WEP Warwick Evans Publishing

Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 9
AIDML 2025

Modeling and Multi Source Signal Processing of High Precision Perception Sensing System in Intelligent Robots

Sitong Man *

School of energy and power engineering, Xi'an Jiaotong University, Xi'an, China

* Corresponding Author Email: stoneman1205@stu.xjtu.edu.cn

Abstract. With the development of intelligent robot technology, the demand for high-precision perception of complex environment is increasing. This article focuses on the modeling of high-precision sensing system, aiming at improving the accuracy and adaptability of robot environmental perception. Compared with the traditional modeling method based on physical characteristics, the data-driven modeling method is emphasized. In this article, a fusion algorithm of convolutional neural network (CNN) and long-term memory network (LSTM) is used. CNN is responsible for processing visual signals, extracting local features through convolution kernel sliding, and constructing feature hierarchy through multi-layer convolution and pooling. LSTM processes auditory and other time series signals, and solves the long-term dependence problem with special gating structure. Experiments show that data-driven modeling is lower than traditional methods in average error and maximum error, and the fusion algorithm of CNN and LSTM improves the multi-source signal processing effect, and the accuracy of target recognition reaches 88%. This verifies the effectiveness of the proposed modeling and multi-source signal processing strategy, and provides strong support for robot complex environment perception and decision-making.

Keywords: Intelligent robot, High-precision sensing system, Multi-source signal processing, Sensor technology.

1. Introduction

In recent years, with the rapid development of science and technology, intelligent robots have been widely used in industrial production, medical services, smart homes and many other fields [1]. The performance of intelligent robot depends largely on the accuracy of its sensing system. High-precision sensing system can make intelligent robot more accurately perceive the surrounding environment and make reasonable decisions, thus improving work efficiency and quality [2]. At present, intelligent robots are faced with complex and changeable working scenes, which puts forward higher requirements for high-precision sensing systems [3]. A single sensor often has limitations, and it is difficult to obtain environmental information comprehensively and accurately; The mutual interference between multi-source sensor signals and how to effectively fuse them have become the key problems to be solved urgently [4]. Therefore, it is of great practical significance to study the modeling and multi-source signal processing of high-precision sensing system in intelligent robots.

In the modeling of high-precision sensing system, accurate model helps to reveal the working principle and performance characteristics of the sensor, and provides theoretical support for optimizing sensor design and improving sensing accuracy [5]. Through the modeling method based on physical characteristics or data-driven, it is expected to build a more accurate perception model [6]. Multi-source signal processing aims to solve the problems of feature extraction, fusion and anti-interference of different sensor signals, so that intelligent robots can obtain effective information from complex multi-source signals.

This article will focus on the modeling and multi-source signal processing of intelligent robot's high-precision sensing system. Through the analysis and summary of the existing technology, a more effective modeling method and multi-source signal processing strategy are proposed, and their effectiveness is verified by experiments, hoping to provide a new method for the development of intelligent robot sensing technology.

2. Intelligent robot high-precision sensing system

The high-precision sensing system of intelligent robot, like its "five senses", is the key bridge for its interaction with the external environment and plays a decisive role in the accuracy and efficiency of robot's task execution in complex environment [7]. The system is mainly composed of various types of sensors. By capturing the optical information in the environment, the visual sensor provides the robot with visual image data such as the shape, color and position of the surrounding objects [8]. The auditory sensor can receive sound signals, realize speech recognition, sound source localization and other functions, so that the robot can communicate with the outside world through sound and obtain acoustic information in the environment. The tactile sensor simulates the human touch, endows the robot with the ability to sense contact force, pressure, texture and other information, and assists the robot to accurately control the strength and movement during operation.

High precision sensing depends on a series of key technologies. In vision, high-resolution imaging technology can improve image clarity and detail capture ability; In the field of hearing, accurate sound positioning technology can make robots accurately judge the direction of sound source; Tactile, sensitive tactile feedback technology helps to achieve more precise operation [9]. At present, the high-precision sensing system of intelligent robot faces many challenges. The difficulty of multi-source signal fusion lies in the great difference of signal characteristics of different types of sensors. How to effectively fuse and extract key information is the key to improve the overall performance of the system.

3. Modeling of high-precision sensing system

In the field of intelligent perception, modeling method is very important to accurately analyze various signals. Although the traditional modeling method based on physical characteristics has clear physical meaning, it is sensitive to parameter changes in complex environment [10]. In contrast, data-driven modeling can improve the adaptability of the model with a large number of actual data. CNN has unique advantages in processing visual signals. The convolution layer extracts local features, such as edges and textures, by sliding the convolution kernel on the image. Taking the two-dimensional image I as an example, the convolution operation can be expressed as:

$$(O_{i,j}^{k}) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} K_{m,n}^{k} \cdot I_{i+m,j+n} \tag{1}$$

Where $O_{i,j}^{k}$ is the value of the k channel of the convolved feature map at (i,j) position, $K_{m,n}^{k}$ is the value of the k convolution kernel at (m,n) position, $I_{i+m,j+n}$ is the value of the input image I at $(i+m,j+n)$ position, and M and N are the sizes of convolution kernels respectively. Through this operation, when recognizing the shape of an object, the convolution layer can capture the key features of the object contour. The pooling layer reduces the dimension of the convolution layer output, which reduces the amount of data and speeds up the calculation while retaining the main features. Through multi-layer convolution and pooling operation, CNN can construct a hierarchical representation of features from low level to high level.

It is difficult for CNN to deal with the data with time series characteristics such as auditory signals effectively. Therefore, LSTM is introduced into this study. LSTM has a special gating structure, including input gate, forgetting gate and output gate. Forgetting gate determines what information to keep from the last cell state; The input gate controls the current input information to enter the cell state; The output gate determines the output cell state information. Taking t time as an example, the calculation formula of forgetting gate f_t is as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2)$$

Where σ is the sigmoid activation function, W_f is the weight matrix of the forgetting gate, h_{t-1} is the hidden state of the last moment, x_t is the input of the current moment, and b_f is the bias of the forgetting gate. This structure enables LSTM to effectively deal with the long-term dependence problem and is suitable for analyzing the time series characteristics of audio signals.

In order to realize the joint modeling of visual and auditory signals, an algorithm combining CNN and LSTM is adopted in this study. Firstly, CNN is used to extract features from visual image data, and the feature vector of the image is obtained. At the same time, the auditory audio data is processed by LSTM to obtain the time sequence feature representation of audio. Then, the feature vectors obtained by the two methods are spliced and fused to form a comprehensive feature vector containing visual and auditory information. Finally, the comprehensive feature vector is classified or regressed by the fully connected layer.

In the model training stage, large-scale visual and auditory data sets are used, and the model parameters are adjusted by random gradient descent algorithm to minimize the model loss function and improve the model performance. In order to prevent over-fitting, regularization method and L2 regularization are used to constrain the model parameters. Through the above modeling and algorithm, the high-precision sensing system can model multi-source sensing signals more comprehensively and accurately, which provides strong support for robot's perception and decision-making in complex environment.

4. Multi-source signal processing strategy

Multi-source signal processing of intelligent robot aims to integrate signals collected by different sensors and extract key information from them, so as to improve the robot's perception and understanding of complex environment. It mainly covers feature extraction, signal fusion algorithm, interference suppression and noise processing.

Feature extraction of multi-source signals is the key first step. Different types of sensor signals have unique characteristics. The tactile signal can extract the characteristics related to the surface texture and hardness of the object by analyzing the changes of parameters such as pressure and vibration. Accurate feature extraction provides a basis for subsequent signal fusion and processing. Signal fusion algorithm is the core to realize the effective integration of multi-source signals. Neural network is widely used in signal fusion, and its powerful nonlinear mapping ability can automatically learn the complex relationship between multi-source signals and realize high-precision fusion.

In practical application, the sensor signal is easily affected by all kinds of interference and noise. For high frequency noise, low-pass filter can be used to filter out; For periodic interference, notch filter can be used for targeted processing. In addition, time-frequency analysis methods such as wavelet transform are also commonly used to remove noise, which can analyze and process signals in time-frequency domain, effectively separate noise components in signals, improve signal quality and reliability, and ensure the accuracy of multi-source signal processing results.

5. Experiment and result analysis

In order to verify the effectiveness of high-precision sensing system modeling and multi-source signal processing strategy, a series of experiments are designed in this section. The experimental environment simulates the common indoor complex scenes of intelligent robots, including various obstacles and different lighting conditions. The experimental equipment selects the mainstream visual, auditory and tactile sensors, builds the platform of sensing system, and realizes the proposed modeling and signal processing algorithm based on this.

5.1. Modeling accuracy experiment

The accuracy of modeling methods based on physical characteristics and data-driven is compared. By measuring the characteristic parameters of the same object many times, the predicted value and the true value of the model are recorded. Taking the measurement of object length by vision sensor as an example, the results are shown in the following Table 1:

Table 1. Comparison of Measurement Errors in Object Length by Visual Sensors under Different Modeling Methods.

Modeling Method	Number of Measurements	Average Error (mm)	Maximum Error (mm)
Physics-Based Characteristics	50	1.2	3.5
Data-Driven	50	0.8	2.1

As can be seen from Table 1, the data-driven modeling method is lower than the modeling method based on physical characteristics in average error and maximum error. This shows that the data-driven method can better adapt to the complex and changeable environment and improve the modeling accuracy by using a large number of actual data for model training.

5.2. Multi-source signal processing performance experiment.

Test the target recognition accuracy of different signal fusion algorithms in complex environment. A variety of target objects are set, and signals are collected and processed by multi-source sensors. The results are shown in Table 2 below:

Table 2. Comparison of Target Recognition Accuracy under Different Signal Fusion Algorithms.

Signal Fusion Algorithm	Target Recognition Accuracy (%)
Weighted Average Method	75
Kalman Filter Method	82
CNN-LSTM Fusion Algorithm	88

Observation Table 2 shows that the target recognition accuracy of CNN and LSTM fusion algorithm is the highest, followed by Kalman filter method, and the weighted average method is relatively low. With its powerful nonlinear fitting ability, the fusion algorithm of CNN and LSTM can fully explore the complex relationship between multi-source signals and realize more accurate fusion. Kalman filter method is based on system state estimation and has obvious advantages in processing dynamic signals. Although the weighted average method is simple, it is difficult to adapt to the diversity of signals in complex scenes because the weight distribution of each sensor signal is relatively fixed, which leads to limited accuracy.

Overall, the experimental results verify the effectiveness of the proposed modeling and multi-source signal processing strategy, but there is still room for further optimization, such as improving the adaptability of the modeling method based on physical characteristics to complex environments.

6. Conclusions

In this study, the modeling of high-precision sensing system is deeply explored. In the modeling method, the traditional modeling method based on physical characteristics is compared with the data-driven modeling method, and it is clear that data-driven modeling has obvious advantages in adapting to complex and changeable environment by training the model with a large number of actual data. It is lower than the traditional method in average error and maximum error index, which strongly proves its effectiveness in improving modeling accuracy. At the level of multi-source signal processing algorithm, CNN and LSTM fusion algorithm are innovatively adopted. Using CNN's local feature extraction ability in visual signal processing and LSTM's long-term dependence on auditory and other signals with time series characteristics, the joint modeling of visual and auditory signals is realized. The experimental results show that this fusion algorithm effectively improves the accuracy of target recognition, reaching 88%, which is much higher than the 75% of the weighted average method, further verifying the effectiveness of the algorithm.

Although the fusion algorithm has achieved good results, the balance between model complexity and computational efficiency needs to be further studied to meet the real-time requirements in practical applications. Generally speaking, this study provides an effective method for high-precision sensing system modeling and is of great significance for promoting the application of intelligent robots in complex environments.

References

[1] Yang S, Yang G, Kim D, et al. Development of Intelligent Soft Wearable Lumbar Support Robot for Flat Back Syndrome[J]. International Journal of Control, Automation and Systems, 2025, 23(4): 1147-1156.

[2] Tran C C, Lin C Y. An intelligent path planning of welding robot based on multisensor interaction[J]. IEEE sensors journal, 2023, 23(8): 8591-8604.

[3] Puttero S, Verna E, Genta G, et al. Preliminary comparison between manual assembly and intelligent human-robot collaborative assemblies in terms of quality and assembly time[J]. Procedia CIRP, 2024, 126: 206-211.

[4] Chaudhuri R, Deb S, Das H. Noble approach on sensor fused bio intelligent path optimisation and single stage obstacle recognition in customized mobile agent[J]. Procedia Computer Science, 2023, 218: 778-787.

[5] Lindemann L, Dimarogonas D V. Barrier function based collaborative control of multiple robots under signal temporal logic tasks[J]. IEEE Transactions on Control of Network Systems, 2020, 7(4): 1916-1928.

[6] Wang F, Chen K, Zhen S, et al. Fuzzy optimization design of adaptive robust control for uncertain cooperative robots with servo constraints[J]. International Journal of Fuzzy Systems, 2024, 26(4): 1158-1174.

[7] Dixit A, Oshiumi W, Shrivastava M, et al. An image processing mechanism for aerial inspection robots to detect submillimeter-width concrete cracks in social infrastructures[J]. Advanced Robotics, 2024, 38(9-10): 698-714.

[8] Zhang Y, Liu W, Tan N. Motion control of legged robots based on gradient central pattern generators[J]. Robotica, 2024, 42(9): 3102-3131.

[9] Staudinger E, Giubilato R, Schuster M J, et al. Terrain-aware communication coverage prediction for cooperative networked robots in unstructured environments[J]. Acta Astronautica, 2023, 202: 799-805.

[10] Ugenti A, Vulpi F, Domínguez R, et al. On the role of feature and signal selection for terrain learning in planetary exploration robots[J]. Journal of Field Robotics, 2022, 39(4): 355-370.