion measurements including acceleration, angular velocity and posture of body segments. Inertial motion-capture technology is relatively mature, and the device is portable, not affected by motion occlusion, and the data frame rate is high. Roetenberg et al. has proposed an inertial capture system based on Kalman filtering that can integrate 17 IMUs [4]. In order to improve the accuracy of inertial motion capture, related studies in recent years mainly focus on automatic calibration of inertial sensor installation parameters [12–14], motion reconstruction constraint conditions [15], and nonlinear optimization-based pose reconstruction methods [16].

Studies have shown that human motion includes a lot of redundant information and can be approximately presented by dimensions lower than the original degree of freedom of human motion [6,17,18], which inspired the study of sparse sensor motion capture. Some researchers developed motion-capture technologies combining sparse inertial sensors with video input [7,19] or optical reflective markers [20] or using only sparse optical markers [21]. Although these studies have achieved acceptable posture reconstruction performance, their applications are limited due to optical occlusion issues and high cost of use. In this study, we decided to use only commercially available IMUs for convenience without location measurements or a depth camera.

Existing research on human body posture reconstruction based on sparse inertial sensors mainly relies on motion prior data and models, including multilayer neural networks [10], local pose linear regression models [22], and deep-learning neural networks. A priori model of human motion based on a bidirectional recurrent neural network (Bi-RNN) using 5–6 IMUs to predict the body posture has been proposed [9]. According to the attitude estimation based on the prior model, von Marcard [8] et al. merged the prior model with inertial measurements, and proposed a motion reconstruction method based on offline optimization. At present, the research of motion-capture technology based on sparse inertial measurement is still in the preliminary stage. Several motion prior models have been proposed, but these models have limited accuracy for human body posture reconstruction in certain application scenarios.

2.2. Training Data Selection

Training data selection is an active research topic in machine learning. Through effective selection of training data, more informative samples are extracted, redundant samples and noise data are eliminated, to achieve better learning performance [23]. The current common training data selection methods include methods based on sampling, clustering, and information theory.

The sampling method is the widely adopted basic strategy to reduce the training set, which is simple, easy and fast. Some methods randomly selected training data according to the spatial nature of the sample [24–26]. Although they can effectively reduce the number of samples, the generalization ability of the obtained model cannot be guaranteed. Based on simple random sampling, the methods of uniformly and randomly selecting a compressed set were also studied [27–29].

Random selection of training samples cannot guarantee that the model works reliably in actual tasks. Therefore, scholars proposed to effectively select the training set by considering the distribution characteristics of the data, such as clustering methods. Zhai et al. presented an instance selection method based on supervised clustering, and the main idea is to select instances belonging to inner boundary and outer boundary of clusters [30]. Some researchers used k-means algorithm to partition the dataset into clusters and picked up data from each cluster [31,32]. While using k-means algorithms, the choice of the number of clusters is a challenging issue. In addition, the clustering results are sensitive to the choices of the initial points.

There are also methods using information theory to evaluate and select data samples. Zheng et al. applied a partial mutual information (PMI) technique to find the optimal dataset [33]. Liu et al. explored frame-level data selection based on the normalized frame-level entropy of Gaussian posterior probabilities obtained from the data [34]. The training data selected by the cross-entropy difference selection method proposed by Robert et al. has a good test performance and only requires a small amount of training data [35]. However, existing data selection methods are mainly used for the data reduction of large datasets to improve the computational efficiency of the general model training. When targeting certain application scenarios, researchers usually need to collect new datasets. Therefore, how to select IID training data for specific application scenarios from general datasets is a new topic that deserves further investigation.

2.3. Sensor Placement Selection

In existing research of sparse inertial sensors human motion reconstruction, sensor placement was usually manually decided, such as on the pelvis, head, forearms, and shanks [9]. In the past few years, some researchers have conducted research on the influence of sensor configuration on activity recognition. Cole et al. studied the choice of the position of the smart watch on the human wrist to predict smoking action [36]. Orha et al. compared the accuracy of placing the three-dimensional accelerometer on the right hand, right thigh, and chest to determine the best sensor placement based on the accuracy of the neural network classification of human activities [37]. On basis of this, some researchers have investigated more candidate positions for activity recognition tasks [38–40]. Banos et al. studied the influence of sensor placement on human motion recognition [41].

Instead of directly comparing the prediction accuracy of the sensor configurations, information-based techniques have been studied. These methods evaluated the value of sensor configurations without obtaining the testing performance in the final task. In this way, the searching efficiency can be improved. Kunze et al. determined the optimal sensor configuration based on mutual information between sensor data and the recognized action type [42], but this method did not consider redundancy within the sensor measurements. At present, the sensor configuration method for human activity recognition has been well studied, but there is still no related work for posture reconstruction. In addition, compared with the classification problem in activity recognition where there is only one output (i.e.,

action category), the output variables have multiple dimensions in the regression problem of human body posture reconstruction.

3. Method Overview

The current inertial motion-capture devices require the subject to wear about 17 sensors [4], and our goal is to reduce the sensor number while achieving acceptable performance in posture reconstruction accuracy. A deep neural network Bi-RNN is adopted as the basic model to map low-dimensional motion measurements to the whole-body posture. The measurements obtained from an IMU include the posture and acceleration of the sensor relative to the world coordinate system. Human bone kinematics are obtained based on installation parameters obtained through calibration. For better generalization, the input data are aligned with the orientation of the person based on the root bone (i.e., pelvis) orientation. The training of the Bi-RNN model requires a large amount of IID data of the target application scenario. To use the accumulated dataset and save time for transfer learning of posture reconstruction, we proposed a method to select useful training data that meet the IID condition from the accumulated motion dataset. At the same time, the position and number of sensors placed on the human body affect the performance of human posture reconstruction, so we also studied the optimal selection of the sensor placement to improve the reconstruction accuracy for target applications. An overview of the entire pipeline of the proposed method is shown in Figure 1. The upper left part of Figure 1 is the training data selection. We collect corresponding sample data according to specific application scenarios, and then select the data which has similar distribution to the sample data with small redundancy from the accumulated motion dataset. The bottom left part of Figure 1 is the optimal sensor configuration. We train the neural network on the right side of Figure 1 based on the selected training data and sensors. In actual use, the measurements of selected sensors is used as the input of the neural network, and finally the neural network outputs the posture.
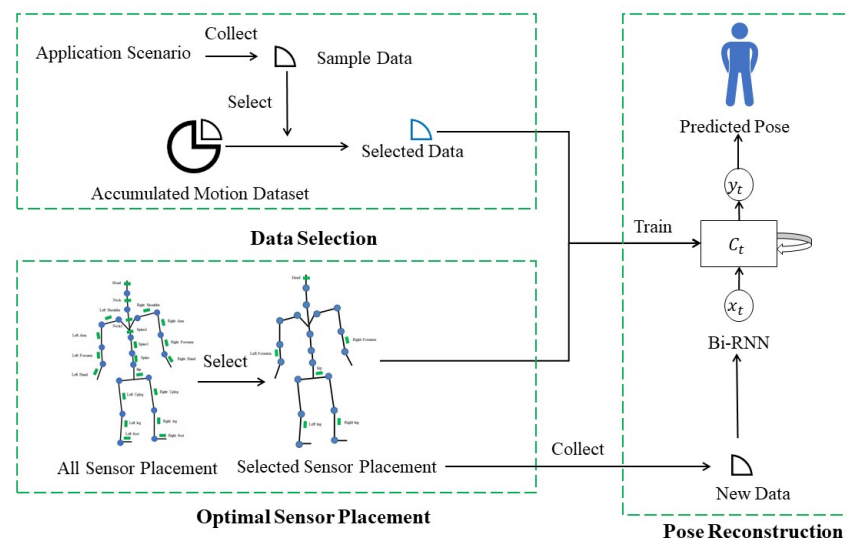
Figure 1. The workflow of data selection, optimal sensor configuration and pose reconstruction.

4. Sparse Sensor Pose Reconstruction

4.1. Data Pre-Processing

In this study, inertial motion-capture suit Perception Neuron Studio is used to obtain human motion data [43]. Data obtained from the inertial sensors include the acceleration $^{W}a_S$ and rotation matrix $^{S}_{W}R$ of the sensor with respect to the world coordinate system W. To express the bone kinematics in W as (1), the kinematics of the sensors must be subjected to a set of calibrations wherein the orientation of sensor with respect to the bone is obtained. To identify the sensor to body alignment, the subject is asked to stand in a known pose for

calibration. The rotation from sensor to body $_{S}^{B}R$ is determined by matching the orientation of the sensor in the world coordinate system $_{W}^{S}R$ with the known orientation of each bone $_{W}^{B}R$ in this pose.

$$_{W}^{B}R =_{W}^{S}R \cdot_{S}^{B}R \tag{1}$$

For better generalization, the input data are normalized with the person's orientation. Then, we standardized all bone orientations with respect to the root bone, which is the pelvis. $_{w}^{Root}R$ denotes the orientation of the root bone with respect to the world coordinate system, and $_{w}^{Root}a$ denotes the acceleration of the root with respect to world coordinate system. The normalized orientation and acceleration of each bone are obtained as (2) and (3).

$$_{Root}^{B}R =_{Root}^{W}R \cdot_{W}^{B}R \tag{2}$$

$$^{Root}a_{B} =_{Root}^{W}R \cdot \left(^{W}a_{B} -^{W}a_{Root}\right) \tag{3}$$

4.2. Bi-RNN Model

Human body motion pattern has temporal characteristics, and the pose of the current frame is often correlated with history and future motion data. Bidirectional Recurrent Neural Network (Bi-RNN) is suitable to learn the temporal relationship between history and future pose data and the current frame pose [44], so Bi-RNN is adopted in this study. The basic structure of our Bi-RNN model refers to the work done by Huang et al. [9], which is shown in Figure 2. Given a training dataset $D = \{(x^i, y^i)\}_{i=1}^{N}$ including N data frames, the objective is to train a $f : x \rightarrow y$ function which can predict the distribution of the kinematic data of unselected sensors y from the measurements of selected sensors x. Huang et al. tried to fit the model to predict the human body Skinned Multi-Person Linear Model (SMPL). SMPL is a parametrized model of 3D human body and pose that takes 72 pose parameters and 10 shape parameters, and returns a mesh with 6890 vertices. Unlike Huang et al. that train the model to predict both pose and shape parameters, we directly predict the body posture. Numbers in brackets in Figure 2 are the input dimensions, output dimensions and number of units in the respective layer. The dropout probability of the input layer is 0.2. The middle layer is composed of Long Short-Term Memory (LSTM) units, and the sequence length of Bi-RNN is 300 frames including 150 frames of both past and future data.

The inputs of the Bi-RNN are the acceleration vector and orientation in rotation matrix format of selected bones, and the outputs are the acceleration norm and orientation in axis angle format of all other unselected bones. Through comparison of the reconstruction performance using different orientation formats in the Bi-RNN, the combination of rotation matrix input and axis angle output achieves the highest accuracy in the experiments.

Figure 2. The structure of the Bi-RNN, and n is the sparse sensor number.

5. Training Data Selection

5.1. Problem Formulation

Conventionally, the entire available dataset is used for deep neural network training. However, there may be redundant data or irrelevant data, which may affect the training performance of the model for specific application scenarios. For example, when learning motion model for walk reconstruction, the motion-capture data of other types of activities, such as run, jump, and dance, are irrelevant and may not contribute to the reconstruction performance of walk. In addition, human walk is a repetitive movement consisting of consecutive similar gait cycles, and a long sequence of normal walk data with consistent

pattern may have large redundancy, which will only increase the training time of the model. Therefore, selecting appropriate training data according to the target scenario is desirable to improve the accuracy of pose reconstruction and reduce model training time.

We denote the accumulated dataset as D_A and the motion feature distribution within the dataset can be different from the target application scenario. To select IID data for a specific application scenario from the accumulated dataset, we first collect a small amount of reference sample data D_{Ref} in this scenario and estimate its a priori distribution of movements features. The goal of training data selection is to re-sample continuous data segments from D_A to obtain actual training data D that has similar distribution to the reference data with small redundancy. Then, the selected training data and reference data are used for the Bi-RNN pose reconstruction model training. The training data selection problem can be formulated into an optimization problem as (4). Here, D is the selected data, which has K continuous segments with various lengths. $H(D)$ is the amount of information of the selected data, i.e., the entropy of D, d_{DP} evaluates the distribution difference between D and D_{Ref}, and $N(D)$ means the amount of data in D. The objective is to select adequate data with enough information for the target scenario.

$$\max_{D \subseteq D_A} \quad \alpha H(D) - \beta d_{DP}(D, D_{Ref}) - \gamma N(D), \tag{4}$$

Assuming D includes K motion data segments and D_A includes M motion data segments, D can be expressed as $\{S_k\}, k \in [1, K]$ where $S_k = [m_k, t_k, \tau_k]$ representing a data segment selected from m_k^{th} data segment from D_A. $S_k = \{f_j\}, j \in [1, \tau_k]$ starts from the t_k^{th} frame and takes τ_k frames in total. Every data frame f_j contains N_f features denoted by $\{\theta_i\}, i \in [1, N_f]$. Because the value function is non-derivable and discontinuous, it is challenging to solve this problem from the perspective of theoretical analysis. Thus, a heuristic algorithm is proposed to solve the problem into two steps. First, a similar and continuous dataset $D_B \subseteq D_A$ is chosen from the dataset using cosine angle to evaluate the difference between the chosen data and the reference data. In this way, we can delete irrelevant motion data from D_A, reducing the search space of heuristic algorithm. Second, we select data with distribution similar to the reference data from D_B by adopting a heuristic algorithm to find an approximate optimum value of the optimization problem. An illustration of the data selection process is shown in Figure 3.
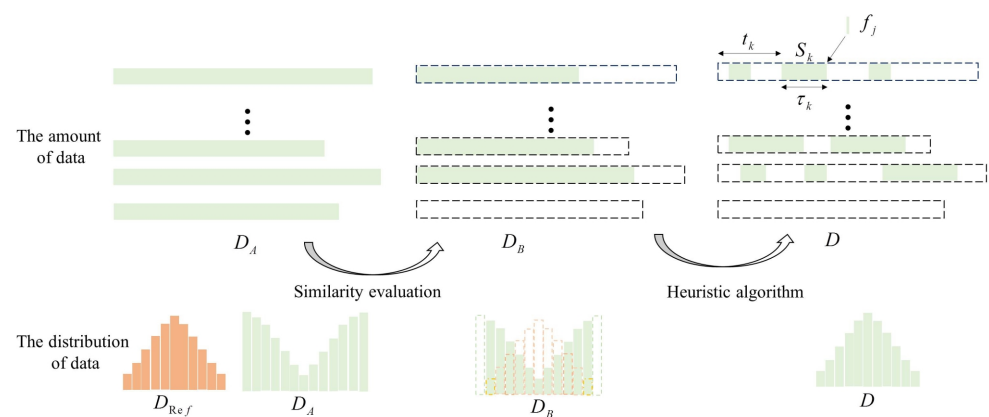

Figure 3. Illustration of the training data selection process. The upper part of the figure is the data segments selected during the procedure. The lower part of the figure shows the data distribution in histograms of both reference data and selected data. In this first step of similar data selection, the selected data segments share the same range of values with the reference data. In the second step of heuristic algorithm of data selection optimization, the selected data segments tend to have similar distribution as the reference data.

5.2. Similar Data Selection

The data selection problem is equivalent to determining a Boolean flag of selected or not for every data frame in the whole dataset. Since the dataset may have a huge number of frames inside, the variable state space to be optimized can be huge as well. Considering that continuous pieces of data segments are preferred in the data selection for Bi-RNN training, we divide each complete segment of data in D_A into data units with a length of $N = 50$ frames (i.e., 0.5 s) to lower down the state space dimension. To select data following the IID condition of the target application scenario, we designed a two-step efficient algorithm. First, we select data units from the whole dataset that are similar to any piece of the reference data unit in terms of motion feature values. The selected data are regarded as the similar dataset D_B. The similar dataset further narrows down the state space of optimal training data searching.

In the implementation, the cosine angle metrics is used to evaluate the similarity between a selected data unit u_1 in D_A and a reference data unit u_2 in D_{Ref} as (5). We calculate the similarity between each pair of u_1 and u_2, and data units u_1 whose maximum similarity with any reference data unit $\eta_{max}(u_1, u_2) \geq \eta_0$ are selected into D_B. Since the input sequence length of Bi-RNN is $L = 300$, only data segments with more than L frames are selected. If η_0 is too large, the searching space will be too large; if η_0 is too small, some useful data will be filtered out. Through experimental tuning, $\eta_0 = 0.8$ is chosen in this study.

$$\eta(u_1, u_2) = \frac{1}{N} \sum_{i=1}^{N} \frac{\langle x_{1i}, x_{2i} \rangle}{\|x_{1i}\| \|x_{2i}\|}, x_{1i} \in u_1, x_{2i} \in u_2, \tag{5}$$

5.3. Optimal Data Selection

After D_B is selected from the original dataset D_A, we need to further select training data $D \subseteq D_B$ to follow the IID condition of the target scenario and be as continuous as possible, contain more information, and has small redundancy at the same time. This optimization problem has been introduced in (4). IID condition requires that the joint distribution of the motion features of the selected training data D should be the same as the reference data D_{Ref}. However, evaluating the similarity of the joint distribution of motion features is difficult due to the high dimensionality of the data. When the motion data features of both D and D_B are Gaussian and have similar value range (guaranteed by similar data selection), and the marginal probability distributions of each feature are similar, then the joint probability distribution should be similar.

The Kullback–Leibler (KL) divergence is a measure of the difference between one probability distribution and another. We use the KL divergence to evaluate the marginal distribution difference between each feature of D and D_{Ref}, and obtain the overall distribution distance as (6). Here, N_f is the number of features, $P_i(x)$ is the probability distribution of the ith feature of D, $Q_i(x)$ is the probability distribution of the ith feature of reference data. To discretize each feature to better express its distribution, we divide the value range of every feature into 20 intervals χ according to the maximum and minimum values of each feature to calculate its probability distribution. The calculation of $P_i(x)$ is given as (7) and (8). $N_i(x)$ counts the number of data falling into the data interval of $x \in \chi$. Similarly, $Q_i(x)$ can be calculated.

$$d_{DP}(D, D_{Ref}) = \frac{1}{N_f} \sum_{i=1}^{N_f} \sum_{x \in \chi} P_i(x) log \frac{P_i(x)}{Q_i(x)}, \tag{6}$$

$$P_i(x) = \frac{\sum_{k=1}^{K} \sum_{j=1}^{l_k} \delta_{kji}(x)}{\sum_{k=1}^{K} l_k}, i \in [1, N_f], \tag{7}$$

$$\delta_{kji} = \begin{cases} 1, & if \quad \theta_{kji} \in x, \\ 0, & otherwise. \end{cases} \tag{8}$$

In this paper, we use information entropy to evaluate the amount of information of the data as (9). In the same way, we divide the features into 20 intervals $x \in \chi$ according to the maximum and minimum values of each feature to calculate the probability distribution of each feature. $N(D)$ is the number of data we selected given in (10), which needs to be minimized to remove redundant data.

$$H(D) = \frac{1}{N_f} \sum_{i=1}^{N_f} \sum_{x \in \chi} P_i(x) log P_i(x). \tag{9}$$

$$N(D) = log(\sum_k l_k). \tag{10}$$

In summary, the cost function of the training data selection optimization problem can be rewritten as (11). Considering maximizing the information term, it is desirable to choose the data distribution P similar to Q and preferably uniformly distributed. When considering minimizing the data amount, data with fewer segments and frames are preferred. Because D is discontinuous and $f_{ds}(D)$ is non-derivable, it is difficult to find the optimal value of f_{ds} from the perspective of theoretical analysis, so we propose a heuristic searching algorithm. To solve the optimization problem, a greedy algorithm is adopted to break down the original problem into a series of sub-problems, and in each iteration, we try to identify the optimal data segment that can be deleted from the existing selected dataset. Initially, set $D = D_{\mathcal{B}}$ which can be expressed as $\{S_{k_B}\}$, $k_B \in [1, K_B]$. To delete the first data segment $S_{k_B,0}$, we need to consider the f_{ds} value and the data segment satisfies (12) will be deleted. The next data segment $S_{k_B}^*$ to be deleted is the data segment satisfying (13). Equation (13) is iterated until $f_{ds}(D \backslash S_{k_B}^*)$ converges to a local minimum.

$$\max_{D \subseteq D_{\mathcal{B}}} \quad f_{ds}(D) = \frac{\alpha}{N_f} \sum_{i=1}^{N_f} \sum_{x \in \chi} P_i(x)[log(P_i^\alpha(x)) - log(\frac{P_i^\beta(x)}{Q_i^\beta(x)})] - log(\sum_k l_k) \tag{11}$$

$$S_{k_B,0} = \underset{S_{k_B}, k_B \in [1, K_B]}{\arg \min} \quad f_{ds}(S_{k_B}) \tag{12}$$

$$S_{k_B}^* = \underset{S_{k_B} \subseteq D, k_B \in [1, K_B]}{\arg \max} \quad f_{ds}(D \backslash S_{k_B}) \tag{13}$$

6. Optimal Sensor Placement

6.1. Problem Formulation

The purpose of the sensor selection is to find a sparse sensor configuration that uses fewer sensors and has high body posture reconstruction accuracy. Generally, when selecting sensors, only the amount of information that can be measured by different sensors is considered. However, our objective is to estimate the measurements of the unselected sensors from the measurements of the selected sensors, thus, we need to maximize the correlation between the measurements of the selected sensors and the unselected sensors. Due to correlation between the sensors, the selected sensors will also contain related measurements. Traditional information-based feature selection methods usually only consider the correlation between the selected sensors and the unselected sensors. In this way, the sensors with more information redundancy are usually selected at the same time, and the sensor with less correlation with the unselected sensors but no redundancy is usually ignored. This is actually not preferred.

In order that the selected sensors can measure as much body motion information as possible, this paper draws on the idea of Max-Relevance and Min-Redundancy (mRMR)

feature selection method in sensor selection [45]. It not only considers the relevance of the candidate sensor and the unselected sensors, but also considers the redundancy among the candidate sensor and the selected sensors. In this section, we introduce the relevance evaluation of the candidate sensor and the unselected sensors, and the measurement of the redundancy between the candidate sensor and the selected sensors. Then, we determine the optimal sensor placement based on the proposed metrics.

All available sensor placements in an inertial motion-capture system are shown in Figure 4. There are a total of 21 positions to be selected, including hands, forearms, arms, thighs, shanks, feet, shoulders, spine, root bone, head and neck. Because it is necessary to calculate the relative posture and acceleration of each bone to the root bone, a sensor placed at the root bone (i.e., pelvis) is required.
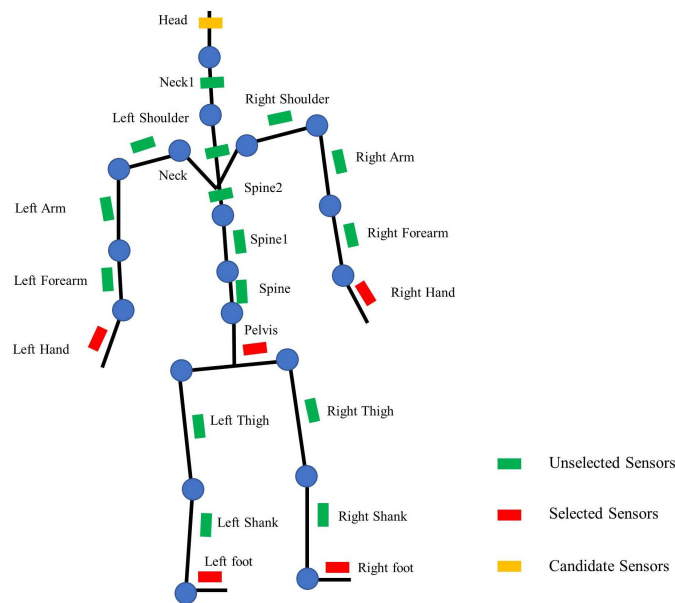
Figure 4. Available positions of inertial sensors in an inertial motion-capture system.

6.2. Maximum Information Coefficient

The maximum information coefficient (MIC) proposed by Reshef et al. can evaluate most of the relationships between two sets of data, including linear and nonlinear relationships [46]. Compared with other coefficients, MIC has the characteristics of universality, fairness and symmetry, and can accurately evaluate the connection between various sensor measurements. The idea of MIC is to discretize variables in a two-dimensional space based on the relationship between two variables and use a scatter plot to represent them. The coefficient will divide the two-dimensional space into a certain number of intervals in the x and y directions, and then check how the current scatter points fall into each grid. Equation (14) gives the definition of the MIC, where $I(X,Y)$ is the mutual information between two variables X and Y, a, b are the number of grids divided in the x, y directions, and B is about 0.6 power of the amount of data.

$$MIC(x;y) = \max_{|a||b|<B} \frac{I(x;y)}{log_2(min(|a|,|b|))} \tag{14}$$

6.3. Max-Relevance and Min-Redundancy

Suppose there are sensor position X and sensor position Y, and the dimension of measurement feature of each sensor is s, then the correlation between the two sensor position measurements is defined as (15). We sum up the MIC between each feature in the measurements of sensor position X and Y to obtain the correlation value.

$$Rele(x;y) = \sum_{i=1}^{s} \sum_{j=1}^{s} MIC(x_i; y_j) \tag{15}$$

We use Max-Relevance and Min-Redundancy (mRMR) to evaluate the measurement information quality of a sensor group. Then, the purpose of optimal sensor selection is to find a sensor group that has greater correlation with the unselected sensors and less redundancy among the selected sensors. First, we find the first sensor with the greatest correlation with all unselected sensors based on the principle of maximum correlation. Second, we use the principle of maximum correlation and minimum redundancy to find the sensor with the largest correlation with the unselected sensor minus the redundancy with the selected sensor as the next sensor to be chosen. This iteration is repeated until the number of sensors and the accuracy of posture reconstruction meet the requirements.

Specifically, the purpose of maximum correlation is to find a sensor group S with the highest correlation with the unselected sensor group C. We add up the correlations between all selected sensors and unselected sensors, and divide by the number of selected and unselected sensors for normalization. The correlation between selected sensors and unselected sensors is defined as (16). On the other hand, the redundancy between sensors in the selected sensor group is defined as (17). The smaller the value of R, the lower the redundancy between sensors. In combination with the correlation and redundancy, the criteria for evaluating the information quality of a sensor group is defined as (18). The larger Φ is, the larger the amount of information can be obtained by the sensor group measurements in the posture reconstruction task.

$$maxT(S,C), \quad T = \frac{1}{|S||C|} \sum_{x_i \in S} \sum_{c_j \in C} Rele(x_i; c_j) \tag{16}$$

$$minR(S), \quad R = \frac{1}{|S|^2} \sum_{x_i \in S} \sum_{x_j \in S} Rele(x_i; x_j) \tag{17}$$

$$max\Phi(T,R), \quad \Phi = T - R \tag{18}$$

6.4. Sensor Selection

A greedy algorithm is adopted to find the optimal solution of Φ. When selecting the first sensor, only the correlation between the sensor and the unselected sensor is considered. The greater the correlation, the more important the sensor is for measuring body motion information. We calculate the correlation between each sensor and the unselected sensor through the previous definition and choose the most relevant as the first selected sensor. For better computational efficiency, incremental search method is used based on mRMR to find the approximate optimal sensors. Assuming that $M-1$ sensor groups S_{m-1} have already been chosen, the goal is to find the mth sensor from the unselected sensors $\{S_0 - S_{m-1}\}$ to maximize Φ. Correspondingly, the incremental algorithm needs to optimize the objective function (19).

$$\max_{x_j \in X - S_{m-1}} \left(\frac{1}{|C|} \sum_{x_i \in S} \sum_{c_j \in C} Rele(x_i; c_j) - \frac{1}{|S|} \sum_{x_i \in S} \sum_{x_j \in S} Rele(x_i; x_j) \right) \tag{19}$$

7. Experiment

7.1. Data Collection and Model Training

In the experiments, we collected motion data in daily activities and sports, including walking, running and jumping, playing basketball, football, table tennis, and so on. The entire dataset contains about 800,000 frames of motion data of 4 people with a total length of 2.5 hA total of 20,000 frames of walking, running, and playing basketball are selected as test set, and the rest data are regarded as the training set. On average, the Bi-RNN model was trained on GPU RTX 2080 for about 6 h until convergence.

7.2. Data Selection Algorithm Performance

To verify the performance of data selection algorithm, we train the model separately with the entire dataset, manually selected data, and the data selected by the proposed data selection algorithm, and then test in the corresponding motion of the test set. In this case, the sensor configuration is the default 6 sensor configuration including head, forearms, shanks and pelvis. The pose reconstruction error and the number of data frames used for model training are shown in Table 1. The pose reconstruction error refers to the relative rotation angle between the reconstructed pose and the actual pose. The manual data selection assumes that the person already knows the type of movement of the dataset, and manually selects the corresponding labeled data files as the training data. The general model is trained by the entire dataset. As the results, the model trained on the algorithm-selected data achieves the best performance in the posture reconstruction accuracy in all the motion types. The model trained on the entire dataset performs better than the model trained on manually selected data.

Moreover, the amount of data selected by our data selection algorithm is less than all the data, but higher than the amount of data selected manually. This indicates that in addition to data files with corresponding labels, other data files of other motion types may also contain useful motion data that are similar to target motion type. Our data selection algorithm can extract these effective data segments from different test files across the entire dataset. The smaller amount of data selected by the proposed algorithm significantly reduces the training time of the model to less than 1/3, compared to the entire dataset.

Table 1. Motion Reconstruction Errors of Models with Default 6 Sensors Configuration.

Motion Type	All Data	Manual Selected Data	Algorithm Selected Data
Walk	7.83° (787,000 frames)	8.84° (98,000 frames)	**7.74°** (170,000 frames)
Run	9.92° (787,000 frames)	11.94° (63,000 frames)	**9.88°** (210,000 frames)
Play Basketball	14.64° (787,000 frames)	15.22° (92,000 frames)	**14.52°** (206,000 frames)

7.3. Sensor Selection Algorithm Performance

To validate the effectiveness of the proposed sensor selection algorithm, we test the performance of the algorithm against some other sensor selection methods using the entire dataset. We compare our method with the modified PCA [47] for feature selection. The modified PCA drawn from the PCA of the input data to leverage the relative importance of the principal components along with the coefficients within the principal directions of the data to uncover the ranking of the features and select the top few features. the relationship curve between the average posture reconstruction angular error and the number of sensors (including the root bone) are shown in Figure 5. In addition, we manually selected the default 6 sensors and 10 sensors configuration (hands, arms, thighs, feet, pelvis and head) by experiences as 'intuition' and evaluated its performance as well. It is shown that our sensor selection algorithm outperforms PCA and at the same time more practical and automated than the default manual sensor configuration. As the number of sensors increases, the pose reconstruction error of our sensor selection algorithm gradually decreases, while PCA fluctuates around a certain value. This may be because PCA did not consider the information redundancy between the sensors, resulting in that the information measured by the sensors selected later cannot reduce the posture reconstruction error.
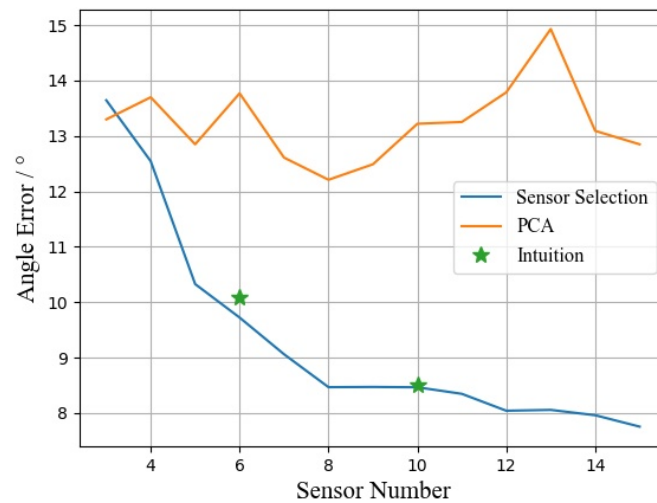
Figure 5. Curve of error and sensor number.

7.4. Performance of Combinations

To evaluate the performance of combinations of both training data selection and sensor placement selection, we further conducted comprehensively comparison study on 6 sensors configuration of different motion types. The experimental results are shown in Figure 6. In the figure, we can see that the algorithm-based sensor selection results is better than 'intuition' and PCA in most cases. The accuracy of the model trained by the proposed data selection method is also better than the model trained on the entire dataset and manually selected data across different motion types.
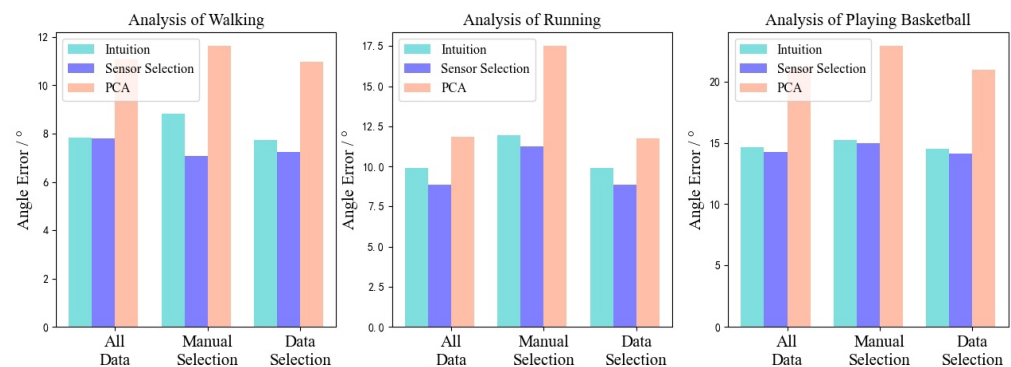
Figure 6. Comprehensive evaluation of algorithm performance.

7.5. Pose Estimation Evaluation

The proposed method has a good posture reconstruction performance for regular exercise reconstructed frames, such as simple walking and running. Table 1 summarizes the posture reconstruction accuracy among different motion types. Figure 7 shows the posture reconstruction performance of some motion frames of playing basketball with obtained optimal 6 sensor configuration. The characters in mixamo [48] are used in the production of the character animation in Figure 7. For some complex sports, such as playing basketball in Figure 7, there are defects in the hand posture. For regular motions, the neural network can easily learn the movement patterns. The complex movements have patterns with more variations, and the amount of complex motion data contained in the training dataset is also limited compared to the regular motions. Thus, the complex motion posture reconstruction performance is worse than simple motion.
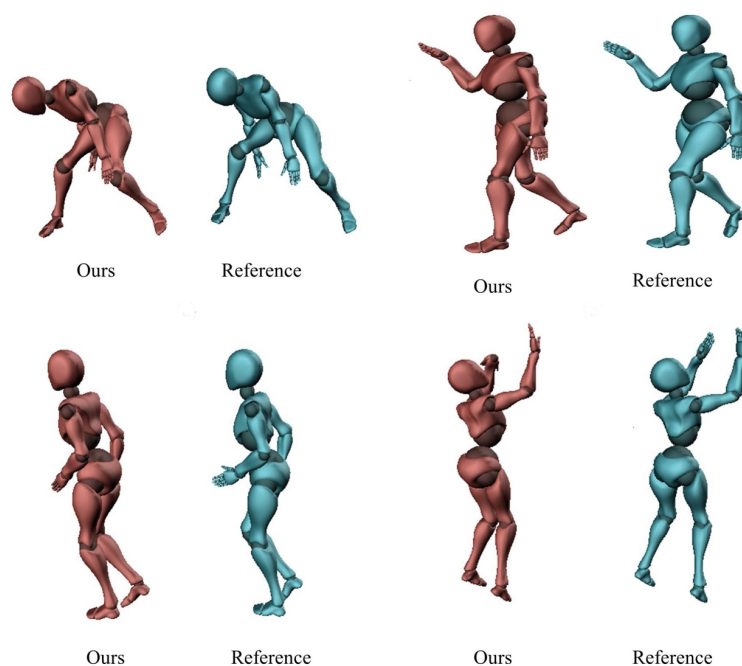
Figure 7. Posture reconstruction results of some motion frames of playing basketball with optimal 6 sensor configuration.

8. Conclusions

In this work, we propose methods for training data selection and sensor position selection in sparse inertial sensor human posture reconstruction under specific application scenarios. The problem of deep-learning training data selection is formulated into an optimization problem. We present a two-step information-based heuristic data selection algorithm by selecting similar data and optimal data from the entire dataset with respect to the reference data collected from the target scenario. Compared with model trained on the entire dataset and manually selected data, the Bi-RNN model trained on the algorithm-selected data have the better performance and less training time.

On the other hand, the sparse sensor position selection is also studied to exploit most information from partial observation of human movement. A greedy algorithm is proposed to search a relative optimum sensor group for maximum motion information and minimum redundancy. Experimental results show that the posture reconstruction performance of the sensor configuration selected by our sensor selection algorithm is better than both PCA and the manually picked default sensor configuration. As the number of sensors increases, the pose reconstruction error of our sensor selection algorithm gradually decreases, which proves that the proposed algorithm takes into account the advantages of sensor measurement information redundancy.

Currently, we proposed heuristic algorithms with greedy strategy to obtain approximate solutions of the established optimization problem. In the future, we will further investigate optimization methods for both training data selection and sensor position selection from a theoretical point of view. Moreover, we will also conduct research on the enhancement of deep neural network structure and the integration of motion analysis for sparse inertial sensor posture construction tasks.

Author Contributions: Conceptualization, Z.Z. and H.M.; Data curation, Z.Z., H.M. and W.Y.; Formal analysis, Z.Z. and H.M.; Methodology, Z.Z. and H.M.; Project administration, H.L.; Software, W.Y.; Supervision, Z.Y. and H.L.; Visualization, W.Y.; Writing—original draft, Z.Z.; Writing—review and editing, Z.Y. and H.L. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable.

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

Abbreviations

The following abbreviations are used in this manuscript:

IID	Independent Identical Distribution
MIC	Mutual Information Theory
PCA	Principal Component Analysis
mRMR	Max-Relevance and Min-Redundancy
Bi-RNN	Bidirectional Recurrent Neural Network
LSTM	Long Short-Term Memory

References

1. Muñoz, J.E.; Chavarriaga, R.; Villada, J.F.; SebastianLopez, D. BCI and motion capture technologies for rehabilitation based on videogames. In Proceedings of the IEEE Global Humanitarian Technology Conference (GHTC 2014), San Jose, CA, USA, 10–13 October 2014; pp. 396–401.
2. Jie, G.; Guiling, C.; Lin, H.; Dong, Z. Application research on motion capture system data reuse in virtual reality environment. In Proceedings of the 2010 Second International Conference on Intelligent Human-Machine Systems and Cybernetics, Nanjing, China, 26–28 August 2010; pp. 114–116.
3. Noiumkar, S.; Tirakoat, S. Use of optical motion capture in sports science: A case study of golf swing. In Proceedings of the 2013 International Conference on Informatics and Creative Multimedia, Kuala Lumpur, Malaysia, 4–6 September 2013; pp. 310–313.
4. Roetenberg, D.; Luinge, H.; Slycke, P. Xsens mvn: Full 6 Dof Human Motion Tracking Using Miniature Inertial Sensors. Xsens Motion Technologies BV. Technical Report. 2009. Available online: http://human.kyst.com.tw/upload/pdfs120702543998066.pdf (accessed on 8 May 2021).
5. Troje, N.F. Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *J. Vis.* **2002**, *2*, 371–387. [CrossRef] [PubMed]
6. Sanger, T. Human arm movements described by a low-dimensional superposition of principal components. *J. Neurosci. Off. J. Soc. Neurosci.* **2000**, *20*, 1066–1072. [CrossRef]
7. Safonova, A.; Hodgins, J.; Pollard, N. Synthesizing physically realistic human motion in low-dimensional, behavior-specific spaces. *ACM Trans. Graph.* **2004**, *23*, 514–521. [CrossRef]
8. Marcard, T.; Rosenhahn, B.; Black, M.; Pons-Moll, G. Sparse inertial poser: Automatic 3D human pose estimation from sparse iMUs. In *Computer Graphics Forum;* Wiley: New York, NY, USA, 2017; pp. 349–360.
9. Huang, Y.; Kaufmann, M.; Aksan, E.; Black, M.J.; Hilliges, O.; Pons-Moll, G. Deep inertial poser: Learning to reconstruct human pose from sparse inertial measurements in real time. *ACM Trans. Graph.* **2019**, *37*, 1–15. [CrossRef]
10. Wouda, F.; Giuberti, M.; Bellusci, G.; Veltink, P. Estimation of full-bod poses using only five inertial sensors: An eager or lazy learning approach? *Sensors* **2016**, *16*, 2138. [CrossRef] [PubMed]
11. Bertsekas, D.P.; Tsitsiklis, J.N. *Introduction to Probability*, 2nd ed.; Number 1 in Optimization and Computation Series; Athena Scientific: Belmont, MA, USA, 2008.
12. Bouvier, B.; Duprey, S.; Claudon, L.; Dumas, R.; Savescu, A. Upper limb kinematics using inertial and magnetic Sensors: comparison of Sensor-to-Segment calibrations. *Sensors* **2015**, *15*, 18813–18833. [CrossRef] [PubMed]
13. Taetz, B.; Bleser, G.; Miezal, M. Towards self-calibrating inertial body motion capture. In Proceedings of the 2016 International Conference on Information Fusion (FUSION), Heidelberg, Germany, 5–8 July 2016.
14. Müller, P.; Bégin, M.; Schauer, T.; Seel, T. Alignment-free, self-calibrating elbow angles measurement using inertial sensors. *IEEE J. Biomed. Health Inform.* **2017**, *21*, 312–319. [CrossRef] [PubMed]
15. Miezal, M.; Taetz, B.; Bleser, G. Real-time inertial lower body kinematics and ground contact estimation at anatomical foot points for agile human locomotion. In Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, 29 May–3 June 2017.
16. Kok, M.; Hol, J.D.; Schön, T.B. An optimization-based approach to human body motion capture using inertial sensors. *IFAC Proc. Vol.* **2014**, *47*, 79–85. [CrossRef]
17. Vlasic, D.; Adelsberger, R.; Vannucci, G.; Barnwell, J.; Gross, M.; Matusik, W.; Popović, J. Practical motion capture in everyday surroundings. *ACM Trans. Graph.* **2007**, *26*, 35. [CrossRef]
18. Trumble, M.; Gilbert, A.; Hilton, A.; Collomosse, J. Deep autoencoder for combined human pose estimation and body model upscaling. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany , 8–14 September 2018; pp. 784–800.

19. Pons-Moll, G.; Baak, A.; Gall, J.; Leal-Taixe, L.; Muller, M.; Seidel, H.P.; Rosenhahn, B. Outdoor human motion capture using inverse kinematics and von mises-fisher sampling. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 1243–1250.

20. Marcard, T.; Pons-Moll, G.; Rosenhahn, B. Human pose estimation from video and IMUs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *38*, 1533–1547. [CrossRef] [PubMed]

21. Andrews, S.; Huerta, I.; Komura, T.; Sigal, L.; Mitchell, K. Real-time physics-based motion capture with sparse sensors. In Proceedings of the 13th European conference on visual media production (CVMP 2016), London, UK, 12–13 December 2016; pp. 1–10.

22. Liu, H.; Wei, X.; Chai, J.; Ha, I.; Rhee, T. Realtime human motion control with a small number of inertial sensors. In Proceedings of the Symposium on Interactive 3D Graphics and Games, San Francisco, CA, USA, 18–20 February 2011; pp. 133–140.

23. Yang, Y.; Lu, Y.; Yao, M.; Xia, Y.; Hao, Y. Training data selection for short term load forecasting. In Proceedings of the 2011 Third International Conference on Measuring Technology and Mechatronics Automation, Shanghai, China, 6–7 January 2011; pp. 1040–1043.

24. Balcázar, J.; Dai, Y.; Watanabe, O. A random sampling technique for training support vector machines. In *Algorithmic Learning Theory*; Abe, N., Khardon, R., Zeugmann, T., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2001; pp. 119–134.

25. Ferragut, E.M.; Laska, J. Randomized sampling for large data application of SVM. In Proceedings of the 2012 11th International Conference on Machine Learning and Applications, Boca Raton, FL, USA, 12–15 December 2012; pp. 350–355.

26. Tsagalidis, E.; Evangelidis, G. The effect of training set selection in meteorological data mining. In Proceedings of the 2010 14th Panhellenic Conference on Informatics, Tripoli, Greece, 10–12 September 2010; pp. 61–65.

27. Lee, Y.J.; Mangasarian, O.L. RSVM: Reduced support vector machines. In Proceedings of the 2001 SIAM International Conference on Data Mining, Chicago, IL, USA, 5–7 April 2001; pp. 1–17.

28. Lee, Y.; Huang, S. Reduced support vector machines: A statistical theory. *IEEE Trans. Neural Netw.* **2007**, *18*, 1–13. [CrossRef] [PubMed]

29. Li, X.; Cervantes, J.; Yu, W. Fast classification for large data sets via random selection clustering and support vector machines. *Intell. Data Anal.* **2012**, *16*, 897–914. [CrossRef]

30. Zhai, J.H.; Xui, H.Y.; Zhang, S.F.; Li, N.; Li, T. Instance selection based on supervised clustering. In Proceedings of the 2012 International Conference on Machine Learning and Cybernetics, Xi'an, China, 15–17 July 2012; pp. 112–117.

31. Koskimäki, H.; Juutilainen, I.; Laurinen, P.; Röning, J. Two-level clustering approach to training data instance selection: A case study for the steel industry. In Proceedings of the 2008 IEEE International Joint Conference on Neural Networks, Hong Kong, China, 1–8 June 2008.

32. Kaneda, Y.; Zhao, Q.; Liu, Y. On-line training with guide data: Shall we select the guide data randomly or based on cluster centers? In Proceedings of the 2016 IEEE Symposium Series on Computational Intelligence (SSCI), Athens, Greece, 6–9 December 2016; pp. 1–7.

33. Zheng, J.; Yang, W.; Li, X. Training data reduction in deep neural networks with partial mutual information based feature selection and correlation matching based active learning. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 5–9 March 2017; pp. 2362–2366.

34. Liu, S.H.; Chu, F.H.; Lin, S.H.; Lee, H.S.; Chen, B. Training data selection for improving discriminative training of acoustic models. In Proceedings of the 2007 IEEE Workshop on Automatic Speech Recognition Understanding (ASRU), Kyoto, Japan, 9–13 December 2007; pp. 284–289.

35. Moore, R.; Lewis, W. Intelligent selection of language model training data. In Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, Uppsala, Sweden, 11–16 July 2010; pp. 220–224.

36. Cole, C.A.; Thrasher, J.F.; Strayer, S.M.; Valafar, H. Resolving ambiguities in accelerometer data due to location of sensor on wrist in application to detection of smoking gesture. In Proceedings of the 2017 IEEE EMBS International Conference on Biomedical Health Informatics (BHI), Orlando, FL, USA, 16–19 February 2017; pp. 489–492.

37. banosDealingEffectsSensor2014, I.; Oniga, S. Study regarding the optimal sensors placement on the body for human activity recognition. In Proceedings of the 2014 IEEE 20th International Symposium for Design and Technology in Electronic Packaging (SIITME), Bucharest, Romania, 23–26 October 2014; pp. 203–206.

38. Dobrucalı, O.; Barshan, B. Sensor-activity relevance in human activity recognition with wearable motion sensors and mutual information criterion. In *Information Sciences and Systems 2013*; Springer: Cham, Switzerland, 2013; pp. 285–294.

39. Ching, Y.T.; Cheng, C.C.; He, G.W.; Yang, Y.J. Full model for sensors placement and activities recognition. In Proceedings of the UbiComp '17 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers, Maui, HI, USA, 11–15 September 2017; Association for Computing Machinery: New York, NY, USA, 2017; pp. 17–20.

40. Murao, K.; Mogari, H.; Terada, T.; Tsukamoto, M. Evaluation function of sensor position for activity recognition considering wearability. In Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication, Zurich, Switzerland, 8–12 September 2013; pp. 623–632.

41. Banos, O.; Toth, M.A.; Damas, M.; Pomares, H.; Rojas, I. Dealing with the effects of sensor displacement in wearable activity recognition. *Sensors* **2014**, *14*, 9995–10023. [CrossRef] [PubMed]

42.　Kunze, K.; Lukowicz, P. Sensor placement variations in wearable activity recognition. *IEEE Pervasive Comput.* **2014**, *13*, 32–41. [CrossRef]

43.　Perception Neuron Studio System Landing Page. 2020. Available online: https://neuronmocap.com/perception-neuron-studio-system (accessed on 9 May 2021).

44.　Bahdanau, D.; Cho, K.; Bengio, Y. Neural machine translation by jointly learning to align and translate. In Proceedings of the 2015 International Conference on Learning Representations (ICLR), San Diego, CA, USA, 7–9 May 2015.

45.　Ding, C.; Peng, H. Minimum redundancy feature selection from microarray gene expression data. *J. Bioinform. Comput. Biol.* **2005**, *3*, 185–205. [CrossRef] [PubMed]

46.　Reshef, D.N.; Reshef, Y.A.; Finucane, H.K.; Grossman, S.R.; McVean, G.; Turnbaugh, P.J.; Lander, E.S.; Mitzenmacher, M.; Sabeti, P.C. Detecting novel associations in large data sets. *Science* **2011**, *334*, 1518–1524. [CrossRef] [PubMed]

47.　Malhi, A.; Gao, R. PCA-based feature selection scheme for machine defect classification. *IEEE Trans. Instrum. Meas.* **2004**, *53*, 1517–1525. [CrossRef]

48.　Mixamo. 2020. Available online: https://www.mixamo.com/#/?page=1&type=Character (accessed on 9 May 2021).