Research Article

Computer Vision for Human-Computer Interaction Using Noninvasive Technology

Janarthanan Ramadoss,[1] J. Venkatesh,[1] Shubham Joshi [ID],[2] Piyush Kumar Shukla [ID],[3] Sajjad Shaukat Jamal [ID],[4] Majid Altuwairiqi,[5] and Basant Tiwari [ID][6]

[1]*Center for Artificial Intelligence and Research, Chennai Institute of Technology, Chennai, Tamil Nadu, India*
[2]*Department of Computer Engineering, SVKM's NMIMS MPSTME Shirpur Campus, Savalade, India*
[3]*Computer Science & Engineering Department, University Institute of Technology, Rajiv Gandhi Proudyogiki Vishwavidyalaya, (Technological University of Madhya Pradesh), Bhopal 462033, India*
[4]*Department of Mathematics, College of Science, King Khalid University, Abha, Saudi Arabia*
[5]*College of Computers and Information Technology, Computer Science Department, Taif University, Taif, Saudi Arabia*
[6]*Department of Computer Science, Hawassa University, Institute of Technology, Hawassa, Ethiopia*

Correspondence should be addressed to Basant Tiwari; basanttiw@hu.edu.et

Computer vision is a significant component of human-computer interaction (HCI) processes in interactive control systems. In general, the interaction between humans and computers relies on the flexibility of the interactive visualization system. Electromyography (EMG) is a bioelectric signal used in HCI that can be captured noninvasively by placing electrodes on the human hand. Due to the impact of complex background, accurate recognition and analysis of human motion in real-time multitarget scenarios are considered challenging in HCI. Further, EMG signals of human hand motions are exceedingly nonlinear, and it is important to utilize a dynamic approach to address the noise problem in EMG signals. Hence, in this paper, the Optimized Noninvasive Human-Computer Interaction (ONIHCI) model has been proposed to predict human motion recognition. Average Intrinsic Mode Function (AIMF) has been used to reduce the noise factor in EMG signals. Furthermore, this paper introduces spatial thermographic imaging to overcome the conventional sensor problem, such as gesture recognition and human target identification in multitarget scenarios. The human motion behavior in spatial thermographic images is examined by target trajectory, and body movement kinematics is employed to classify human targets and objects. The experimental findings demonstrate that the proposed method reduces noise by 7.2% and improves accuracy by 97.2% in human motion recognition and human target identification.

1. Introduction

Nowadays, with the rapid development of information technology, human beings are trying to communicate with computers more naturally [1]. The conventional human-computer interaction input devices such as the mouse, keyboards, and remote devices lack flexibility, and there is no longer a natural way of interacting [2]. In general, voice commands and body language are natural ways for people to communicate with computers, including many online commercial products [3]. The interaction between humans and computers is the most important application of computer vision for autonomous structures [4]. It is essential to acquire precise data like shape, behavior, and motions for efficient human-computer interaction [5]. An effective characteristic analysis of these human targets can accurately recognize and identify the targets [6]. Human target identification and objects surrounding play a crucial role and

pose many challenges before interaction between computers and humans [7]. We rapidly determine the number of important facts and qualities about each other during human-to-human interaction, including identification, age, facial expressions, and gestures. These visual cues/features have an impact on the content and flow of a conversation, and they provide contextual information such as situation and speech context. A gesture or a facial expression, for example, could be intended as a signal of understanding, or the gaze direction can be used to differentiate between the object referred to like this and the direction over there in speech. As a result, other communication channels such as speech and gestures are both coexpressive and complementary to the visual channel.

Conversely, conventional sensors do not deliver an acceptable field of view (FOV) to monitor various targets to examine human movement and body features [8]. The process of human movement identification needs sufficient space to map command gestures and different human targets [9]. It is significant to recognize the human targets and surrounding objects, and the computer aims to satisfy the requirements of the human interaction environment [10]. The user's gesture would be consistent with the physical space of the virtual world; i.e., the user's action should be matched with the gesture in the virtual field, and it is more appropriate to estimate the gesture for human-computer interaction [11]. A rich user experience and more effective and efficient interaction can be obtained by integrating visual information with other input modalities (such as keyboard and mouse). In addition to standard desktop computing, vision-based interaction could be beneficial in a variety of scenarios, including mobile, immersive, and ubiquitous computing.

Presently, sensor technology such as electromyographic (EMG) and signal processing has been extensively utilized in the field of human-computer interaction and multifunctional prosthetic hand control [12, 13]. Electromyographic (EMG) signal collects superficial muscle and nerve trunk activity bioelectric signals through electrodes on the surface of the skin and performs muscle processes evaluations and simulations via the recorded, filtered, amplified, transmitted, and feeding of the collected bioelectric signals [14, 15]. Since the EMG signaling of the human leg or hand movements during object usage easily interferes with noise [16]. The main problem to accomplish a difficult understanding of hand motion is successfully gathering signals, extracting features, and classifying diverse hand movements for human-computer interaction [17]. From the object's characteristics such as weight, size, and shape, it is possible to identify human targets/emotions [18].

In this paper, the Optimized Noninvasive Human-Computer Interaction (ONIHCI) model has been proposed to address gesture recognition and human target identification problems. The electromyographic (EMG) signal can represent the muscles' active conditions; the data of the neural activities can be determined [19]. The advantage of electromyographic is noninvasive; thus, it executes well in studying neurological rehabilitation, motion detection, and artificial control. Besides, the AIMF algorithm has been

employed to reduce noise in the EMG signal. The present models focus on human-computer interaction, emphasizing a specific target or the behavioral analysis of a set of targets [20]. The spatial thermographic images for human motions are analyzed to explain the trajectory actions and the kinematics motion of the human and objects. Our approach defines human targets precisely and allows them to improve their restricted vision and overcome traditional methodological problems related to gesture recognition and human target identification [21]. Target trajectory examines human motion behavior in spatial thermographic images, and body movement kinematics is used to classify human targets and objects.

The rest of the paper is arranged as follows. Section 1 and Section 2 discussed the overview of computer vision for human-computer interaction and related works. In Section 3, the Optimized Noninvasive Human-Computer Interaction (ONIHCI) model has been suggested. In Section 4, the experimental results have been performed. Finally, Section 5 concludes the research paper.

2. Related Works and Features of This Research Article

Qi et al. [22] introduced the linear discriminant analysis and extreme learning machine (LDA-ELM) method for smart human-computer interaction based on surface EMG gesture recognition. The proposed method can minimize the useless data in Surface Electromyography (SEMG) signals and enhance identification accuracy and efficiency. This paper concentrates on time variances optimization in surface [23]. EMG pattern recognition and the numerical outcomes are advantageous to decreasing the time variances in gesture identification based on surface EMG. Chen et al. suggested the Motor Unit Spike Trains with Blind Source Separation Algorithm (MUST-BSSA). That is how well high-density EMG signals recognize motor unit movements during hand postures. They characterize the precision in recognition of motor unit actions during hand postures from high-density EMG signals. The results demonstrate the possibility of recognizing motors during the assigned motor operations and the high precision of hand gestures classification for human-computer interface perspectives [24].

Song et al. [25] discussed the Guidance framework for tracking by detection (GFTD) for hand detection based on the thermal image. They introduced an Adaptive Hand Detection (AHD) based automatic tracking-by-detection algorithm utilizing the Kernelized correlation filters tracker to enhance the performance of the proposed model. The proposed model detects hands in real time by decreasing calculation utilizing a single sensor instead of fusing manifold sensors, enables precise tracking, and enhances hand tracking precision. Xiao et al. [26] proposed the variational mode decomposition and composite permutation entropy index (VMD-CPEI) method to classify hand movements. The approach suggested in this paper uses the VMD procedure for decomposing the initial SEMG signal into multiple VMFs and measuring the related CPEI of each signal component. The model proposed

can enhance the quality of life of amputees, disabled persons, and others.

Sekhavat et al. [19] introduced Affective User Interface Design and Evaluation (AFFUIDE) to evaluate the effect of using facial expressions emotions as a user interface (UI) and system input in virtual scenarios. The data suggests that the traditional usable user interface is the most useful and that the full affective user interface feels the best fun and user experience. Choudhary et al. [20] recognized someone even when they were not in a neutral condition or had any facial expression. Using a hidden Markov model with singular value decomposition, they could identify persons whose top half face is visible (HMM-SVD). Singular value decomposition parameters were used in this article to build a series of blocks for each picture of a face [21]. To cover the entire face, a seven-state HMM was employed. In this paper, the Optimized Noninvasive Human-Computer Interaction (ONIHCI) model has been proposed to address gesture recognition and human target identification [23]. This paper proposes a novel approach to determine the human targets in spatial thermographic images [24, 27] using vision perception. This system estimates the capacity of the human targets to view when they have restricted visualization and supports targets perceiving their surroundings by informing them about their condition through the visualization system [28, 29]. Identifying human targets is the main aim and, at the same time, recognizing the scenario for human targets. The signals of human hand motions have been gathered utilizing an EMG acquisition device [30, 31]. A motor unit classification-based gesture recognition method was proposed in [32]. MUSTs were first classified into 11 categories, one per motion. The averaged discharge timings of MUSTs in each group are then used to measure the activation level of the cerebral drive to each motion. By comparing the estimated activation levels of each motion, the output gesture class was determined. AIMF has been used to denoise the obtained real signs and feature sets for the hand movement classification process [22, 25, 33]. The proposed method is briefly described in the following section.

3. Optimized Noninvasive Human-Computer Interaction Model (ONIHCI)

In this paper, the Optimized Noninvasive Human-Computer Interaction (ONIHCI) model has been proposed to address gesture recognition and human target identification problems. The requirement for computer vision-based human-computer interaction and human movement recognition has been increased in fields like intelligent monitoring, security, and surveillance system. The most significant and common human movement is walking and running. Many studies aimed to develop a computer model of motion. The role of human movement recognition is a challenge in human-computer interaction. Motion analysis includes measuring, analyzing, and evaluating the motion functions associated with walking or running activity to determine the human target in an HCI system.

Case 1. Human kinematics analysis and trajectory analysis for human target identification

Solution 1. Analysis of human kinematics is the initial measure to determine a human in the course of its movement or static prospect. Human frames are extracted from the spatial thermographic image, and areas are calculated from dynamic targets to explore trajectories and kinematics to calculate the proposed approach. The key targets know the instructions and are in an upright position. This paper utilizes three human targets for the identification of human targets.

Figure 1 shows the human body kinematics analysis. The human body orientation around the arm region $[X_B Y_B Z_B]$, legs $[X_{L,R} Y_{L,R} Z_{L,R}]$, and head for the upper part of the targets spot $[X_V Y_V Z_V]$ is determined for every human.

> *Human Target 1 (G_1).* The human body orientation around the arm region is denoted as $[X_{B_1} Y_{B_1} Z_{B_1}]$, left leg $[X_{L_1} Y_{L_1} Z_{L_1}]$, right leg $[X_{R_1} Y_{R_1} Z_{R_1}]$, and head target sport $[X_{V_1} Y_{V_1} Z_{V_1}]$
>
> *Human Target 2 (G_2).* The human body orientation around the arm region is denoted as $[X_{B_2} Y_{B_2} Z_{B_2}]$, left leg $[X_{L_2} Y_{L_2} Z_{L_2}]$, right leg $[X_{R_2} Y_{R_2} Z_{R_2}]$, and head target sport $[X_{V_2} Y_{V_2} Z_{V_2}]$
>
> *Human Target 3 (G_1).* The human body orientation around the arm region is denoted as $[X_{B_3} Y_{B_3} Z_{B_3}]$, left leg $[X_{L_3} Y_{L_3} Z_{L_3}]$, right leg $[X_{R_3} Y_{R_3} Z_{R_3}]$, and head target sport $[X_{V_3} Y_{V_3} Z_{V_3}]$

The human body provides a few signs of motion or static state disposition. The inclination for direction is extracted from the determination of the direction between these kinematics. If the legs and head position are open at a certain angle, the motion direction from the slope of the head is provided by the angle between the legs and head. Figure 1 shows the human body kinematics; here, the human target G_1 is being demonstrated with a running direction to the head and right sloped toward a similar direction. In the instance of a static condition of the target, body, legs, and head orientation comprising arms might be parallel to human targets G_2 and G_3. The interactive system point of view transferred kinematics is provided with G_W and the kinematics from the initial interactive system, $[X_G, Y_G, Z_G]$ are multiplied with orientation and translation, and the interactive systems and rotation between the interactive systems are expressed as

$$G_{W_1} = \begin{bmatrix} X_{G_1} & Y_{G_1} & Z_{G_1} \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} x_t & y_t & z_t \end{bmatrix},$$

(1a)

$$G_{W_2} = \begin{bmatrix} X_{G_2} & Y_{G_2} & Z_{G_2} \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} x_t & y_t & z_t \end{bmatrix},$$

(1b)
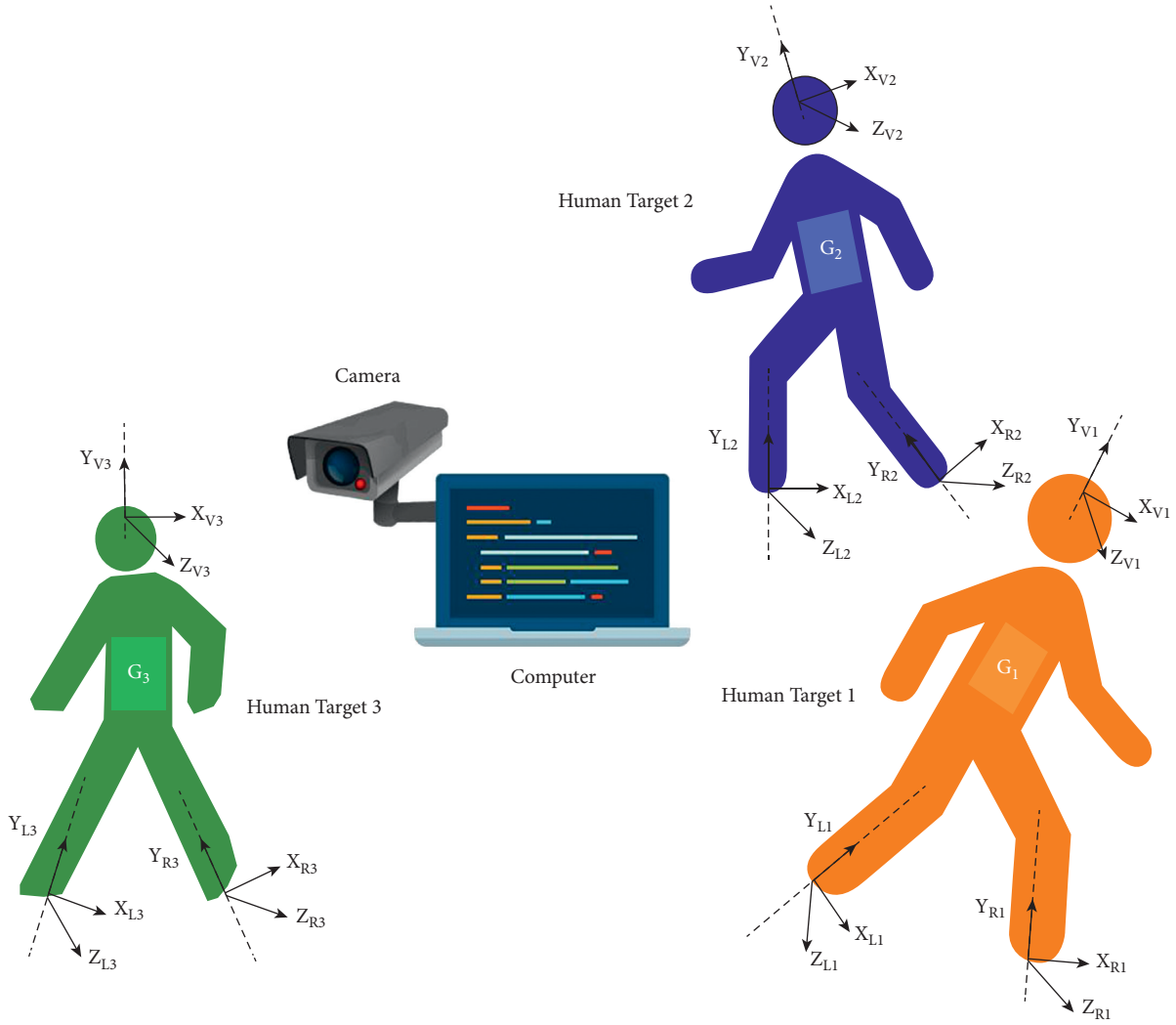
FIGURE 1: Human body kinematics analysis.

$$G_{W_3} = \begin{bmatrix} X_{G_3} & Y_{G_3} & Z_{G_3} \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} x_t & y_t & z_t \end{bmatrix}.$$

(1c)

As inferred from equations (1a), (1b), and (1c), the human body target rotation matrix, kinematics, and translation vector of the transformation vector can be computed. The relationship between two groups of kinematics can be utilized to choose the categorized targets from two interactive systems to connect similar targets in two various thermographic views.

Figure 2 shows the position data of every computer or interactive system T_t, and R_t is utilized to convert the human body kinematics of the other view direction. In Figure 2, the spatial thermographic interactive system coordinate system is denoted as $Q_{ST} = [X_{ST}Y_{ST}Z_{ST}]$, the single perspective interactive system is denoted as $Q_W = [X_W Y_W Z_W]$, and the interactive stereo system denoted is as $Q_{ST} = [X_Q Y_Q Z_Q]$. The rotation and translation between interactive systems are denoted as R_Q, T_Q and R_W, T_w. The human target trajectories

are examined in the thermographic scene concerning the trajectory trend of the respective target's feature points. The height and width of the target area are evaluated for a ratio to determine the first data about the trajectory target trend. The ratio variations are reliant on the target alignment through the center rotation. The height and width variations for the target's backward and forward motion concern the interactive system.

In Figure 3, the backward and forward motion calculations and the magnitude and direction vector are provided for every interactive system denoted as dotted lines. The variations of ratios from every interactive systems perspective are presented in Figure 3. Stereo vision-based identification combines features extracted from two-dimensional stereo images with reassembled three-dimensional object features to sense humans in an interactive setting. The interactive stereo system is already monitoring the target G_1, the computer endowed perspective sensor; thermographic camera altered its direction to the absorbed target as a portion of the human-computer interaction task. The human target coordinate system is represented by the direction vector U_1, width s_w, and the height g_w; the ratio $(g_{q,1}/s_{q,1})$ indicates the target G_1 moving toward a single
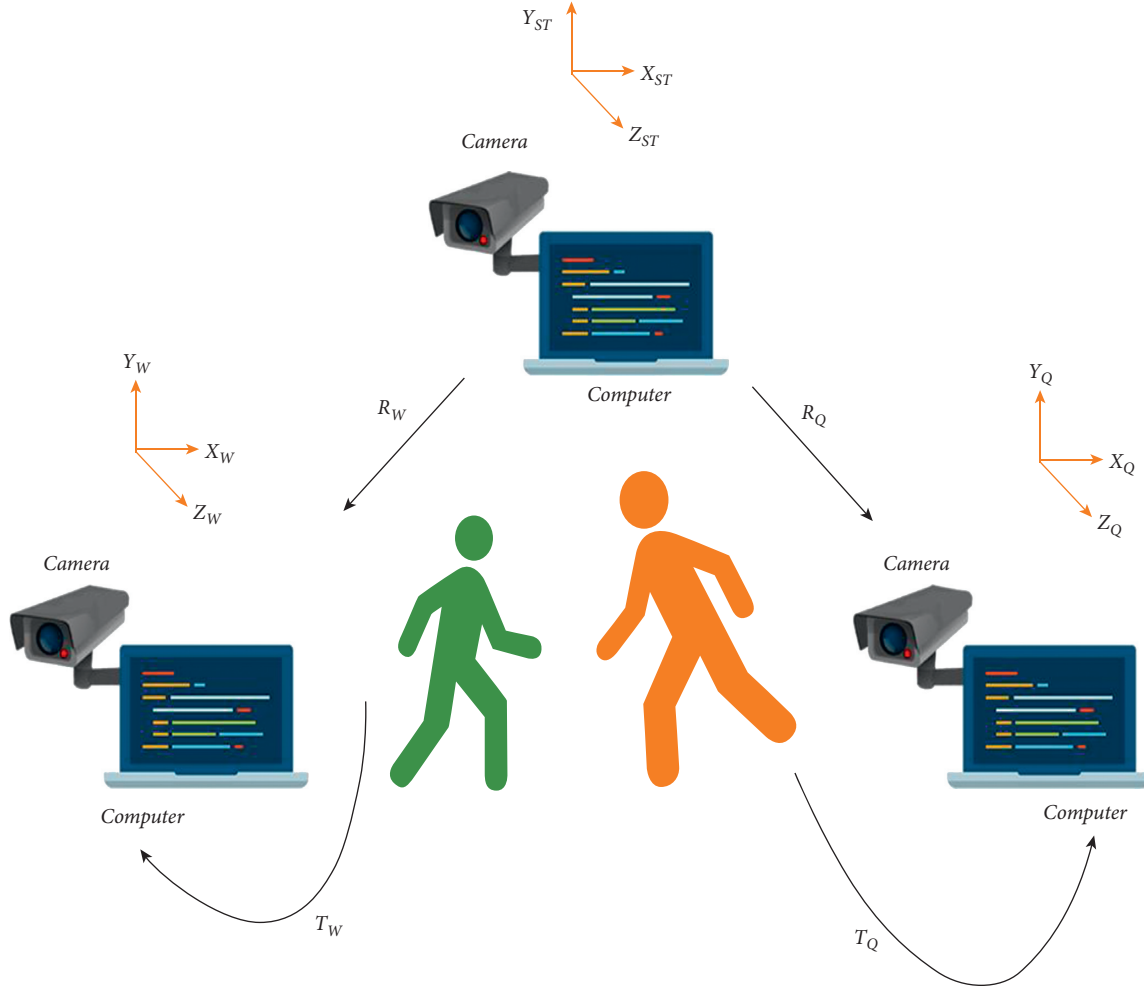
FIGURE 2: Three interactive systems equipped with a single thermal camera, stereo thermal sensor, and IR sensor. Rotation and translation between interactive systems are exposed with a corresponding spatial thermographic interactive system.

perspective interactive system. The ratio $(g_{w,1}/s_{w,1})$ indicates the target G_1 moving toward the interactive stereo system; c_s is the direction. The rotation and translation between interactive systems are denoted as R_Q, T_Q and R_W, T_W. These variations are noted during the movement of the target G_1 and updated for a time r by equation (2). To determine the method vector direction U_s, the height g_w and width s_w are utilized in the thermographic image.

$$U_s = \begin{cases} s_{w,r} - s_{w,r-1}, & \text{if} \left(\dfrac{g_w}{s_w}\right)_r - \left(\dfrac{g_w}{s_w}\right)_{r-1} = 0, \\ \\ 0, & \text{if} \left(\dfrac{g_w}{s_w}\right)_r - \left(\dfrac{g_w}{s_w}\right)_{r-1} \neq 0. \end{cases} \tag{2}$$

As shown in equation (2), s_w denotes rotation width, g_w denotes rotation height, and (g_w/s_w) denotes the rotation ratio. If there is no variance between successive proportions, the direction vector U_s is acquired from the variance between the successive target's width in successive pictures. If the proportion is varying, let us think that the target is creating a rotation around itself.

The horizontal target motion is utilized for right and left directions in its trajectory. Let us assume the overall extracted feature point for this particular target is F and the feature extracted point for a particular target is x_f. The horizontal motion feature point x_f is followed after the evaluation of their mean coordinates until the overall number of F. The final horizontal position average at the time $r - 1$ is deducted from the present mean. This variation supports the determination of the direction and horizontal vector magnitude c_w from the next expression.

Condition 1:

$$c_w = \frac{\left(\sum_{f=1}^{F} x_{f,r}\right)}{F}. \tag{3a}$$

Condition 2:

$$c_w = \frac{\left(\sum_{f=1}^{F} x_{f,r-1}\right)}{F}. \tag{3b}$$
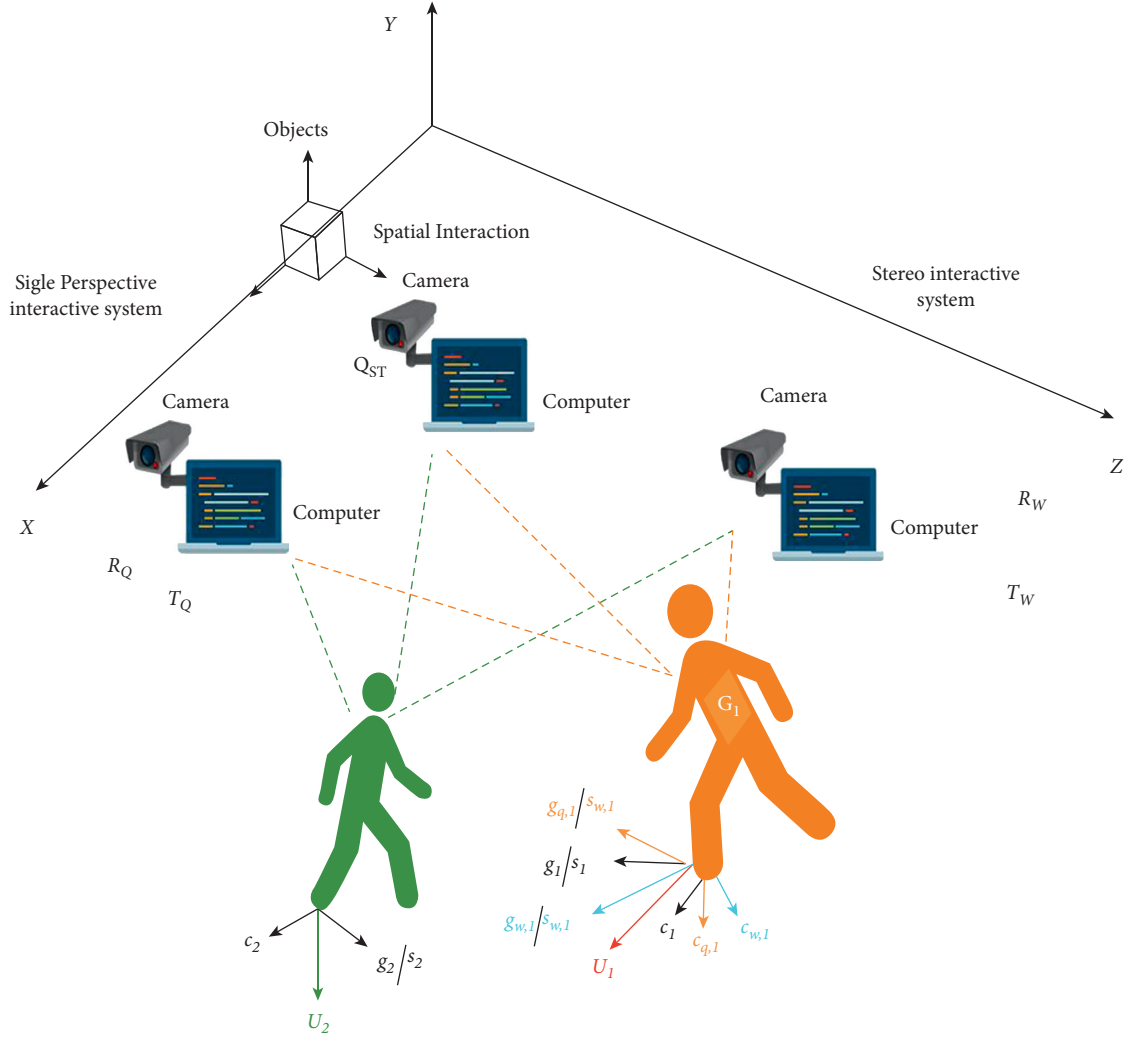
After computing

FIGURE 3: Interactive system identifying the trajectory vector for every target.

$$c_w = \frac{\left(\sum_{f=1}^{F} x_{f,r}\right)}{F} - \frac{\left(\sum_{f=1}^{F} x_{f,r-1}\right)}{F}, \tag{3c}$$

$$U_w = \sqrt{U_s^2 + c_w^2}. \tag{3d}$$

As discussed in equations (3a), (3b), and (3c) x_f denotes the extracted feature points for a particular target. F denotes the overall extracted features point for specific targets. The final horizontal position average at period $r - 1$ is deducted from the present mean r. c_w denotes the direction and horizontal vector magnitude. U_w denotes the final trajectory vector. A trajectory trend is produced for every target in the spatial thermographic picture, and it is equated with another interactive system trajectory assessment. Every interactive system assessment utilizes a horizontal and method vector to determine the last trajectory vector.

The target areas identified are labeled with two criteria for their characteristics. The human target areas are identified and bounded by a rectangle frame. The frame is split into $W = (Ga/m_{\text{ver}})(Sa/m_{\text{hor}})$ number of cells to examine every segment distinctly. The number of cells in the horizontal direction m_{hor} and the vertical direction is provided with m_{ver}. The human target is chosen utilizing associated elements and distinct from added substances in the target frame. The pixels in each cell are added by equation (4), and the overall human target field is determined for the respective cell. Let us consider that w is the respective cell, A_w is the overall cell pixel value, and G_a and S_a are the height and width of the target frame, respectively. Each row and the cell column are added to a bird's eyesight and perception target spectrum as an added target indistinguishable signature:

$$A_w = \sum_{x=A_x}^{x_k} \sum_{y=A_y}^{x_k} q(x, y). \tag{4}$$

As shown in equation (4), A_w is the overall cell pixel value. Each pixel in a cell is provided with the pixel q, and the

coordinates of this pixel in the cells are (x, y). The final pixel directs from the bottom and left of each cell can be evaluated from the subsequent expression:

$$(xy)_k = f\big((xy)_k\big) = \begin{cases} x_k, & \text{if } (m_{\text{hor}}) = A_x + w_x\left(\dfrac{S_a}{m_{\text{hor}}}\right), \\[2ex] y_k, & \text{if } (m_{\text{ver}}) = A_y + w_y\left(\dfrac{G_a}{m_{\text{ver}}}\right), \end{cases}$$
(5a)

$$x_k = A_x + w_x\left(\frac{S_a}{m_{\text{hor}}}\right) y_k = A_y + w_y.$$
(5b)

Equations (5a) and (5b) show the thermographic image final pixel coordinates of each cell, where the number of cells in the horizontal direction m_{hor} and the vertical direction is provided with m_{ver}. G_a and S_a are the target frame height and width, respectively. The beginning coordinates of every cell A_x and A_y are determined from the present directs of the target frame and frame height and width from the resulting expression:

$$(A)_{xy} = f\big((A)_{xy}\big) = \begin{cases} A_x, & \text{if } (m_{\text{hor}}) = \dfrac{S_a}{m_{\text{hor}}}(w_x - 1), \\[2ex] A_y, & \text{if } (m_{\text{ver}}) = \dfrac{G_a}{m_{\text{ver}}}(w_y - 1), \end{cases}$$
(6a)

$$A_x = \frac{S_a}{m_{\text{hor}}}(w_x - 1),$$
(6b)

$$A_y = \frac{G_a}{m_{\text{ver}}}(w_y - 1),$$
(6c)

As derived in equations (6a), (6b), and (6c), every cell index is evaluated from $w_y = [w - 1/m_{\text{ver}}] + 1$ utilizing floor function after division and $w_y = w - w_y m_{\text{hor}}$. After determining the kinematics for arm, leg, and head areas, corr chooses the relationship of these kinematics for each thermographic image of a similar target. Additional correlation outcome originates trajectories utilizing the human target trajectory directions $[X\ Y\ Z]$ with corr( ) to choose the last trajectory direction with the maximum association outcome. Human motion has been identified based on human body kinematics and target trajectory analysis; human activity has been identified accurately for the human-computer interaction process. With the collective thermographic images of pixel and edge variances, human motion detection has been performed. The spatial thermographic images have many noises, low image resolution, and low contrast to resolve these problems. AIMF algorithm has been proposed.

Case 2. EMG based human motion recognition.

Solution 2. A better human-computer interaction can be accomplished utilizing electromyographic (EMG) based human motion recognition. The signals of common human gestures have been gathered using the EMG acquisition device. Electrode's locations have been chosen in line with the musculoskeletal of these 5 muscles and definite by contractions of the muscle-specific, which involve physically repelled finger abduction and extension. EMG signals are bioelectrical reactions affected by muscle fiber movement during multiple muscle contractions.

Figure 4 displays the human-computer interaction procedure based on the EMG signal. Data acquisition, preprocessing, classification model, and feature extraction are the significant and main stages in human gesture classification. The gesture classification can identify discrete body gestures and cannot be utilized for control of the interactive system and, thus, continuous motion regression, which has assessments for more motion information. Compared to the EMG signal-based musculoskeletal model for gesture recognition, the mapping between EMG and angular acceleration, joint moment, joint angular velocity, and angle can be recognized. The commonly utilized feature can be primarily separated into a time-frequency domain feature and a frequency domain feature. The initial EMG signals gathered include a variety of noise signals, like electromagnetic noise, electrode noise, and pervasive noise, caused by an external environment and acquisition system. Therefore, reducing the noise of the real signal is necessary to generate actual and efficient information for the extraction of features.

As shown in Algorithm 1, the proposed AIMF algorithm with an adaptive time-frequency data analysis can be distinctly a time series into a finite number of elements, known as Empirical Mode Decomposition (EMD). For the self-adaptive decomposition, the data processing method is used for the nonstationary and nonlinear signals, where t is the acquisition channel, input $E(t)$ is the EMG signal, and output D_{IMF} is the intrinsic mode function. To reduce the noise in EMG signal, let us compute all the extreme points of $E(t)$; i.e., $K(t) = \{K_{\min}(j), K_{\max}(j)\}$. We fit the high points by interpolation approach: upper envelope represented as $K_{\max}(j) \longrightarrow O_{\max}(j)$ and lower envelope represented as $K_{\min}(j) \longrightarrow O_{\min}(j)$; subsequently, we compute average envelope value; i.e., $V(t) = (O_{\min}(j) + O_{\max}(j))/2$; after that, we obtain a stationary data sequence; $S(t) = E(t) - V(t)$. The algorithm decomposes the actual signal divides noise from the efficient signal with high time-frequency resolution and better adaptability in various intrinsic mode functions. Therefore, the collecting the features of the EMG signal, the AIMF algorithm is a perfect method for decreasing the EMG signal noise.

As shown in Algorithm 1, Empirical Mode Decomposition is executed on every EMG signal channel; a set of AIMF can be determined as

$$D(t) = \sum_i^m D_{\text{IMF}}(i) + \gamma_m.$$
(7)

As inferred from equation (7), γ_m denotes the residual of the actual EMG signal after extracting m AIMFs D. The data signal determined after noise reduction using Algorithm 1 guarantees the feature extraction of original motion signals.

> **Input:** the actual data sequence of the EMG signal, $E(t)$, $(t = 1, \ldots, m)$
> **Output:** intrinsic mode function, $D_{\text{IMF}}(i)$;
> **Repeat**
> **While** $j < L$ **do**
> **for** all j such as $1 \le j \le l$ **do**
> Compute $E(t)$ i.e., $K(t) = \{K_{\min}(j), K_{\max}(j)\}$;
> Obtain upper envelope: $K_{\max}(j) \longrightarrow O_{\max}(j)$,
> Obtain lower envelop: $K_{\min}(j) \longrightarrow O_{\min}(j)$;
> Compute average envelope value $V(t) = (O_{\min}(j) + O_{\max}(j))/2$;
> Obtain a stationary data sequence, $S(t) = E(t) - V(t)$;
>   Update $t = t + 1$;
> **Until** convergence

ALGORITHM 1: Average intrinsic mode function algorithm for electromyographic signal noise reduction.

Figure 5 shows the communication between interactive systems. Without loss of generalization, it is preassumed that the l th HCI system records data of K heterogeneous variable to be evaluated and I heterogeneous variable to interact, which produce a computation vector $c_l' = [c_{l,1}, \ldots, c_{w,I}']^T$ and communication vector $d_l' = [d_{l,1}', \ldots, d_{l,I}']^T$, where c_l' and d_l' are calculated values of the k th computation variable and i th communication variable at l th HCI system, respectively. As a reminder for computation, the interactive system engages itself in computing k th targeted function through the following equation:

$$p_k = f_k \left(\sum_{l=1}^{L} h_{l,k}(c_{l,k}) \right), \quad k = 1, \ldots, K. \tag{8}$$

As discussed in equation (8), p_k is the ideal computation output, and $h_{l,k}(\cdot)$ and $f_k(\cdot)$ denote the preprocessing function at the l th interaction system. Let us consider $d_l = [h_{l,k}(c_{l,k}), \ldots, h_{l,K}(c_{l,K})]^T$ indicating the preprocessed computation signal at the l th interaction system. For simplicity of exploration and without loss of generalization, let us preassume that the communication signals and computation signals are distributed with the unit norm; that is, $E\{d_l d_l^H\} = J$ and $E\{d_l' d_l'^H\} = 1$. Therefore, the l th interactive system builds the coded transmit EMG signal y_l as

$$y_l = S_l d_l + \sum_{i=1}^{I} u_{l,i} d_{l,i}'. \tag{9}$$

As inferred from equation (9), $S_l \in E^{N \times K}$ and $u_{l,i} \in E^{N \times 1}$ indicate the transmit beams for the communication and computation signal, respectively.

$$\|S_l\|_F^2 + \sum_{i=1}^{I} \|u_{l,i}\|^2 \le \frac{Z_l}{T/2}, \quad \forall l, i. \tag{10}$$

Thus, the received signal at the interactive system is expressed as

$$x = \sum_{l=1}^{L} H_l y_l + m = \underbrace{\sum_{l=1}^{L} H_l S_l d_l}_{\text{computation signal}} + \underbrace{\sum_{l=1}^{L} \sum_{i=1}^{I} H_l u_{l,i} d_{l,i}'}_{\text{communication signal}} + m. \tag{11}$$

As shown in equation (11), m is the noise vector with variances σ_m^2. Initially, the processing of the computation signals has been discussed. Because of the one-to-one mapping between $c = \sum_{l=1}^{L} c_l$ and $p = [p_1, p_2, \ldots, p_K]^T$ in equation (8), let us take a precise c at the interactive system as the targeted function signal. It is anticipated to execute a receive beam at the interactive approach to reduce the distortion of the targeted function signal affected by channel noise, fading, and interference. Therefore, the received signal for computation at the interactive system is expressed as

$$\hat{c} = V^L \sum_{l=1}^{L} H_l S_l d_l + V^L \left(\sum_{l=1}^{L} H_l \sum_{i=1}^{I} u_{l,i} d_{l,i}' + m \right). \tag{12}$$

As derived in equation (12), $V \in E^{M \times K}$ is a receive beam for computation outcomes at the interactive system. As a rule, the distortion of computation at the interactive system is calculated by the Root Mean Square Error (RMSE) between c and $\hat{c}$ which is stated as

$$\text{RMSE}(d, \hat{d}) = Z\{tr((\hat{d} - d)(\hat{d} - d)^H)\}. \tag{13}$$

Replacing (12) into (13), the computation distortion can be calculated as the resulting RMSE function of transmitting and receiving beam:

$$\text{RMSE}(V, S_l, u_{l,i}) = \sum_{l=1}^{L} \|V^H H_l S_l - J\|_F^2 + \sigma_m^2 \|V\|_F^2 + \sum_{l=1}^{L} \sum_{i=1}^{I} \|V^H H_l u_{l,i}\|^2. \tag{14}$$

Subsequently, communication signal processing has been discussed. The received EMG signal for communication at the interactive system can be evaluated as

$$x_{l,i}' = q_{l,i}^H H_l u_{l,i} d_{l,i}' + q_{l,i}^H \sum_{j=1,\, j \neq l}^{L} H_j \sum_{n=1,\, n \neq i}^{I} q_{j,n} d_{j,n}'$$
$$+ q_{l,i}^H \sum_{j=1}^{L} H_j S_j d_j + q_{l,i}^H m. \tag{15}$$

As inferred from equation (15), $q_{l,i} \in E^{M \times 1}$ indicates the receive beam vector for communication signal $c_{l,i}'$ at the interactive system. As a result, the received EMG signal to inference and noise proportion at the communication receiver system can be calculated using

$$\zeta_{l,i} = \frac{\left| q_{l,i}^H H_l u_{l,i} \right|^2}{\sum_{j=1,\, j \neq l}^{L} \sum_{n=1,\, m \neq i}^{I} \left| q_{l,i}^H H_j u_{j,n} \right|^2 + \sum_{j=1}^{L} \left\| q_{l,i}^H H_j S_j \right\|^2 + \sigma_m^2 \| q_{l,i} \|^2}. \tag{16}$$

Besides, the transmit signal of EMG relies on the energy beam sent by the human-computer interaction system.

EMG signal-based human motion recognition, Lyapunov exponent, is utilized to detect the numerical features in a complex system and denoted the system sensitivity to the first value as the parameter progress with a period t. The M-dimensional systems have M Lyapunov exponents, creating an exponential spectrum. Thus, it is broadly utilized in system fault diagnosis along with muscle activity and muscle contraction identification. It is stated as

$$\lambda_{\max} = \frac{1}{\Delta t} \sum_{p=1}^{N} \frac{L(t_p)}{L(t_{p-1})}. \tag{17}$$

As discussed in equation (17), $L(t_p)$ denotes the distance between two adjacent 0 points at the time t_p, Δt is the sampling time, and N indicates the overall step length. The distance between end-to-end paths is generally reproduced by the forecast fault on the log function to attain the Lyapunov exponent of the complete set of IMF. It is stated as

$$k(p) = \frac{1}{Mt} \sum_{m=1}^{M} 1a \frac{\Gamma_m(p)}{\Delta m}. \tag{18}$$

As shown in equation (18), Δm denotes the distance between phase points, E_{m+1} is the adjacent point to E_m, and $\Gamma_m(p)$ is the distance between E_m and E_{m+1} after p convolution step length time. M indicates the aggregate number of phase points. The two typical EMG signal features embedded in dimension and delay time are two essential variables for evaluating the Lyapunov exponent. The motion characteristics of the EMG signal have been chosen and represented using these two parameters. The delay time was calculated using the mutual information technique as follows:

$$B(\tau) = k(Y_m, Y_{m+\tau}) \ln \frac{k(Y_m, Y_{m+\tau})}{k(Y_m)k(Y_{m+\tau})}. \tag{19}$$

As shown in Figure 6 and equation (19) representing the delay time calculation, $B(\tau)$ denotes the EMG signal delay time and $k(Y_m)$, $k(Y_{m+\tau})$, and $k(Y_m, Y_{m+\tau})$ are likelihood values.

By Algorithm 1, the embedded dimension has been evaluated, which is stated as

$$O_1(n) = \frac{O(n+1)}{O(n)}. \tag{20}$$

As discussed in equation (20), $O(n+1)$ and $O(n)$ are the mean thresholds of the distance between every two adjacent neighbors in the recreated $n+1$ dimensional space and n-dimensional space, respectively.

The proposed AIMF algorithm reduces the EMG signal noise and obtains active human motion features for the human-computer interaction. Finally, the proposed Optimized Noninvasive Human-Computer Interaction (ONIHCI) model addresses the problems such as accurate gesture recognition and human target identification and reduces the noise in EMG signal for an effective HCI process. The following section briefly describes the experimental results.

4. Experimental Results and Discussion

The proposed Optimized Noninvasive Human-Computer Interaction (ONIHCI) model experimental results have been performed in a computer vision-based human-computer interaction environment. Different situations have been used analyzed using the training and testing dataset [21]. From the large dataset, a 70 : 30 ratio of training and testing data has been formed. A different dataset was randomly chosen among the testing data and observed various performance metric values. The following figure 7 and table 1 give the results observed from this analysis. Human targets are running, walking, and slow waving; the data is divided into multiperson and single-person behavior. The analysis of human trajectory and human kinematics has been discussed in this section. The human targets are detected in multiple thermographic images.

Furthermore, EMG-driven musculoskeletal model-based motion recognition has been utilized, to the map between EMG and joint angular velocity, angle, joint moment, or angular acceleration. The experimental results show that the HCI suggested method achieves lesser noise and enhances the accuracy in human motion recognition and identifying human targets with high performance. Simulation parameter has been used for the human motion recognition based on EMG signal acquisition with a simulation time of 4 seconds, and sampling frequency utilized is 25 kHz, muscle length is 200 mm, the number of electrodes used is 64, muscle radius is 20 mm, muscle fiber length is 45 mm, muscle fiber diameter is 35.77, and fiber length is 150 mm.

4.1. Root Mean Square Error (RMSE) Ratio Analysis. The Root Mean Square Error is often used in continuous motion prediction. It compares the actual values to the projected values. The conflict is squared to prevent canceling negative
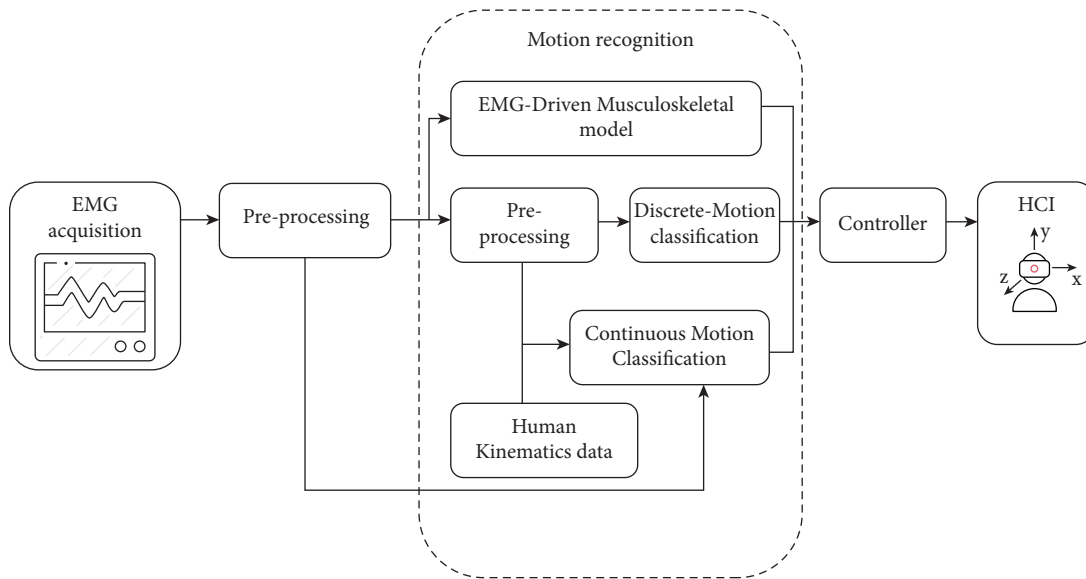
FIGURE 4: The process of EMG signal-based human-computer interaction.
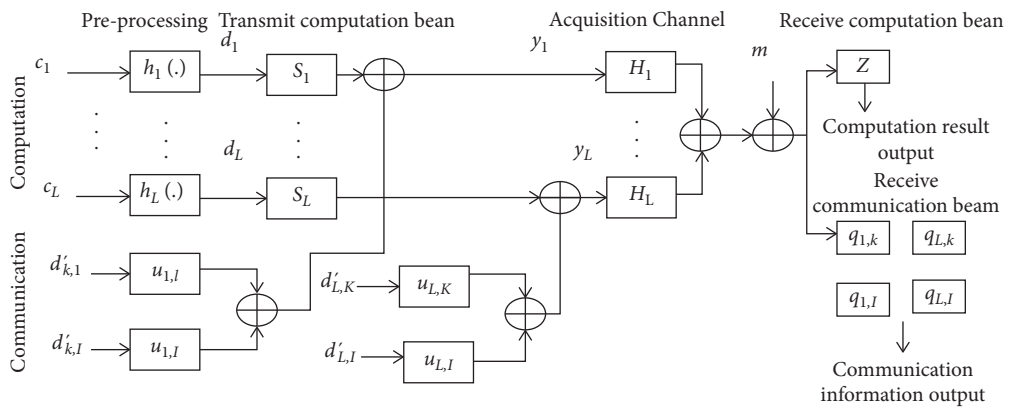

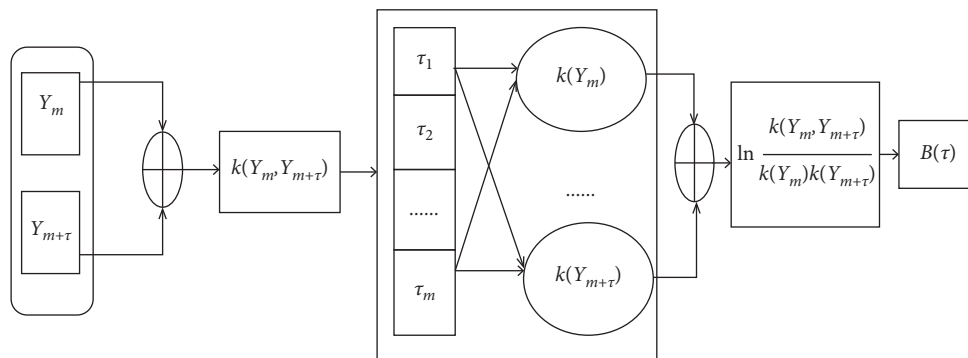FIGURE 5: Communication between interactive systems.
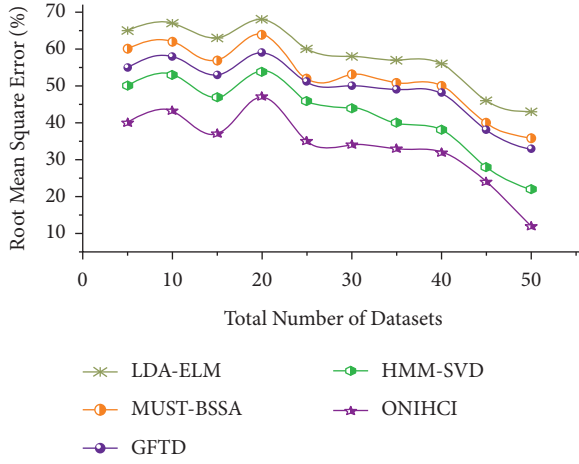

FIGURE 6: Delay time calculation.

FIGURE 7: Mean square error ratio.

and positive values. The distortion of the computation at the interactive system is evaluated by the mean square error d and $\widehat{d}$ which is stated as

$$\text{RMSE}(d, \widehat{d}) = Z\{tr((\widehat{d} - d)(\widehat{d} - d)^H)\}. \quad (21)$$

Replacing (12) into (13), the computation distortion can be calculated as the following RMSE function of receive and transmit beam:

$$\text{RMSE}(V, S_l, u_{l,i}) = \sum_{l=1}^{L} \|V^H H_l S_l - J\|_F^2 + \sigma_m^2 \|V\|_F^2 \quad (22)$$
$$+ \sum_{l=1}^{L} \sum_{i=1}^{I} \|V^H H_l u_{l,i}\|^2.$$

Therefore, the computation outcome should have a low RMSE. Figure 7 demonstrates the RMSE using the suggested system.

The performance ratio evaluation of the recommended ONIHCI system is shown in Table 1. To detect the movement of a human, it is necessary to track the body's motion throughout physical activity. Data processing to better portray such movement aids in the operation's detection, which helps this study effectively. Compared to other technologies, the human-computer interface performance based on human motion recognition findings is quite successful. Table 1 shows that the highest performance ratio of 95.4% by the proposed model outperforms the existing models in the literature survey. For all sets of data inputs, the ONIHCI gives the highest performance ratio.

4.2. Recognition Accuracy Ratio Analysis.
The human thermographic view is examined for comprehending the trajectory behaviors and the motion kinematics during the target's movement. Humans are initialized using the kinematics system by features like the number of limbs, degrees of freedom, limb length, etc. Humans are represented as images, and personality traits like form or region are extracted and stored in the image model. The acknowledgment, as well as the complexity of the motion, can be

determined effectively. The accuracy from every sensor is provided in Table 2 concerning every sensor's kinematics and trajectory analysis utilizing thermographic images. The simulation outcomes demonstrated that the suggested approach could assess the human target angle with high accuracy compared to other existing approaches. Figure 8 demonstrates the recognition accuracy ratio using proposed ONIHCI methods.

4.3. Delay Time Determination and Noise Reduction Ratio Analysis.
In reflection of the nonlinear dynamic EMG signal, the actual EMG signal has been decomposed into a set of IMF when noise reduction, as demonstrated in Figure 9(b). The actual EMG signal $X(t)$ has been decomposed into $D_{\text{IMF}}(i)$, $(i = 1, 2, \ldots, 6)$ and deviation γ_m. $D_{\text{IMF}}(i)$ is an oscillation function with various frequencies and amplitudes. γ_m is a monotonic signal, denoting the drift element determined by deducting every $D_{\text{IMF}}(i)$ from the actual signal, and $X(t)$ no longer meets the decomposition states. Embedded dimension and delay time are two significant constraints for the manipulative Lyapunov exponent. Moreover, the chosen delay time is too short and not advantageous to EMG signal optimization.

The two typical EMG signal features embedded in dimension and delay time are essential for evaluating the Lyapunov exponent. These two parameters are appropriate for depicting EMG gesture time sequence data; the motion features of the EMG signal have been selected and represented.

The mutual information approach has been utilized for calculating delay time which is stated as

$$B(\tau) = k(Y_m, Y_{m+\tau}) \ln \frac{k(Y_m, Y_{m+\tau})}{k(Y_m) k(Y_{m+\tau})}. \quad (23)$$

As shown in the above equation, $k(Y_m)$, $k(Y_{m+\tau})$, and $k(Y_m, Y_{m+\tau})$ are likelihood values.

Based on Algorithm 1, the embedded dimension has been evaluated, which is stated as

$$O_1(n) = \frac{O(n+1)}{O(n)}. \quad (24)$$

As discussed in the above equation, $O(n+1)$ and $O(n)$ are the mean thresholds of the distance between every two adjacent neighbors in the recreated $n + 1$-dimensional space and n-dimensional space, respectively. Figure 9(a) shows the delay time using the proposed ONIHCI method.

The quality of the EMG signal measurement is demonstrated by the ratio of the EMG signal calculated to unwanted environmental noise inputs. A high-quality signal offers more information to predict the intention so that it increases prediction accuracy. Nevertheless, noises from various sources are possible, and the analysis of EMG signals may be contaminated. To maximize the signal-to-noise relation in this respect, amplifiers are designed and used to reject or eliminate noises. The accuracy of EMG signals is affected by noise and artifacts from various causes (including electric devices, power lines, and physiological factors),

TABLE 1: Performance ratio evaluation.

No. of datasets	LDA-ELM [22]	MUST-BSSA [32]	GFTD [25]	HMM-SVD [26]	ONIHCI (proposed approach)
5	54.2	56.1	67.1	77.7	80.1
10	51.3	53.3	60.4	67.8	63.5
15	63.9	56.5	78.4	87.9	52.4
20	76.2	77.6	78.5	73.7	72.3
25	83.1	54.7	56.9	76.5	77.1
30	54.7	50.8	53.8	57.4	81.2
35	43.9	44.4	46.7	47.3	82.2
40	51.8	54.3	56.7	60.2	89.2
45	65.3	65.2	76.6	77.1	90.3
50	67.2	66.1	77.5	87.2	95.4

TABLE 2: Precision ratio analysis.

No. of datasets	LDA-ELM	MUST-BSSA	GFTD	HMM-SVD	ONIHCI
5	32.1	34.1	35.7	37.7	70.1
10	53.3	54.3	69.4	67.8	61.5
15	64.9	54.5	79.4	87.9	42.4
20	77.2	72.6	72.5	73.7	72.3
25	82.1	44.7	53.9	76.5	77.1
30	53.7	42.8	54.8	57.4	79.2
35	42.9	49.4	45.7	47.3	82.2
40	51.8	59.3	56.7	60.2	77.2
45	64.3	67.2	76.6	77.1	92.3
50	68.2	62.1	77.5	87.2	96.4
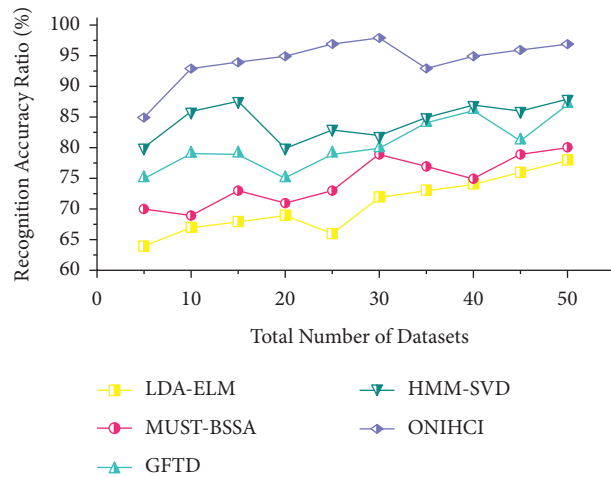
FIGURE 8: Recognition accuracy ratio analysis.

which may contribute to the inaccurate analysis of data or a misunderstanding of motion parameters. In the proposed ONIHCI method, the AIMF algorithm reduces the noise level seen in the raw EMG signal. Figure 9(b) shows the noise reduction ratio using the proposed method.

Table 2 shows the precision ratio using the proposed ONIHCI method. The present system implemented in this study is user-friendly compared to a command-based system or standard device and robust in detection and recognition. The proposed AIMF algorithm is extensively utilized in the classification and regression of HCI because of its simple execution, high precision, and antinoise capability.

4.4. Normalized Computation Error Evaluation.
Normalized computation error is a statistical assessment utilized to compare proficiency testing outcomes where the uncertainty in the measurement outcomes is included. The cause for error with the spatial thermographic sensor has the minimal image size of the targets and unexpected changes. Single perspective and stereo sensors provided a benefit to decreasing the computation error with surplus views. The distance between the human target and interactive system positions at every degree of target force direction is averaged over a tracking cycle known as normalized tracking error. The proposed method has lesser

Figure 9: (a) Delay time determination. (b) Noise reduction ratio analysis.

Figure 10: Normalized computation error evaluation.

singular value decomposition coefficients," in *Advances in Computational Intelligence and Communication Technology*, pp. 549–561, Springer, Singapore, 2021.

[21] https://www.rami-khushaba.com/electromyogram-emg-repository.html.

[22] J. Qi, G. Jiang, G. Li, Y. Sun, and B. Tao, "Intelligent human-computer interaction based on surface EMG gesture recognition," *IEEE Access*, vol. 7, pp. 61378–61387, 2019.

[23] A. Aggarwal, V. Sharma, A. Trivedi et al., "Two-way feature extraction using sequential and multimodal approach for hateful meme classification," *Complexity*, vol. 2021, Article ID 5510253, 7 pages, 2021.

[24] D. Singh and V. Kumar, "Image dehazing using Moore neighborhood-based gradient profile prior," *Signal Processing: Image Communication*, vol. 70, pp. 131–144, 2019.

[25] E. Song, H. Lee, J. Choi, and S. Lee, "AHD: thermal image-based adaptive hand detection for enhanced tracking system," *IEEE access*, vol. 6, pp. 12156–12166, 2018.

[26] F. Xiao, D. Yang, Z. Lv, X. Guo, Z. Liu, and Y. Wang, "Classification of hand movements using variational mode decomposition and composite permutation entropy index with surface electromyogram signals," *Future Generation Computer Systems*, vol. 110, pp. 1023–1036, 2020.

[27] D. Singh and V. Kumar, "A comprehensive review of computational dehazing techniques," *Archives of Computational Methods in Engineering*, vol. 26, no. 5, pp. 1395–1413, 2019.

[28] M. Agrawal, A. U. Khan, and P. K. Shukla, "stock price prediction using technical indicators:A predictive model using optimal deep learning," *International Journal of Recent Technology and Engineering*, vol. 8, no. 2, pp. 2297–2305, July 2019.

[29] P. Tiwari and P. Shukla, "Artificial neural network-based crop yield prediction using NDVI, SPI, VCI feature vectors," in *Information and Communication Technology for Sustainable Development*, M. Tuba, S. Akashe, and A. Joshi, Eds., vol. 933, Advances in Intelligent Systems and Computing, Springer, Singapore, 2020.

[30] X. Tang, F. Li, T. G. Seetharam, and C. C. Vignesh, "Internet of things-assisted intelligent monitoring model to analyze the physical health condition," *Technology and Health Care*, 2021.

[31] E. Benli, Y. Motai, and J. Rogers, "Visual perception for multiple human–robot interaction from motion behavior," *IEEE Systems Journal*, vol. 14, no. 2, pp. 2937–2948, 2019.

[32] C. Chen, Y. Yu, S. Ma et al., "Hand gesture recognition based on motor unit spike trains decoded from high-density electromyography," *Biomedical Signal Processing and Control*, vol. 55, Article ID 101637, 2020.

[33] R. Bhatt, P. Maheshwary, P. Shukla, P. Shukla, M. Shrivastava, and S. Changlani, "Implementation of fruit fly optimization algorithm (FFOA) to escalate the attacking efficiency of node capture attack in wireless sensor networks (WSN)," *Computer Communications*, vol. 149, pp. 134–145, 2020.