

The Rehabilitation Exoskeletons Based on Intention Recognition

Tongtian Ying

College of Engineering, Shanghai Ocean University, Shanghai, China

binah911iron@gmail.com

Abstract. With the global aging population and the increasing number of stroke and gout patients, the number of individuals with mobility impairments is rising. Rehabilitation exoskeleton, as an emerging human-machine fusion assistive technology, has been widely used in recovery training for patients with neurological injuries. Motion intention recognition is the core component for achieving human-machine interaction, with its accuracy directly determining the safety and effectiveness of rehabilitation training. Driven by advancements in artificial intelligence and multimodal sensing technologies, single-modality motion intention recognition is evolving toward multimodal fusion. This paper systematically reviews the development of intent recognition for rehabilitation exoskeletons, highlighting the limitations of single-modality approaches such as electromyography and electroencephalography. Consequently, researchers have shifted focus to multimodal intent recognition. The current state of multimodal intent recognition is presented, encompassing sensory modalities, feature extraction and fusion methods, and representative application cases. It also summarizes the current difficulties and challenges in development, along with potential future trends, aiming to outline developmental trajectories and provide reference and insights for subsequent research.

Keywords: Rehabilitation exoskeletons, multimodal intent recognition, unimodal intent recognition.

1. Introduction

Population aging is regarded as a global issue. According to World Health Organization statistics, the number of people aged 60 and above surpassed that of children aged 5 and under in 2020. Suggesting that by 2050, the global population aged 60 and above will reach 2.1 billion. While patients with physical disability and frailty have also become social problems in many countries, from 1990 to 2019, the absolute number of stroke cases increased by 70.0%, the number of people affected rose by 85.0%, stroke-related deaths increased by 43.0%, and stroke-attributable DALYs increased by 32.0% [1]. With the aging problem and the concerns about patients with chronic diseases, the demand for therapy for movement disorders has significantly grown. Traditional rehabilitation relies on experienced therapists, but the intensity and repeatability of the training cannot be guaranteed, so the outcomes are worrisome. In recent years, exoskeleton robots have gradually entered the field of rehabilitation. A rehabilitation exoskeleton is a mechanical device that can enhance recovery effects [2]. These devices assist with repetitive movements while also collecting patients' physiological and movement data, thereby providing valuable insights into their needs and supporting personalized care strategies [3]. As a result, the development of Rehabilitation exoskeletons has begun to receive attention among scholars.

The accuracy of intent recognition determines the efficiency and safety of the rehabilitation experience. This paper summarizes the basic modalities applied in intent recognition. Furthermore, it lists current challenges such as signal variability, time delay, and real-time processing, and proposes future research directions and potential solutions aimed at improving the accuracy and robustness.

The essence of intent recognition lies in accurately perceiving human movement intentions to achieve natural and precise human-robot interaction. In rehabilitation exoskeletons, intent recognition is the essential part that enables devices to promote both safety and effectiveness. With the development of advanced sensing technologies, researchers have proposed various approaches to improve the accuracy and adaptability of intent recognition in medical and rehabilitation [4].

2. Unimodal Motion Intention Recognition

2.1. Muscle Signals

Since the electromyographic signal was recorded in 1922 by Professor Joseph Erlanger, Electromyography (EMG) technology has been widely applied in clinical medical diagnosis [5]. Japan created the first exoskeletons equipped with this technology. Besides, in 2023, China also deployed exoskeletons equipped with electromyographic signals in hospitals, and patients are highly satisfied with the excellent rehabilitation effects [6]. However, the electromyographic signal also has several disadvantages; it can be distracted by external influences, so it needs special processing methods to avoid interference coming from muscle movement, electrical noise, and environmental noise. In addition to these, when muscles become exhausted, the electromyographic signals may be difficult to correctly record, leading to errors in recognizing motion intentions.

2.2. Neural Signals

Neural signals represent the earliest stage of motor intention generation, as they are directly associated with cortical activity responsible for planning and initiating movement. By capturing these brain signals, rehabilitation exoskeletons can predict user intentions before muscular activation, enabling more natural and proactive human–machine interaction.

EEG technology has gained great attention in recent years, and it has a significant effect on scanning personal intent, leading to various applications in brain-computer interface and human-computer interaction [8]. It records electrical activity from the scalp with millisecond-level temporal resolution, and it has been extensively studied as a biomarker for motor intention recognition [7]. EEG's high temporal precision allows pre-movement detection, making it particularly suitable for exoskeletons that require anticipatory responses. However, EEG signals are often contaminated by artifacts (e.g., eye blinks, muscle activity, environmental noise), and long-term use requires advanced signal processing and adaptive algorithms. To address this issue, Chen et al. proposed an EEG-based adaptive algorithm for detecting motor intentions [8]. In addition, researchers have introduced a multi-task deep recurrent neural network that can improve the accuracy of intention recognition by scanning more EEG signals of different frequencies [9].

2.3. Deep Learning

Deep learning could process complex information and data, which is hard for traditional machine learning methods to process. Recent progress in deep learning has significantly advanced motion intention recognition. Convolutional neural networks (CNNs) effectively capture patterns in EMG and EEG signals, while recurrent neural networks (RNNs) and long short-term memory (LSTM) models are particularly suited for time-series decoding. Such models have demonstrated real-time decoding of user intentions even under non-stationary and noisy conditions [8].

3. Multimodal Motion Intention Recognition

With the rise of the complexity of human-robot interaction scenes, unimodal motion intention recognition can't meet the demand for practical scenes. Research based on unimodal intention recognition overlooks the complexity of real-world multimodal scenarios. For example, conclusions drawn solely from textual information may differ significantly from those reached after incorporating other modalities [3].

And as a result, scholars start to focus on multimodal intention recognition, which combines many different sources of information such as vision, EMG signals, EEG signals, and so on, so multimodal intention recognition can capture more accurate and total intention of the users, increasing robustness.

Zhao et al. classify studies in multimodal intent recognition, and they proposed a processing workflow comprising three stages: feature extraction, multimodal representation learning, and intent

classification. Representation learning is the core area of multimodal modeling, where different modalities are aligned, fused, and jointly encoded to form a unified representation [10].

3.1. Feature Extraction

Feature extraction is the core component of multimodal intent recognition, aiming to transform raw high-dimensional signals into discriminative low-dimensional representations for efficient identification of motion intentions by subsequent classification or regression models. Signals from different modalities exhibit distinct characteristics in the time domain, frequency domain, and spatial structure, necessitating tailored feature extraction methods.

With the advancement of computational capabilities and the proliferation of large-scale datasets, deep learning has become the predominant paradigm for multimodal intention recognition. It enables hierarchical feature abstraction from raw physiological and kinematic signals through end-to-end optimization, thereby reducing manual intervention and improving accuracy.

3.2. Multi-Task Coordination Methods

In practical applications, unimodal intent recognition is often unable to meet demand. The multi-task learning framework enhances overall performance and improves model generalization by sharing parameters across multiple related tasks within a single model. Multi-task coordination methods solve the obstacles of single-task learning, enhancing robustness in multimodal settings [10]. Illustrated in Fig. 1.

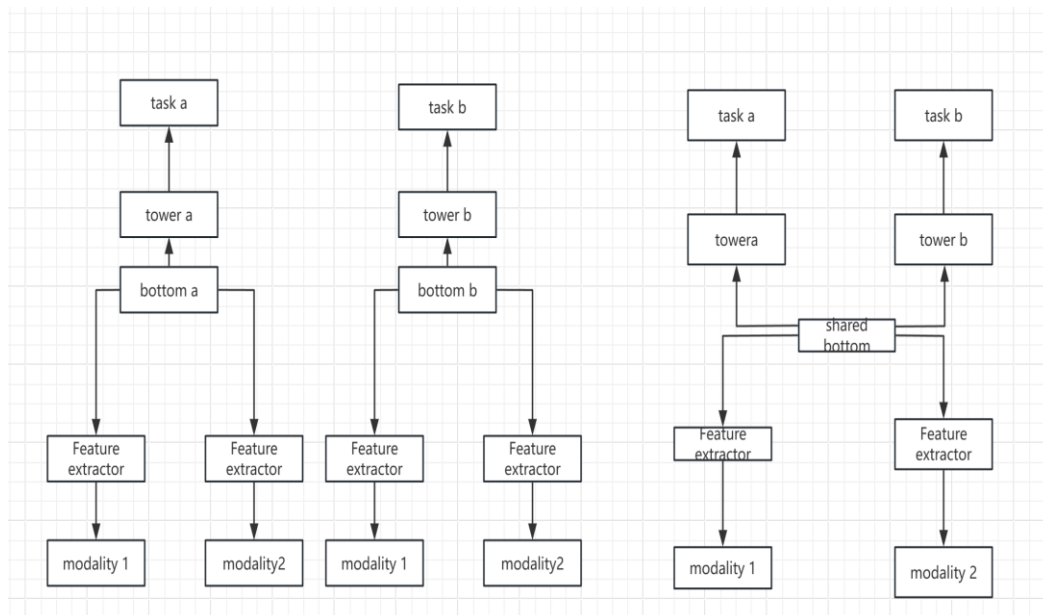


Fig. 1 Flowchart of Multi-Task Coordination Methods [10]

3.3. Application

Wang et al. designed an exoskeleton framework that utilizes muscle signals and neural signals for intent recognition [10]. It can address the muscle fatigue caused by relying solely on muscle signals for intent recognition. This framework combines continuous brain stimulation from rehabilitation devices, guides patients to imagine limb movements, and promotes neural recovery in the brain. This fosters the formation of new neural circuits around damaged brain regions, recovering patients' motor functions.

Chen et al. designed a multimodal intent recognition scheme based on machine vision and hearing [11]. Using PocketSphinx speech recognition technology, Kinect visual data acquisition, and yolov3 algorithm to track the position of the hand exoskeleton and the grasped object in real time. They decrease the loss function to 0.03 and raise the confidence level to over 90% [11]. Assist stroke patients in rehabilitation training while incorporating appropriate daily activities.

Langlois et al. analyzed that integrating pressure sensors and electromyography sensors into the interface enhances classification accuracy for upper limb lifting and extension tasks, achieving a maximum classification accuracy of 96.8% across all subjects and trials [12]. Based on this, an integrated multimodal interface for upper limb exoskeletons is proposed [12].

4. Challenges and Future Directions

Despite the great progress in intention recognition that has been made in recent years, many obstacles remain. Humans' intentions are often vague and ambiguous, which makes intention recognition hard to infer accurately. User intentions are often dynamic [9]. Words of praise can also be used to attack or mock others. Meanwhile, modal fusion and model training still face challenges. For example, many intention categories are underrepresented in current training data. System delay and real-time performance are also of concern. At present, many multimodal intention recognition exoskeletons have not been widely put into practice and are still in the laboratory stage.

In the future, it will likely focus on multimodal fusion, personalization, and large-scale clinical deployment to realize the full potential of rehabilitation exoskeletons in improving the quality of life for patients worldwide.

5. Conclusions

This article systematically explained the importance of intention recognition. Intention recognition is the cornerstone of intelligent rehabilitation exoskeletons, bridging human neural signals with robotic control. The next parts illustrate the obstacles it faces when using unimodal intention recognition that which can't correctly interpret intentions in complex environments development of intention recognition, which evolves from unimodal intention recognition to multimodal intention recognition. Introducing feature extraction, fusion methods, multi-task coordination, specific application cases for multimodal intent recognition, and current applications in rehabilitation exoskeletons. Listing challenges such as multi-intent and compositional intent recognition, out-of-domain intent detection, and clinical validation remain to be addressed. Future efforts may focus on multimodal sensor fusion for robust and adaptive perception, personalized and adaptive learning algorithms that evolve with user progress, and integration with virtual reality and digital twins for immersive and data-driven rehabilitation.

References

- [1] Cross M, Ong K L, Culbreth G T, Steinmetz J D, Cousin E, Lenox H, ... Woolf A D. Global, regional, and national burden of gout, 1990–2020, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. *The Lancet Rheumatology*, 2024, 6: e507-e517.
- [2] Bogue R. Robotic exoskeletons: a review of recent progress. *Industrial Robot: An International Journal*, 2015, 42: 5-10.
- [3] Pei S, Wang J, Tian C, Li X, Guo B, Guo J, Yao Y. Assist-as-Needed Controller of a Rehabilitation Exoskeleton for Upper-Limb Natural Movements. *Applied Sciences*, 2025.
- [4] Huang X, Ma T, Jia L, Zhang Y, Rong H, Alnabhan N. An effective multimodal representation and fusion method for multimodal intent recognition. *Neurocomputing*, 2023, 548: 126373.
- [5] Zhang X, Qu Y, Zhang G, Wang Z, Chen C, Xu X. Review of sEMG for Exoskeleton Robots: Motion Intention Recognition Techniques and Applications. *Sensors*, 2025.
- [6] Zhang W, Li J, Gao F, Yang D, Guo Y. Research progress of wearable technology in rehabilitation medicine. *Chin J Rehabil Theory Pract*, 2017.
- [7] Al-Quraishi M S, Elamvazuthi I, Daud S A, Parasuraman S, Borboni A. EEG-based control for upper and lower limb exoskeletons and prostheses: a systematic review. *Sensors*, 2018, 18: 3342.

- [8] Jiang Z, Liu Y. A review of human action recognition based on deep learning. 9th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), 2024, 9.
- [9] Chen W, Wang S, Zhang X, Yao L, Yue L, Qian B, Li X. EEG-based motion intention recognition via multi-task RNNs. In: Proceedings of the 2018 SIAM International Conference on Data Mining. SIAM, 2018: 279–287.
- [10] Zhao J, Wen Y, Li Q, Hu M, Zhou Y, Xue J, ... Li Y. Deep Learning Approaches for Multimodal Intent Recognition: A Survey. arXiv preprint arXiv:2507.22934, 2025.
- [11] Wang Y, Fang Y, Huang W, Liu Z, Zhao C, Du M, Zhu L. Development of lower limb exoskeleton rehabilitation robot framework based on multi-modal motion intent detection. IEEE Access, 2025.
- [12] Langlois K, Geeroms J, Van De Velde G, Rodriguez-Guerrero C, Verstraten T, Vanderborght B, Lefeber D. Improved motion classification with an integrated multimodal exoskeleton interface. Frontiers in Neurorobotics, 2021, 15: 693110.