Mechanical Engineering Science

Publisher: Viser Technology Pte. Ltd.

ISSN: 2661-4448(online)
2661-443X(print)

Frequency: Semi-annual

Add.: 195 Pearl's Hill Terrace, #02-41,
Singapore 168976

https://www.viserdata.com/

Mechanical Engineering Science

Volume 7 No.2 (2025)

CONTENTS

A Survey of Remote Sensing Image Segmentation Based on Deep Learning 1

Shibo SUN, Yunzuo ZHANG

Research on Dual-motor Synchronization Based on Fuzzy Neural PID Control 11

Xiaoqiang WU

Research on Design Method of Dynamic Shop Floor Scheduling System Based on Human-computer Interaction 17

Songling TIAN, Zhuke CAI, Xiaoqiang WU, Xiaoqian QI

The Research on CAD Design System of Shaper Cutter Based on VB and Matlab 27

Xiaoqiang WU, Rui XUE, Kan XING, Tiegang WANG, Peng WANG, Fenghe WU, Yabin GUAN, S. ZHANG

A Simple and Reliable Eccentric Locking Mechanism 34

Mengjiao NIU, Yong ZHAO, Yongliang YUAN

VISER

Viser Technology Pte. Ltd.
Mechanical Engineering Science
DOI: 10.33142/mes.v7i2.17492

Research Article Open Access

A Survey of Remote Sensing Image Segmentation Based on Deep Learning

Shibo SUN, Yunzuo ZHANG*

School of Information science and Technology, Shijiazhuang Tiedao University, Shijiazhuang, Hebei, 050043, China

***Corresponding Author:** Yunzuo ZHANG, E-mail: zhangyunzuo888@sina.com*

Abstract

Remote sensing image segmentation has a wide range of applications in land cover classification, urban building recognition, crop monitoring, and other fields. In recent years, with the booming development of deep learning, remote sensing image segmentation models based on deep learning have gradually emerged and produced a large number of scientific research achievements. This article is based on deep learning and reviews the latest achievements in remote sensing image segmentation, exploring future development directions. Firstly, the basic concepts, characteristics, classification, tasks, and commonly used datasets of remote sensing images are presented. Secondly, the segmentation models based on deep learning were classified and summarized, and the principles, characteristics, and applications of various models were presented. Then, the key technologies involved in deep learning remote sensing image segmentation were introduced. Finally, the future development direction and application prospects of remote sensing image segmentation were discussed. This article reviews the latest research achievements in remote sensing image segmentation from the perspective of deep learning, which can provide reference and inspiration for the research of remote sensing image segmentation.

Keywords: *Remote sensing image segmentation; Deep learning; Split tasks; Model classification; Key technology*

1 Introduction

Remote sensing image segmentation is the process of dividing remote sensing images into different regions, and it is an important basis for remote sensing image analysis. The quantity, quality and diversity of remote sensing images have been improved with the updating of remote sensing platforms and sensors, bringing more data and higher requirements for remote sensing image segmentation. However, remote sensing image segmentation also faces challenges such as data scarcity, high annotation cost, large scale variation, class imbalance, complex background, etc., which affect the performance and adaptability of remote sensing image segmentation. In order to overcome these challenges, improve the accuracy and efficiency of remote sensing image segmentation, deep learning, a technique of automatic feature learning, has been widely applied and developed in the field of remote sensing image segmentation. Deep learning can use large amounts of data and complex network structures to automatically learn the features and patterns of remote sensing images, achieve end-to-end training and inference, and be applicable to various segmentation tasks. This paper reviews the concept, task, dataset, model classification, technical characteristics, development direction and application prospect of remote sensing image segmentation from the perspective of deep learning, and provides a reference perspective for the research and application of remote sensing image segmentation.

Remote sensing images are images of the earth's surface observed from a distance by sensors mounted on remote sensing platforms such as aircraft or artificial satellites. Remote sensing images have characteristics such as high spatial resolution, high spectral resolution, and high spatiotemporal coverage, and can reflect the physical, chemical, biological and other information of the earth's surface. Image segmentation is the basis of many visual understanding systems, and remote sensing image segmentation applies image segmentation techniques to the field of remote sensing, achieving pixel-level classification of remote sensing images, which has important applications in fields such as environmental monitoring, urban planning, land resource utilization, etc. Remote sensing image segmentation, as a special image segmentation task, encounters the following three main problems in the research process:

The objects of interest in remote sensing images have large scale variations, ranging from a few pixels to thousands of pixels, which leads to the multi-scale problem, making it difficult to locate and identify the objects of interest.

Received on July 12, 2025; Accepted on August 20, 2025

The background in remote sensing images is more complex and diverse, due to the influence of factors such as terrain, landform, season, illumination, etc., there are large intra-class differences and small inter-class differences in the background, [1] which leads to low class separability, making it easy for the objects of interest to be confused with the background.

The foreground objects in remote sensing images occupy a relatively small proportion, compared to the target objects in natural images, the foreground objects in remote sensing images often only occupy a small part of the image, [2] which leads to the foreground-background imbalance problem, making it easy for the objects of interest to be ignored or occluded.

2 Remote Sensing Images and Remote Sensing Image Segmentation Tasks

In this section, we mainly introduce the types of remote sensing images and the common tasks and datasets of remote sensing image segmentation.

2.1 Common types of remote sensing images and their characteristics

According to the different ways of sensor mounting, imaging, and sensing electromagnetic waves, remote sensing images can be divided into various types, each of which has its own characteristics and application fields. Here are some common types of remote sensing images and their characteristics:

Hyperspectral remote sensing images: Hyperspectral remote sensing images refer to remote sensing images that are imaged by remote sensing sensors in continuous narrow bands, usually containing hundreds to thousands of bands. Due to the complex characteristics of hyperspectral data, the accurate classification of hyperspectral data is challenging for traditional machine learning methods. In addition, the spectral information captured by hyperspectral imaging has a nonlinear relationship with the materials it corresponds to. In recent years, deep learning has been recognized as a powerful feature extraction tool that can effectively solve nonlinear problems. Driven by these successful applications, deep learning has also been introduced to hyperspectral remote sensing image classification and has shown good performance 3. Figure 1 shows an example of a hyperspectral remote sensing image.[3,4]

Synthetic aperture radar remote sensing images: Synthetic aperture radar remote sensing images refer to remote sensing images that are imaged by electromagnetic waves in the microwave band, usually obtained by synthetic aperture radar sensors mounted on platforms such as satellites or aircraft. Synthetic aperture radar remote sensing images can reflect the geometric features, roughness, dielectric constant, etc. of the objects, and are suitable for monitoring and analysis of terrain, landform, earthquake, landslide, etc. The characteristics of synthetic aperture radar remote sensing images are that they are not affected by weather, illumination, etc., and can achieve all-weather, all-day observation 4, but they also have problems such as complex scattering mechanism, low image quality, interference fringes, etc. Figure 2 shows an example of a synthetic aperture radar remote sensing image.

Lidar remote sensing images: Lidar remote sensing images refer to remote sensing images that are imaged by electromagnetic waves in the laser band, usually obtained by lidar sensors mounted on platforms such as satellites or aircraft. Lidar remote sensing images can reflect the elevation, morphology, structure, etc. of the objects, and are suitable for monitoring and analysis of surface elevation, buildings, vegetation, aerosols, etc.[5] The characteristics of lidar remote sensing images are that they have high accuracy, high resolution, high sensitivity, etc., and can achieve three-dimensional reconstruction, change detection, target recognition, etc., but they also have problems such as high cost, large data volume, complex processing, etc. Figure 3 shows an example of a lidar remote sensing image.

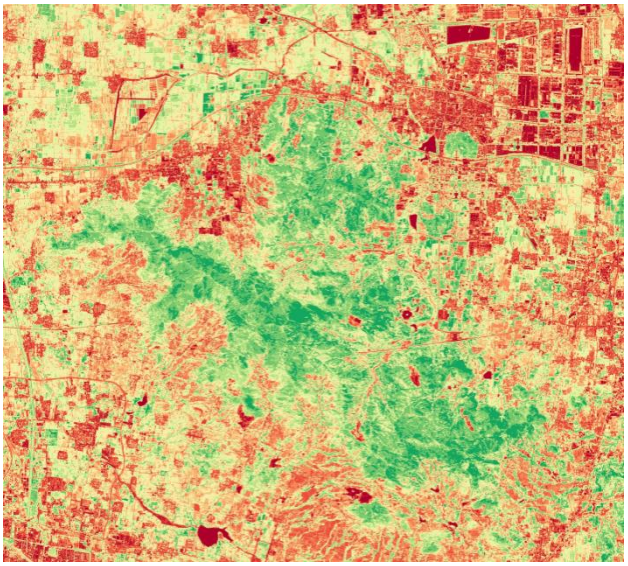


Figure 1 Xample of hyperspectral remote sensing images

Figure 2 Xample of Synthetic Aperture Radar Remote Sensing

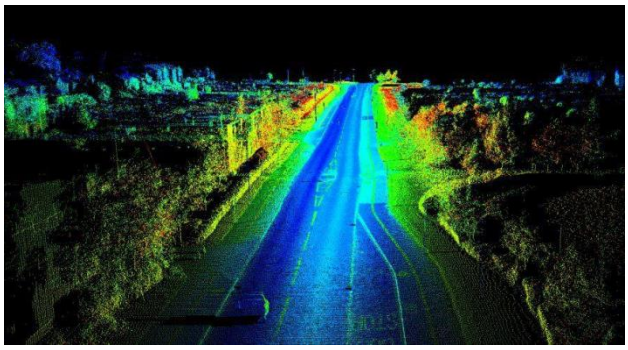


Figure 3 Example of LiDAR remote sensing images

2.2 Common tasks and data sets of remote sensing image segmentation

The common tasks and datasets of remote sensing image segmentation include:

Small object segmentation: Small objects occupy fewer pixels in remote sensing images, such as airplanes, vehicles, ships, etc. The difficulties of this task lie in the small size, irregular shape, complex background, and high similarity among the objects. [6,7]

Ship segmentation: Ship segmentation refers to the process of separating the ship objects from the water background in remote sensing images, which is the basis of maritime target monitoring. The difficulties of ship segmentation lie in the diverse shapes, different sizes, uneven colors, and low contrast with the water. [8]

Building segmentation: Building segmentation requires separating the building objects from the ground background in remote sensing images, which is an important means for urban planning and management. The difficulties of building segmentation lie in the complex shapes, different sizes, similar colors, and confusion with shadows and trees. [9]

Natural environment segmentation: Natural environment segmentation refers to the process of separating the natural environment objects from other backgrounds in remote sensing images, which have various types of objects, from clouds to vegetation to ice, with large differences. The difficulties of natural environment segmentation lie in the diverse classes, uneven distribution, fuzzy boundaries, and confusion with other objects. [10,11]

Road extraction: Road extraction refers to the process of separating the road objects from other backgrounds in remote sensing images, which is an important basis for traffic planning and management. The difficulties of road extraction lie in the different widths, complex shapes, uneven colors, and confusion with buildings and shadows. [12,13]

The commonly used datasets for remote sensing image segmentation include:

DOTA: DOTA is a large-scale aerial image dataset, containing 2806 images, with a total of 188282 small objects, divided into 15 categories, such as airplanes, ships, bridges, vehicles, etc. The images of DOTA come from Google Earth, with diverse perspectives, scales, backgrounds, and object densities.

SeaShips: SeaShips is a large-scale ship segmentation dataset, containing 31000 images, with a total of 59000 ship objects, divided into four categories, namely cargo ships, tankers, fishing boats, and speedboats. The images of SeaShips come from Planet Labs, with high resolution, multiple time phases, multiple angles, and multiple regions.

WHU Building Dataset: WHU Building Dataset is a high-resolution building segmentation dataset, containing 326 images, with more than 10000 building objects. The images of WHU Building Dataset come from QuickBird and WorldView-2, with different resolutions, angles, illuminations, and seasonal changes.

LandCoverNet: LandCoverNet is a global annual land cover classification dataset, containing 31000 images, each image is 256×256 pixels, divided into 7 categories, namely impervious surface, agriculture, forest, soil, water, wetland, and ice and snow.

The SpaceNet Datasets: The SpaceNet Datasets are a large-scale remote sensing image dataset, containing high-resolution satellite images and corresponding road network labels of multiple cities. The images of SpaceNet come from DigitalGlobe, with different resolutions, angles, illuminations, and seasonal changes. Table 1 lists the image sizes, categories, and numbers of each dataset.

Table 1 Introduction to Common Datasets and Their Content

Dataset name	Image size	Number of categories	Number of images
DOTA-v1.0	800～20 000	15	2806
DOTA-v1.5	800～20 000	16	2806
DOTA-2.0	800～20 000	18	11268
SeaShips	1920×1080	6	31455
WHU Building Dataset	800～20 000	1	2806
LandCoverNet	256×256	7	31000
The SpaceNet Datasets	900～1300	18	21346

3 Classification of Remote Sensing Segmentation Models Based on Deep Learning

In this section, we classify the segmentation models based on deep learning according to different network structures. The application fields and representative networks of each type of model are summarized in Table 2.

3.1 Segmentation models based on CNN

Convolutional neural networks (CNN) are neural networks that use convolution operations to extract image features, which have advantages such as translation invariance, parameter sharing, and sparse connection, and are suitable for processing image data. CNN networks were very popular in the early deep

learning models, and segmentation models based on CNN were one of the important breakthroughs of deep learning in the field of image segmentation, which can achieve end-to-end training and inference, and are applicable to various segmentation tasks. Common backbone networks such as VGG[14], ResNet[15], GoogLeNet[16] and others adopt the convolutional structure of CNN. In recent years, the rise of Tramsformer has brought new ideas and methods for image segmentation. Zhang[17] and others combined Tramsformer and CNN in their research and applied it to high-resolution remote sensing image semantic segmentation, and the method proposed was very close to the state-of-the-art methods in terms of overall accuracy.

3.2 Segmentation models based on GAN

Generative adversarial networks (GAN) are a technique that uses two competing neural networks to generate new data, where one network is called a generator, responsible for generating fake data, and the other network is called a discriminator, responsible for judging the authenticity of the data. Figure 4 shows the structure of the GAN network. Segmentation models based on GAN usually combine the traditional multi-class cross-entropy loss with the adversarial network, first pre-train the adversarial network, and then use the adversarial loss to fine-tune the segmentation network. Segmentation models based on GAN can use the generator's ability to enhance or reconstruct remote sensing images, thereby improving the quality and effect of segmentation. Tasar[18] and others proposed a color mapping-based method, using ColorMapGANs to transform the source domain remote sensing images into the target domain style, and then used a pre-trained segmentation network to segment the transformed images. The method achieved significant performance improvement on two high-resolution remote sensing datasets. In addition, GAN is also often used to solve the problem of domain transfer feature mismatch. Zhang[19] and others proposed a method of using GAN structure combined with cheap available data to train the model for the segmentation task under unsupervised conditions, which can effectively reduce the domain gap.

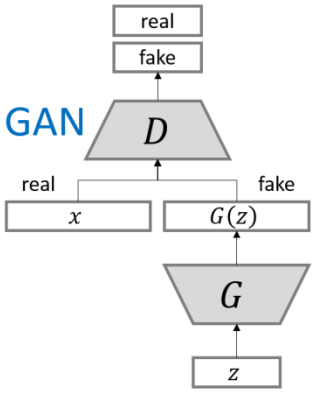


Figure 4 GAN network architecture diagram

3.3 Segmentation models based on Tramsformer

Tramsformer is a neural network that uses self-attention to achieve sequence-to-sequence mapping, which has advantages such as parallel computing, long-distance dependence, and position encoding. Segmentation models based on Tramsformer use Tramsformer to achieve image segmentation, which usually divide the input image into multiple sub-regions, and then use the encoder and decoder of Tramsformer to extract and reconstruct the features of each sub-region. This type of model can use the self-attention mechanism to perform global context understanding of remote sensing images, and improve the consistency and accuracy of segmentation. Figure 5 shows the structure of the Tramsformer network. Robin[20] and others proposed the Segmenter model based on Vision Tramsformer, which achieved excellent results in semantic segmentation. Liu[21] and others proposed a new visual Tramsformer, which can serve as a general backbone for computer vision, and its performance greatly exceeded the previous state-of-the-art techniques. He[22] and others embedded Swin-Tramsformer into the UNet network to form a new dual-encoder structure.

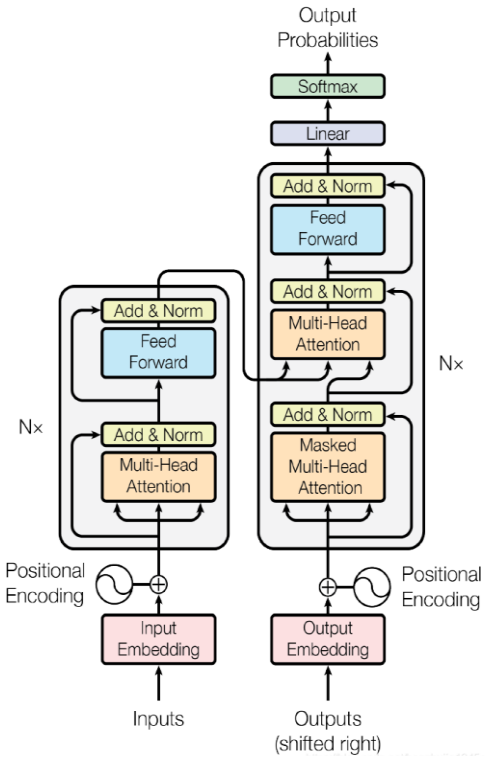


Figure 5 Tramsformer network architecture diagram

3.4 Segmentation models based on pre-trained models

Segmentation models based on pre-trained models are an important way to alleviate the problem of large-scale labeled data scarcity, which usually use pre-trained models as encoders or feature extractors, and then add a decoder or segmentation head after them. This

type of segmentation model can use the ability of pre-trained models to effectively extract features from remote sensing images, thereby improving the efficiency and effect of segmentation. Early pre-trained models include VGG, ResNet, DenseNet[23], etc., which have relatively few parameters. Li[24] and others used ResNeXt-101 instead of ResNet as the backbone in their model, enhancing the feature extraction ability. With the continuous improvement of data collection technology, pre-trained models based on large-scale datasets have been applied to the remote sensing field. ViT-G12[25] is one of the pre-trained models with a considerable amount of parameters. ViT-G12 was trained on the Million-AID dataset, with a parameter amount of 2.4B, and has a strong feature extraction ability.

4 Key Technologies for Remote Sensing Image Segmentation

The In this section, we summarize the key technologies in remote sensing image segmentation. Table 3 lists the representative models and the datasets they use based on the following six key technologies. As mentioned in Section 2, remote sensing image types and segmentation tasks are diverse, so different models use different key technologies. For example, small object segmentation tasks tend to pay more attention to semantic information, and long-distance dependence is particularly important when dealing with such tasks; while feature fusion operations can preserve spatial details and better complete boundary recognition. Remote sensing image segmentation key technologies can be classified according to different focuses, and this section will introduce the following six types:

Attention mechanism: Attention mechanism can make the network automatically focus on the most important parts when processing images, thereby improving the accuracy and robustness of segmentation. Attention mechanism is widely used in segmentation tasks, which can be divided into spatial attention and channel attention, the former focuses on different positions in the image, and the latter focuses on different features in the image. Figures 6 and 7 show the module structures of spatial attention and channel attention. Segmentation techniques based on attention mechanism can solve the problems of target size difference, complex background, occlusion, etc. in remote sensing images, and improve the details and boundaries of segmentation. Lei Ding[35] and others used adaptive attention mechanism in their research, which bridged the gap between high-level and low-level features, and kept the spatial details and semantic information during the feature fusion process. Although attention mechanism can improve the accuracy and robustness of segmentation, it also increases the computation and parameter amount of the network, resulting in slower training and inference speed of the network. In addition, attention mechanism may also over-focus on some specific areas and ignore other information, leading to overfitting problems.

Table 2 Application fields and representative models of various segmentation models

Category	Common areas of application	Representative model
CNN-based segmentation model	Feature classification, land cover monitoring, urban planning, etc	ConvNeXtV2[26] RepLKNet[27]
GAN-based segmentation model	Image enhancement, image reconstruction, image style conversion, etc	Pix2Pix[28] SPGAN-DA[29] PU-GAN[30]
Segmentation model based on Tramsformer	Remote sensing scene understanding, remote sensing target detection, remote sensing video analysis, etc	Conv2Former[31] Swin-Tramsformer[32]
Based on pre-trained segmentation models	Zero-shot segmentation of remote sensing images, multi-source fusion of remote sensing images, and cross-domain migration of remote sensing images	CMID[33] DGCC-EB[34]

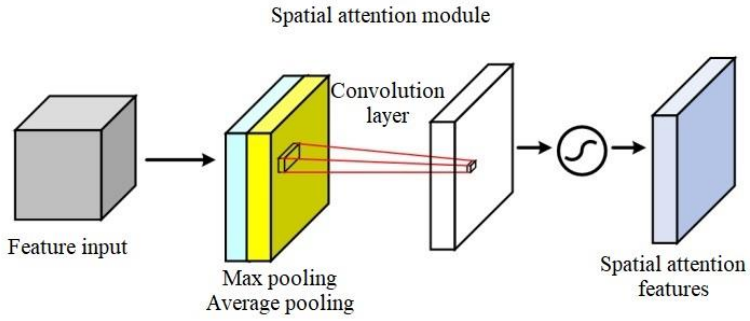


Figure 6 Spatial attention module

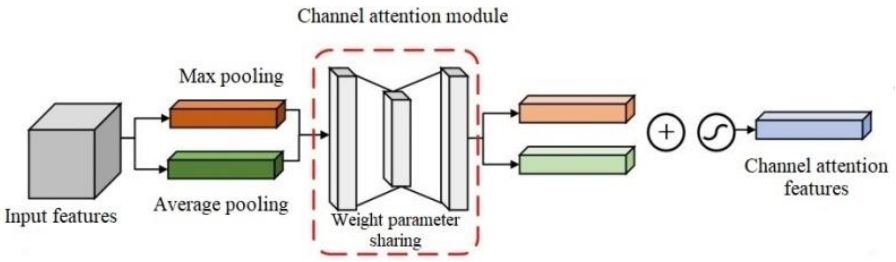


Figure 7 Channel attention module

Dilated convolution: Dilated convolution is a convolution operation that introduces dilation in the convolution kernel, which can increase the receptive field of the convolution kernel without increasing the parameters and computation, and capture more context information. Segmentation techniques based on dilated convolution can solve the problems of target detail loss, resolution reduction, etc. in remote sensing images, and improve the accuracy and clarity of segmentation. Figure 8 shows the dilated convolution model with dilation rates of 1, 2, and 4. Common segmentation techniques based on dilated convolution models include DeepLab[36], DenseASPP[37], etc. Zhao[38] and others used dilated convolution to enlarge the receptive field, and at the same time reduced the number of downsampling layers as much as possible to prevent the loss of detail information. Although dilated convolution plays a huge role in enhancing the receptive field, the size and stride of dilated convolution, the distribution and arrangement of dilated convolution, etc. are also problems that need to be solved. In addition, dilated convolution itself may also cause uneven receptive field of the network, and jagged phenomenon of the boundary.

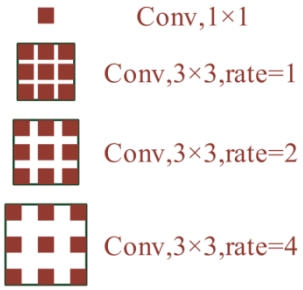


Figure 8 Dilated convolution with different magnifications

Feature fusion: Feature fusion uses different levels of feature information for effective fusion and utilization, thereby improving the details and boundaries of segmentation. Feature fusion techniques can solve the problems of target detail loss, resolution reduction, boundary blur, etc. in remote sensing images, and improve the accuracy of segmentation. There are three common ways of feature fusion, namely addition, concatenation, and fusion based on attention mechanism. Figure 9 shows the fusion process of high-level features and low-level features under the channel attention weight matrix. Common segmentation models that use feature fusion techniques include U-Net[39], RefineNet[40], etc. In this process, the way of feature fusion is very important, and inappropriate fusion methods may lead to noise introduction[41] Peng[42] and others proposed a cross-fusion module in their model, which used high-level feature maps to compensate for the receptive field of low-level feature maps, making the low-level feature maps have a similar receptive field to the high-level feature maps, thus significantly improving the semantic information capture ability of the low-level feature maps.

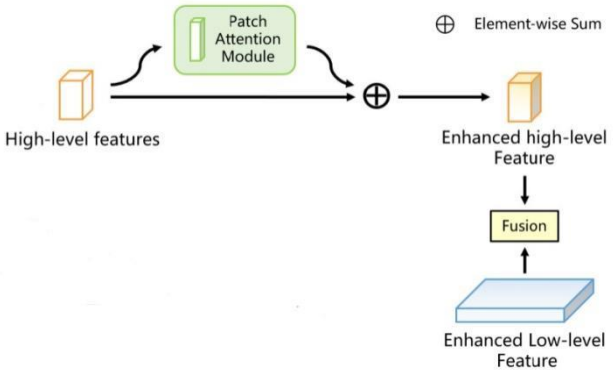


Figure 9 Fusion based on attention mechanism

Multi-task learning: Multi-task learning performs multiple segmentation tasks simultaneously, such as semantic segmentation, instance segmentation, boundary segmentation, etc., by sharing features and optimizing objectives, thereby improving the consistency and accuracy of segmentation. Multi-task learning segmentation techniques can solve the problems of target diversity, class imbalance, semantic ambiguity, etc. in remote sensing images, and improve the robustness and generalization of segmentation. For example, boundary segmentation and object segmentation tasks can promote each other[43] Li[44] and others used adaptive weight mechanism for multi-task learning, and separated the boundary information from the semantic features, and then used the corresponding loss to supervise.

Encoder-decoder structure: Encoder-decoder structure is a common structure in deep learning models, which uses two sub-networks, encoder and decoder, to perform feature extraction and feature reconstruction, respectively, to achieve pixel-level prediction of the

input image. Encoder-decoder structure segmentation technique is one of the important breakthroughs of deep learning in the field of image segmentation, which can achieve end-to-end training and inference, and is applicable to various segmentation tasks. Common segmentation techniques based on encoder-decoder structure models include U-Net, SegNet[45], etc. Liu[46] and others used encoder-decoder structure as the backbone network of their model.

Contrastive learning: Contrastive learning is a technique that uses positive and negative sample pairs in images to learn the feature representation of images by maximizing the similarity of positive sample pairs and minimizing the similarity of negative sample pairs. Figure 10 shows the feature extraction based on the existing support image to help the unknown image segmentation. Contrastive learning can solve the problems of target annotation insufficiency, unsupervised learning, self-supervised learning, etc. in remote sensing images to some extent, and improve the scalability and adaptability of segmentation. Tang[47] and others used the contrastive learning method of the self-supervised learning framework to improve the model's representation ability at the local pixel level.

5 Prospect

In this section, we mainly discuss the future development direction and application prospect of remote sensing image segmentation from a technical perspective.

5.1 Future development direction of remote sensing image segmentation

Remote sensing image segmentation has important applications in land use classification, urban planning, environmental monitoring, resource management, and other fields. The progress of remote sensing image segmentation technology has improved the accuracy and efficiency of remote sensing image segmentation. However, remote sensing image segmentation also faces problems such as data scarcity, high annotation cost, large scale variation, class imbalance, complex background, etc., which limit the performance and generalization ability of remote sensing image segmentation. To solve these problems, future research directions can be explored from the following three aspects:

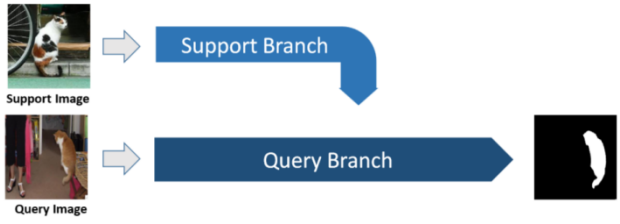


Figure 20 Contrastive learning

Table 3 Common key technologies and their representative models

Key technology	Applicable scenarios	Network structure	Training dataset
Attention mechanisms	Differences in target size, complex background, occlusion, etc.	DANet[48]	Cityscapes
		SENet[49]	ImageNet
		AGNet[50]	HDR+ Burst Photography Dataset
Dilated convolution	Loss of target detail, reduced resolution, etc	DeepLab	PASCAL VOC
		DenseASPP	Cityscapes
		DPC	COCO
Feature fusion	Loss of target detail, reduced resolution, blurred boundaries, etc	U-Net	Cell Tracking Challenge Dataset
		PSPNet[51]	ADE20K
		FPN[52]	COCO
Multi-task learning	Goal diversity, category imbalance, semantic ambiguity, etc	Mask R-CNN[53]	COCO
		FarSeg[54]	BraTS 2018
		DSSNet[55]	PASCAL VOC
Encoder-decoder structure	Various segmentation tasks	SegNet	CamVid
		DeconvNet[56]	PASCAL VOC
		FCN	PASCAL VOC
Contrastive learning	Insufficient target labeling, unsupervised learning, self-supervised learning, etc	MoCo[57]	ImageNet
		SimCLR[58]	ImageNet
		BYOL[59]	ImageNet

Multi-source data fusion: Remote sensing images contain multiple types of data, such as optical data, radar data, hyperspectral data, infrared data, etc., and different types of data have different features and advantages. By fusing different types of data, the robustness and accuracy of remote sensing image segmentation can be improved by using their complementarity. Multi-source data fusion methods include feature-level fusion and decision-level fusion, the former is to fuse different types of data at the feature extraction stage, and the latter is to fuse different types of data at the segmentation result stage. The key problem of multi-source data fusion is how to design effective fusion strategies, balance the differences and consistency between different data, and how to deal with the incompleteness of data.

Low-sample and zero-sample learning: One of the difficulties of remote sensing image segmentation is class imbalance, because the land cover classes in remote sensing images have diversity and complexity, and some classes have few samples, or even do not appear in the training data, which makes it difficult for the model to learn the features of these classes, affecting the accuracy of segmentation. Facing this problem, low-sample and zero-sample learning are promising research directions.

Low-sample learning uses a small number of samples to learn, and improves the generalization ability of the model by using transfer learning. Zero-sample learning uses auxiliary information, such as semantic information or attribute information, to learn, and establishes the association between classes, to achieve the recognition of unseen classes. The key problem of low-sample and zero-sample learning is how to design effective feature extraction and feature matching methods, use information from different sources, and how to deal with the differences and similarities between classes.

Interpretability and reliability: One of the objectives of remote sensing image segmentation is to provide reliable information support for remote sensing applications, therefore, the results of remote sensing image segmentation require not only accuracy, but also interpretability and reliability. However, the wide application of deep learning makes the remote sensing image segmentation model more and more complex, making the internal mechanism and output results of the model difficult to understand and verify, which may cause users' distrust. To improve the interpretability and reliability of remote sensing image segmentation, research needs to be conducted from three aspects: model, data, and result. Model aspect, it is necessary to design interpretable model structure, or provide model visualization and explanation methods, to reveal the working principle and key factors of the model. Data aspect, it is necessary to provide data quality assessment and uncertainty analysis, to evaluate the reliability and applicability of data. Result aspect, it is necessary to provide result confidence assessment and error detection, to evaluate the credibility and risk of the result.

5.2 Application prospect of remote sensing image segmentation

With the development of remote sensing technology, the application level of remote sensing image segmentation will also be continuously improved, providing more intelligent remote sensing services for various industries. The following are some application scenarios of remote sensing image segmentation:

Land cover change detection: Land cover refers to the natural and artificial coverings on the earth's surface, such as water, vegetation, bare soil, buildings, etc., which is an important indicator of the physical, chemical and biological processes on the earth's surface. The change of land cover will affect climate change, ecosystem, resource utilization, urban development, and other aspects, therefore, monitoring the change of land cover is of great significance for understanding and managing the earth environment. Remote sensing image segmentation can segment remote sensing images of different time phases, extract the categories and ranges of land cover, and then compare and analyze the changes of different categories, thus realizing the detection of land cover change.

Urban planning and management: Remote sensing image segmentation can finely extract and classify the elements of the city, such as buildings, roads, green spaces, water bodies, etc., to obtain the information of the city's spatial structure, functional distribution, ecological environment, etc., and provide scientific basis and reference for urban planning and management. Remote sensing image segmentation can also monitor and evaluate the changes of the city dynamically, and analyze the characteristics of the city's expansion, renewal, density, morphology, etc., and provide guidance and suggestions for the optimization and adjustment of the city.

Disaster monitoring and assessment: Remote sensing image segmentation can segment the remote sensing images before and after the disaster, and extract the information of the disaster type, range, degree, impact, etc., thus realizing the rapid identification, quantitative assessment and loss analysis of the disaster. Remote sensing image segmentation can also track and predict the evolution process of the disaster, and analyze the dynamic changes, development trends and potential risks of the disaster, and provide support and basis for disaster early warning and prevention.

Ecological environment protection: Remote sensing image segmentation can segment the various elements of the ecological environment, such as vegetation, soil, water, etc., and extract the information of the ecological environment structure, function, quality, service, etc., thus obtaining the information of the ecological environment status, change, problem, etc., and provide scientific data and methods for ecological environment protection. Remote sensing image segmentation can also monitor and evaluate the restoration and improvement of the ecological environment, and analyze the restoration effect, improvement measures, optimization schemes, etc., and provide decision basis and suggestions for ecological environment protection.

6 Conclusion

In the past decade, deep learning has achieved explosive development, greatly stimulating the research and application of deep learning in remote sensing image segmentation tasks. This paper introduces some basic concepts of remote sensing image segmentation, as well as its common tasks and datasets, and reviews the research status of remote sensing image segmentation models and techniques based on deep learning. Finally, several possible directions for future research are discussed. We hope that this research can provide valuable insights for researchers and inspire researchers to make more progress.

References

[1] L. Huang, B. Jiang. Deep Learning-based Semantic

Segmentation of Remote Sensing Images: A Survey [J]. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 2023:1-28.
[2] Q. An, Z. Pan. DRBox-v2: An Improved Detector With Rotatable Boxes for Target Detection in SAR Images [J]. IEEE Trans. Geosci. Remote Sens., 2019, 57(11): 8333-8349.
[3] S. Li, W. Song. Deep Learning for Hyperspectral Image Classification: An Overview [J]. IEEE Trans. Geosci. Remote Sens., 2019, 57(9): 6690-6709.
[4] R. Shang, M. Liu. Region-Level SAR Image Segmentation Based on Edge Feature and Label Assistance [J]. IEEE Trans. Geosci. Remote Sens., 202260: 1–16.
[5] F. Gaudfrin, O. Pujol. A New Lidar Technique Based on Supercontinuum Laser Sources for Aerosol Soundings: Simulations and Measurements—The PERFALIS Code and the COLIBRIS Instrument [J]. IEEE Trans. Geosci. Remote Sens., 2023, 61: 1-21.
[6] A. Ma, J. Wang. FactSeg: Foreground Activation-Driven Small Object Semantic Seg-mentation in Large-Scale Remote Sensing Imagery [J]. IEEE Trans. Geosci. Remote Sens., 2022, 60: 1-16.
[7] J.-H. Kim, Y. Hwang. GAN-Based Synthetic Data Augmentation for Infrared Small Target Detection [J]. IEEE Trans. Geosci. Remote Sens., 2022, 60: 1-12.
[8] N. Wang, B. Li. Ship Detection in Spaceborne Infrared Image Based on Lightweight CNN and Multisource Feature Cascade Decision [J]. IEEE Trans. Geosci. Remote Sens., 2021, 59(5): 4324-4339.
[9] J. Kang, et al. DisOptNet: Distilling Semantic Knowledge From Optical Images for Weather-Independent Building Segmentation [J]. IEEE Trans. Geosci. Remote Sens., 2022, 60: 1-15.
[10] K. Heidler, L. Mou. HED-UNet: Combined Segmentation and Edge Detection for Monitoring the Antarctic Coastline [J]. IEEE Trans. Geosci. Remote Sens., 2022, 60: 1-14.
[11] Z. Lu, et al. An Iterative Classification and Semantic Segmentation Network for Old Landslide Detection Using High-Resolution Remote Sensing Images [J]. IEEE Trans. Geosci. Remote Sens., 2023, 61: 1-13.
[12] DiResNet: Direction-Aware Residual Network for Road Extraction in VHR Remote Sensing Images.
[13] Y. Wei, S. Ji. Scribble-Based Weakly Supervised Deep Learning for Road Surface Extraction From Remote Sensing Images [J]. IEEE Trans. Geosci. Remote Sens., 2022, 60: 1-12.
[14] K. Simonyan, A. Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition [J]. arXiv, 2015.
[15] K. He, X. Zhang. Deep Residual Learning for Image Recognition [J]. arXiv, 2015.
[16] C. Szegedy, et al. Going deeper with convolutions [C]. In 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE: Boston, MA, USA.
[17] C. Zhang, W. Jiang. Transformer and CNN Hybrid Deep Neural Network for Semantic Segmentation of Very-High-Resolution Remote Sensing Imagery [J]. IEEE Trans. Geosci. Remote Sens., 2022, 60: 1-20.
[18] O. Tasar, S.L. Happy. ColorMapGAN: Unsupervised Domain Adaptation for Semantic Segmentation Using Color Mapping Generative Adversarial Networks [J]. IEEE Trans. Geosci. Remote Sens., 2020, 58(10): 7178-7193.
[19] L. Zhang, M. Lan. Stagewise Unsupervised Domain Adaptation With Adversarial Self-Training for Road Segmentation of Remote-Sensing Images [J]. IEEE Trans. Geosci. Remote Sens., 2022, 60: 1-13.
[20] R. Strudel, R. Garcia. Segmenter: Transformer for Semantic Segmentation [C]. In 2021 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE: Montreal, QC, Canada.
[21] Z. Liu, et al. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows [C]. In 2021 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE: Montreal, QC, Canada.
[22] C. Zhang, L. Wang. SwinSUNet: Pure Transformer Network for Remote Sensing Image Change Detection [J]. IEEE Trans. Geosci. Remote Sens., 2022, 60: 1-13.
[23] Y. Zhu, S. Newsam. DenseNet for dense flow [C]. In 2017 IEEE International Conference on Image Processing (ICIP), IEEE: Beijing.
[24] R. Li, et al. Multiattention Network for Semantic Segmentation of Fine-Resolution Remote Sensing Images [J]. IEEE Trans. Geosci. Remote Sens., 2022, 60: 1-13.
[25] K. Cha, J. Seo. A Billion-scale Foundation Model for Remote Sensing Images [J]. arXiv, 2023.
[26] S. Woo, et al. ConvNeXt V2: Co-designing and Scaling ConvNets with Masked Autoencoders [J]. arXiv, 2023.
[27] X. Ding, X. Zhang. Scaling Up Your Kernels to 31x31: Revisiting Large Kernel Design in CNNs.
[28] P. Isola, J.-Y. Zhu. Image-to-Image Translation with Conditional Adversarial Networks [J]. arXiv, 2018.
[29] Y. Li, T. Shi. SPGAN-DA: Semantic-Preserved Generative Adversarial Network for Domain Adaptive Remote Sensing Image Semantic Segmentation [J]. IEEE Trans. Geosci. Remote Sens., 2023, 61: 1-17.
[30] L. Zhou, H. Yu. PU-GAN: A One-Step 2-D InSAR Phase Unwrapping Based on Conditional Generative Adversarial Network [J]. IEEE Trans. Geosci. Remote Sens., 2022, 60: 1-10.
[31] Q. Hou, C.-Z. Lu. Conv2Former: A Simple Transformer-Style ConvNet for Visual Recognition [J]. arXiv, 2022.
[32] Z. Liu, et al. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows [C]. In 2021 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE: Montreal, QC, Canada.
[33] D. Muhtar, X. Zhang. CMID: A Unified Self-Supervised Learning Framework for Remote Sensing Image Understanding [J]. IEEE Trans. Geosci. Remote Sens., 2023, 61: 1-17.
[34] H. Zhang, et al. DGCC-EB: Deep Global Context Construction With an Enabled Boundary for Land Use Mapping of CSMA [J]. IEEE Trans. Geosci. Remote Sens., 2022, 60: 1-15.
[35] L. Ding, H. Tang. LANet: Local Attention Embedding to Improve the Semantic Segmentation of Remote Sensing Images [J]. IEEE Trans. Geosci. Remote Sens., 2021, 59(1):

426-435.
[36] L.-C. Chen, G. Papandreou. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs [J]. arXiv, 2017.
[37] M. Yang, K. Yu. DenseASPP for Semantic Segmentation in Street Scenes [C]. In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE: Salt Lake City, UT, USA.
[38] Q. Zhao, J. Liu. Semantic Segmentation With Attention Mechanism for Remote Sensing Images [J]. IEEE Trans. Geosci. Remote Sens., 2022, 60: 1-13.
[39] O. Ronneberger, P. Fischer. U-Net: Convolutional Networks for Biomedical Image Segmentation [J]. arXiv, 2015.
[40] G. Lin, A. Milan. RefineNet: Multi-path Refinement Networks for High-Resolution Semantic Segmentation [C]. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE: Honolulu, HI.
[41] J. Fu, X. Sun. An Anchor-Free Method Based on Feature Balancing and Refinement Network for Multiscale Ship Detection in SAR Images [J]. IEEE Trans. Geosci. Remote Sens., 2021, 59(2): 1331-1344.
[42] C. Peng, K. Zhang. Cross Fusion Net: A Fast Semantic Segmentation Network for Small-Scale Semantic Information Capturing in Aerial Scenes [J]. IEEE Trans. Geosci. Remote Sens., 2022, 60: 1-13.
[43] J. Zheng, A. Shao. Remote Sensing Semantic Segmentation via Boundary Supervi-sion-Aided Multiscale Channelwise Cross Attention Network [J]. IEEE Trans. Geosci. Remote Sens., 2023, 61: 1-14.
[44] A. Li, L. Jiao. Multitask Semantic Boundary Awareness Network for Remote Sensing Image Segmentation [J]. IEEE Trans. Geosci. Remote Sens., 2022, 60: 1-14.
[45] V. Badrinarayanan, A. Kendall. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation [J]. IEEE Trans. Pattern Anal. Mach. Intell., 2017, 39(12): 2481-2495.
[46] W. Liu, F. Su. Bispace Domain Adaptation Network for Remotely Sensed Semantic Segmentation [J]. IEEE Trans. Geosci. Remote Sens., 2020: 1-11.
[47] M. Tang, K. Georgiou. Semantic Segmentation in Aerial Imagery Using Multi-level Contrastive Learning with Local Consistency [C]. In 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), IEEE: Waikoloa, HI, USA.
[48] H. Xue, C. Liu. DANet: Divergent Activation for Weakly Supervised Object Localization [C]. In 2019 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE: Seoul, Korea (South).
[49] J. Hu, L. Shen. Squeeze-and-Excitation Networks [J]. arXiv, 2019.
[50] S. Zhang, et al. Attention Guided Network for Retinal Image Segmentation [J]. arXiv, 2019.
[51] H. Zhao, J. Shi. Pyramid Scene Parsing Network [J]. arXiv, 2017.
[52] T.-Y. Lin, P. Dollár. Feature Pyramid Networks for Object Detection [J]. arXiv, 2017.
[53] K. He, G. Gkioxari. Mask R-CNN [J]. arXiv, 2018.
[54] Z. Zheng, Y. Zhong. Foreground-Aware Relation Network for Geospatial Object Segmentation in High Spatial Resolution Remote Sensing Imagery.
[55] B. Pan, X. Xu. DSSNet: A Simple Dilated Semantic Segmentation Network for Hyperspectral Imagery Classification [J]. IEEE Geosci. Remote Sens. Lett., 2020, 17(11): 1968-1972.
[56] H. Noh, S. Hong. Learning Deconvolution Network for Semantic Segmentation [J]. 2015.
[57] K. He, H. Fan. Momentum Contrast for Unsupervised Visual Representation Learning [J]. arXiv, 2020.
[58] T. Chen, S. Kornblith. A Simple Framework for Contrastive Learning of Visual Representations [J]. arXiv, 2020.
[59] J.-B. Grill. Bootstrap your own latent: A new approach to self-supervised Learning [J]. 2020.

Viser Technology Pte. Ltd.
Mechanical Engineering Science
DOI: 10.33142/mes.v7i2.18368

Research Article Open Access

Research on Dual-motor Synchronization Based on Fuzzy Neural PID Control

Xiaoqiang WU*

Ordos INSTITUTE OF TECHNOLOGY College of Mechanical and Transportation Engineering, Ordos, Inner Mongolia, 017000, China

***Corresponding Author:** Xiaoqiang WU, E-mail: wangzai8402@163.com*

Abstract

In order to solve the problem of double motor synchronous error in the hydraulic lifting system of large crane, fuzzy control and neural network control are combined to realize the dynamic correction of PID parameters. With the use of cross-coupling control method in the control process based on the dynamic characteristics of the hydraulic system, both the pressure difference of hydraulic motor outlet and displacement of steel wire rope are regard as control index on the simulation and experimental research to improve the accuracy of synchronous control. The results show that this control strategy has strong ability of anti-interference, and effectively improving the synchronization control precision of the two motors.

Keywords: *Crane; fuzzy neural network; Cross coupling; Synchronous control*

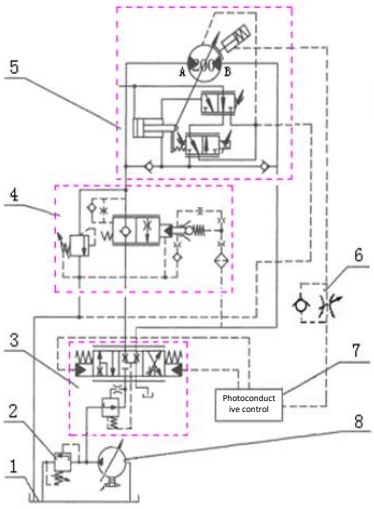
1 Introduction

The large crane hoisting system often uses two hydraulic motor to improve the load, but for the negative influence of factors included with leakage and prone to synchronization error, the crane's working performance is limited, or even cause accidents. In order to ensure the operation safety, effective control strategy ought to be adopted to keep the two motors with high synchronization precision.

The control method for now is applied the master-slave control mode to the hydraulic synchronization, whose References is cited in the text just by square brackets [1].Although it has better control effect, it is not suitable for large variations in load situations. Conventional PID control is the most widely used in control strategy, which is simple and easy to implement [2], however, because of the fixed parameter, real-time adjustment application is limited. Through the analysis for the dynamic characteristics of the crane hoisting system, we find out the factors influencing the synchronization accuracy; According to the characteristics of hydraulic transmission, nonlinear time-varying fuzzy control and neural network control combined fuzzy neural PID control strategy is proposed, for the purpose of achieving real-time adjusting PID parameters. At the same time, the cross-coupling control method is used to simulate the hydraulic pressure and the displacement of the steel wire rope.

2 Characteristic Analysis of Hoisting System

The lifting system of the large crane is composed of two independent groups of pump control motor system, which are coordinated by the controller to realize the synchronous control [3]. The structure of the subsystem is shown in figure 1.



1- variable pump; 2- reversing valve; 3- balance valve; 4- variable motor, 5-Variable motor;6-One-way throttle valve;7-pilot control valve;8-Constant power variable pump

Figure 1 Structure of crane hydraulic lifting system

Received on June 5, 2025; Accepted on July 20, 2025

In order to get the ideal synchronization control precision, it is essential to control the speed synchronization of two hydraulic motors. The variable mechanism of variable pump is composed of proportional valve and variable cylinder [4,5]:

Flow continuity equation of variable cylinder:

$$Q_1 - C_{tg} p_1 + C_{ig} p_2 - = \frac{dV_1}{dt} + \frac{V_1}{\beta_e} \cdot \frac{dp_1}{dt} \tag{1}$$

$$C_{ig} p_1 - C_{tg} p_2 - Q_2 = \frac{dV_2}{dt} + \frac{V_2}{\beta_e} \cdot \frac{dp_2}{dt} \tag{2}$$

Where: V_1—Volume of rodless cylinder of variable cylinder;
V_2—Volume of the cylinder in the variable cylinder;
β_e—Elastic modulus of hydraulic oil;
C_{ig}—Cylinder internal leakage coefficient;
C_{tg}—Cylinder leakage coefficient;
Q_1, Q_2—Flow rate of oil inlet and return chamber.

Force balance equation of cylinder.

$$m\frac{d^2X}{dt^2} + B\frac{dX}{dt} + KX + F_L = Ap_1 - Ap_2 \tag{3}$$

Where: m—Load and total piston mass;
B—Viscous damping coefficient of hydraulic oil;
F_L—Load force;
X—Piston displacement;
K—Spring stiffness;
A—Effective working area of piston.

Force balance equation of valve core:

$$m_v \frac{d^2Y}{dt^2} + B_v \frac{dY}{dt} + K_v Y = K_i I - K_y Y \tag{4}$$

Where: m_v—Spool quality;
Y—Spool displacement;
K_v—Spring stiffness;
K_i—Current gain;
I—Control current;
K_y—Displacement force gain;
B_v—Damping coefficient.

3 Control Strategy Research

3.1 Control mode

At present, there are three kinds of control methods used in hydraulic synchronization, which are equal control, master-slave control and cross coupling control [6-8].

3.1.1 Equal-status

Equal-status refers to the hydraulic system in the implementation of several components at the same time with an ideal input signal as the target, tracking output, so as to achieve synchronization control, as shown in figure2.

3.1.2 Master-slave

Master-slave mode refers to a way achieved by the synchronization control of the output including two components: one is the target, the other is the tracking control, to achieve a way, as shown in figure 3.

3.1.3 Cross-coupling

Cross coupling control mode refers to an ideal input as the goal, by coupling variables, to achieve synchronization control and implementation of all the components of the tracking control, then the output of the two actuators are compared, whose deviation signal is obtained as an additional signal feedback, such as shown in figure 4.

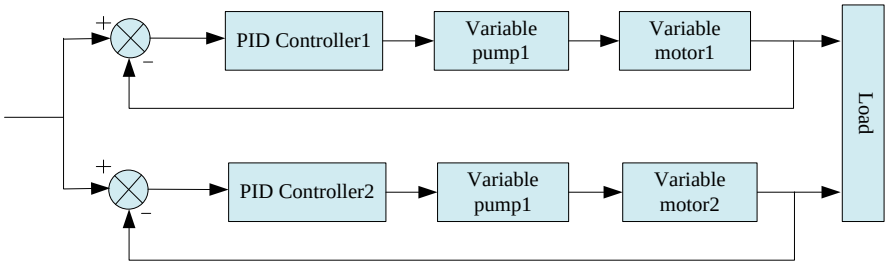


Figure 2 Equal-status control principle

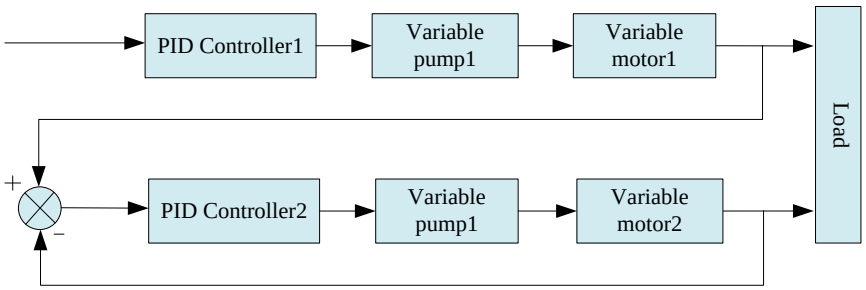


Figure 3 Master-slave control principle

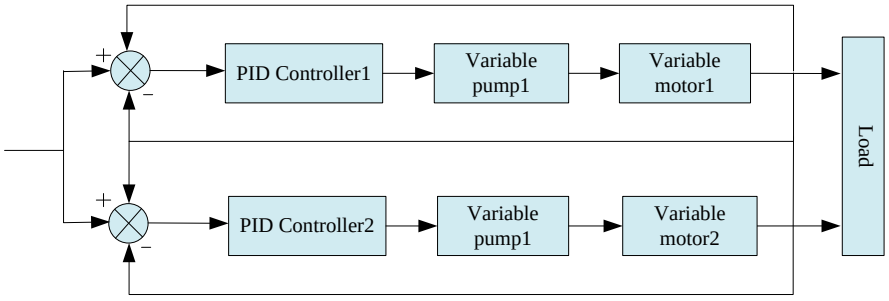


Figure 4 Cross-coupling control principle

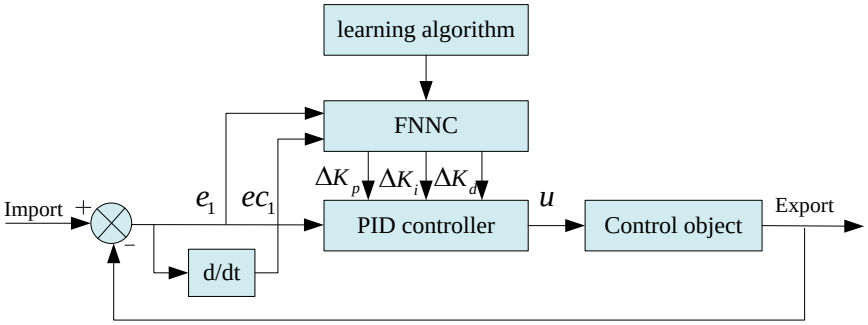


Figure 5 PID structure of fuzzy neural network

The above three methods can achieve both double motor synchronous control and satisfactory control effect, the first two methods have the advantages of simple structure and easy realization, but they are not suitable for load changes in the larger situation, which causes that it is difficult to achieve synchronization in the true sense. While by the cross-coupling control method, a coupling signal is added between each subsystem, which can improve the accuracy of the synchronization control.

3.2 Control strategy

Hydraulic lifting system is a very complex system, so as that it is difficult to establish accurate mathematical model. In order to obtain the optimal control effect, it is necessary to adjust the parameters in real time to counteract the disturbance caused by the disturbance. A parameter self-learning PID controller is designed based on fuzzy neural network, as shown in figure 5. The weights of the neural network are calculated by fuzzy rules, and the parameters of the traditional PID controller can be adjusted online in real time [9].

The fuzzy neural network model in this paper adopt five layers of structure, there are two input nodes and three output nodes, the input node corresponding to the error and error change rate of the output, after processing [10-12], the three parameters of the output node corresponding to the PID controller, as shown in figure 6.

The first layer is the input layer with two nodes, which are connected with the input vector, and the function is to change the error and error change rate of the input network.

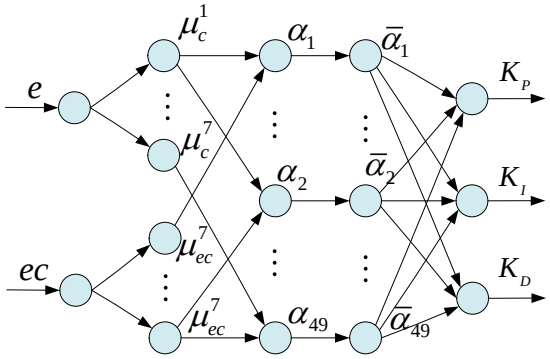


Figure 6 Topological structure of fuzzy neural network

The second layer is the language variable layer, also called fuzzy layer. Each node corresponds to a language variable value, whose function is to determine the amount of input into fuzzy vector as the input of the controller, using the normal function as its membership function:

$$\mu_i^j = \exp[-(x_i - c_{ij})^2] / \sigma_{ij}^2 \tag{5}$$

Where: c_{ij}—Center of membership function;

σ_{ij}^2—Width of membership function.

The third layer is the fuzzy inference layer, and each node represents a fuzzy rule. By connecting with the fuzzification layer to complete the matching of fuzzy rules, the corresponding weighting coefficients are obtained through the combination of each node, and the calculation of each application of the rules.

The fourth layer is the normalized layer, which is normalized.

The fifth layer is the output layer, when the fuzzy variables has been clear, it has the function to determine the parameters of the PID controller.

Obviously, the designed fuzzy neural network is also a kind of multilayer forward feed network, which can adjust the parameters by the method of error feedback:

$$e = \frac{1}{2}\sum_{i=1}^{3}(r_i - y_i)^2 \tag{6}$$

Where: r_i —Expected output;

y_i —Actual output.

The learning rule of connection weights is:

$$\omega_{ij}(k+1) = \omega_{ij}(k) + \eta(r_i + y_i) + \lambda[\omega_{ij}(k) - \omega_{ij}(k-1)] \tag{7}$$

Where: η —Learning rate;

λ —Smoothing factor, and $0 < \lambda < 1$.

4 Simulation Analyses

Based on the cross-coupling control mode, the fuzzy neural network control strategy is used in the double motor synchronous control system of the crane, and combined with the traditional PID control, respectively, to control the two hydraulic motors, as shown in figure 7.

Taking the variable pump displacement of 0～145mL/r, the engine speed of 1000r/min, motor displacement of 160mL/r, sampling frequency of 50Hz, and the motor pressure cut-off valve set value of 20MPa, simulation analysis is carried out under different working conditions and control strategies.

(1) Taking the minimum displacement difference of wire rope as the target

When the displacement of two steel ropes is not at the same time, using the deviations with two hydraulic pump displacement to control the controller, thereby together with changing the system flow, the motor speed changes. The displacement is equal to two steel ropes, and the displacement difference is zero, through which it can achieve synchronous control.

Assuming that the initial displacement difference between the two steel ropes is 1.2mm, crane weights 200 tons, the results shown in figure 8～9.

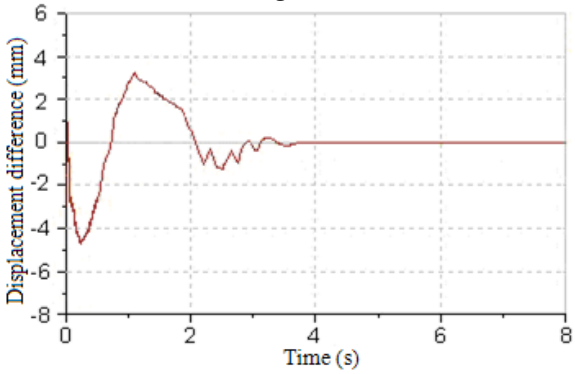


Figure 8 Traditional PID control

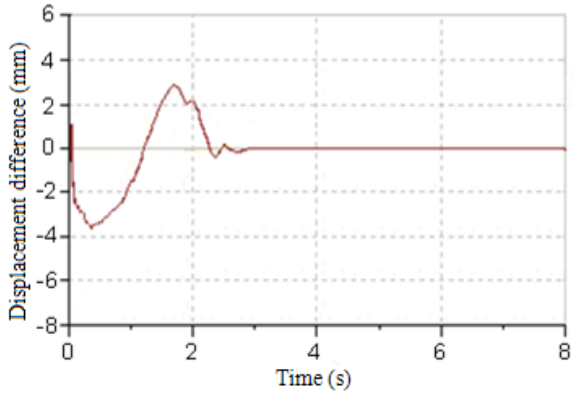


Figure 9 Fuzzy neural PID control

The simulation results show that the displacement difference of the two steel ropes will be close to 0 after the synchronization control. When using the traditional PID control strategy, the difference between the maximum displacement of steel wire and rope change rate reached 4.8mm, in 3.7 seconds to 0, while the fuzzy neural PID control strategy when the maximum displacement difference range is 3.6mm, up to 0 values in 2.8 seconds, the control precision and the convergence speed is improved obviously.

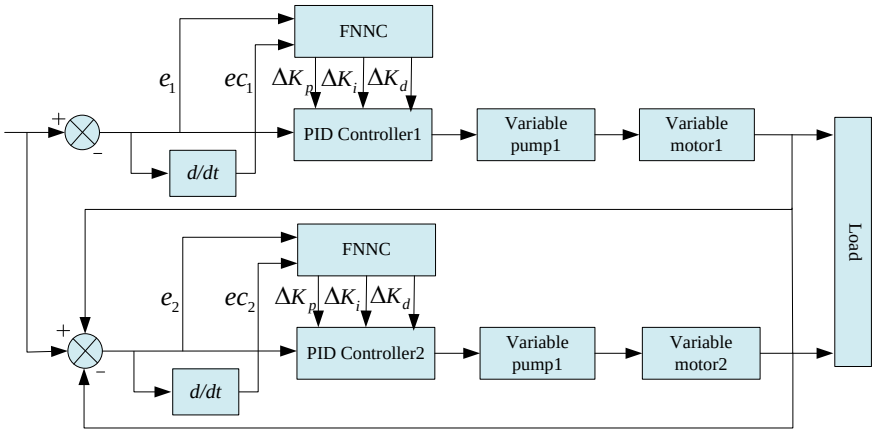


Figure 7 Diagram of double motor synchronous control system

(2) Taking the minimum pressure difference as the simulation target

When there is a synchronization error, the load will be inclined, which will lead to the unbalanced force of the wire rope, so that the load cannot be evenly distributed to the two hydraulic systems, leading to the inconsistence of the pressure of their corresponding position. Taking the exit pressure of two hydraulic motor as the index, the result is shown in figure 10～12.

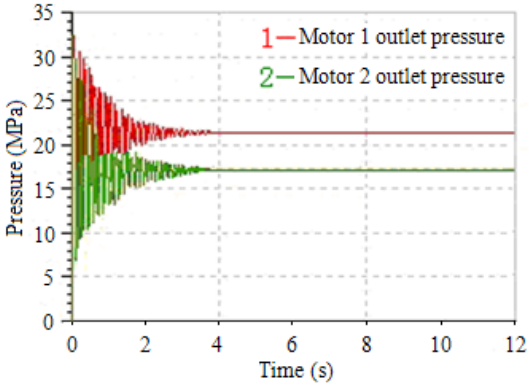


Figure 10 No synchronization control

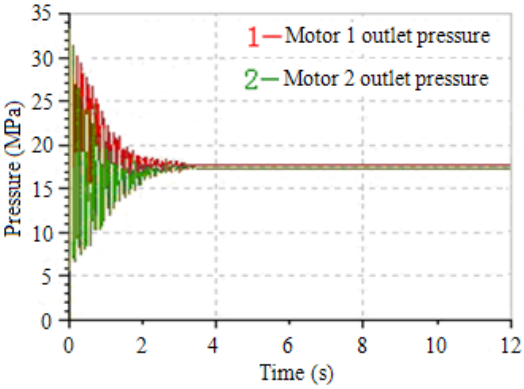


Figure 11 Traditional PID control

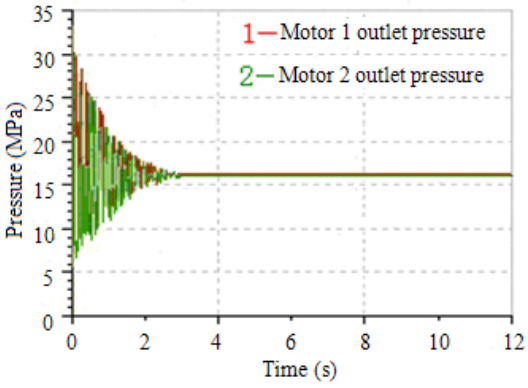


Figure 12 Fuzzy neural PID control

The results show that when the system synchronization error occurs, it will result in the deviation of the pressure of the two hydraulic systems. Compared with the traditional PID control strategy, the fuzzy neural PID control effect is better.

5 Experimental Study

In order to verify the correctness of the simulation analysis and the practicability of the control method, a crane is taken as the research object, and the experiment is carried out with the same parameters. The results are shown in figure 13～14.

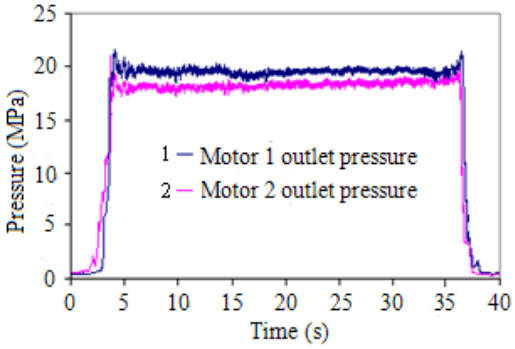


Figure 13 Traditional PID control

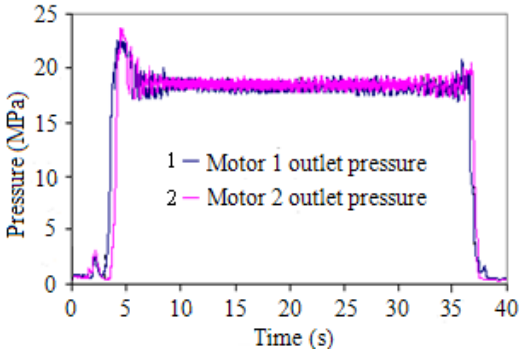


Figure 14 Fuzzy neural PID control

The results of above experiments show that the fuzzy neural PID control strategy has strong anti-interference ability, which greatly reduces the pressure difference between the two subsystems, further improves the control accuracy, and verifies the correctness of the simulation results.

6 Conclusion

In order to improve the precision of double motor synchronous control, the dynamic characteristics of the hydraulic lifting system are analyzed, and we find out the relationship among the parameters. By combining fuzzy control with neural network control, a fuzzy neural PID control strategy is proposed, which overcomes the shortcoming of the traditional PID control strategy lacking of adjustment on line. The simulation and experimental results show that the method has good dynamic response and robustness, high control precision, furthermore, it has certain value for applications.

Acknowledgement: Acknowledgement: This work was supported by the project of the Central Government Guides Local Science and Technology Development

Plans of Inner Mongolia (2022ZY0013), 2022 Autonomous Region "Grassland Talents" Young Innovative Talents Level 1 (2023QNCXRC04), 2022 Western Light Talent Training Program of the Organization Department of the CPC Central Committee "Western Young Scholars" (S24001).

References

[1] HUANG Liqin, LIU Rong, CHEN Ying. Research on pipeline pressure shock in valve control hydraulic positioning system [J]. Mechanical & Electrical Engineering Magazine, 2009,22(3):234-236.

[2] Favennec G, Alirand M. Optimal response of pressure reducer and stability influence of the downstream line dynamics [J]. Modena, Italy, 2002(7):1-10.

[3] S.Guo,L.Huang. Periodic oscillation for discrete-time Hopfield neural networks [J]. Physics Lettesr A,2004,329(3):199-206.

[4] Li Jun-wei, Zhao Ke-ding, Wu Sheng-lin. Re- search on Dual Electro hydraulic Motors Synch- ronization via Fuzzy Control [J]. Machine Tool & Hydraulics,2003,25 (1):115-116.

[5] ZHAO L,LIU X H,WANG T J. Influence of counterbalance valve parameters on stability of the crane lifting system[C]. International Conference on Mechatronics and Automation Xi'an, China, 2010.

[6] Lu Ren, James K. Mills,Dong Sun. Adaptive Synchronized Control for a Planar Parallel Manipulator [J]. Theory and Experiments, 2006, 128(4):976-979.

[7] K.Hirasawa, S.Mabu, J.Hu. Propagation and control of stochastic signals through universal learning networks [J]. Neural Networks, 2006,19(4): 487-499.

[8] Di Zhou,Tielong Shen,Katsutoshi Tamura.Adaptive Nonlinear Synch- ronization Control of Twin-Gyro Precession [J]. Journal of Dynamic Systems, Measurement and Control,2006,128(3):592-599.

[9] QU J Y,REN C B,et al.Parameters optimization method for variable displacement pump/motor and transmission of hydraulic braking energy regeneration system [J]. International Forum on Computer Science Technology and Applications,2009(3):19-22.

[10] S.Guo, L.Huang. Periodic oscillation for discrete-time Hopfield neural networks [J]. Physics Lettesr A, 2004,329(3):199-206.

[11] Kayacan E, Cigdem O, Kaynak O. Sliding mode control approach for online learning as applied to type-2 fuzzy neural networks and its experimental evaluation [J]. IEEE Transactions on Industrial Electronics, 2012,59(9):3510-3520.

[12] AHN K W, HYUN J H.Optimization of double loop control parameters for a variable displacement hydraulic motor by genetic algorithms [J]. JSME, International Journal Series C-Mechanical Systems Machine Elements and Manufacturing,2 005,48(1):81-86.

Viser Technology Pte. Ltd.
Mechanical Engineering Science
DOI: 10.33142/mes.v7i2.18369

Research Article Open Access

Research on Design Method of Dynamic Shop Floor Scheduling System Based on Human-computer Interaction

Songling TIAN[1], Zhuke CAI[2], Xiaoqiang WU[3], Xiaoqian QI[1*]

1. School of Control and Mechanical Engineering, Tianjin Chengjian University, Tianjin, 300384

2. Shuyunke (Tianjin) Technology Co., Ltd., Tianjin, 300192, China

3. College of Mechanical and Transportation Engineering, Ordos Institute of Technology, Ordos, Inner Mongolia, 017000, China

Corresponding Author: Xiaoqian QI, E-mail: qixiaoqian@tju.edu.cn

Abstract

The shop floor dynamic scheduling system based on human-computer interaction is the use of computer-aided decision-making and human-computer interaction to solve the dynamic scheduling problem. A human-computer interaction interface based on Gantt chart is designed, which can not only comprehensively and quantitatively represent the scheduling process and scheduling scheme, but also have friendly human-computer interaction performance. The data transmission and interaction architecture is constructed to realize the rapid response to shop floor disturbance events. A priority calculation algorithm integrating priority rules and dispatcher preference is proposed, which realizes the automatic calculation of priority for the dispatcher's reference and reduces their burden. A man-machine interactive shop floor dynamic scheduling strategy is proposed. When solving the dynamic flexible job shop scheduling problem caused by machine tool breakdown and urgent order, the origin moments obtained by using this strategy are 0.4190 and 0.3703 respectively. As can be seen from the origin moment indicator, the dynamic shop floor scheduling system based on the human-computer interaction is efficient and reliable in solving dynamic scheduling problems, and related strategies of this system are also feasible and stable.

Keywords: *Human-computer interaction; Dynamic scheduling; Flexible shop floor scheduling; Perturbation events*

1 Introduction

With the continuous deep integration and application of information technology and advanced manufacturing technology, the relationship between human and manufacturing system is undergoing great changes, and the manufacturing industry has also entered the era of intelligent manufacturing characterized by Intelligent IoT, virtual-real integration and human-machine integration. Scientific and reasonable human-machine cooperation mode is a powerful guarantee for the effectiveness of intelligent human-machine system. The merits and disadvantages of human-machine integration cooperation mode will directly affect people's decision-making judgment and behavior efficiency, and determine the security of intelligent human-machine system and the feasibility and completion of major decision-making tasks. CPS (Cyber-Physical systems) are the basic framework of an intelligent manufacturing system. Its implementation is to build a virtual manufacturing system that is parallel to the physical manufacturing system, and to optimize the physical manufacturing system through intelligent decision, scheduling and control. People are both the participating subject and the service subject of the intelligent manufacturing system. Therefore, to realize the application of the intelligent manufacturing system, it is necessary to solve the human-computer interaction problems under the CPS architecture, such as the design of the human-computer interaction shop floor scheduling system, the design of the human-computer interaction shop floor system monitoring, and the design of the human-computer interaction shop floor system operation and maintenance system.

Scheduling is part of the fundamental aspects of intelligent manufacturing systems and is very important for modern companies. The main methods for solving scheduling problems are artificial intelligence scheduling [1] and automated forms of dealing with such as simulation-based [2]. Artificial intelligence methods can solve common scheduling problem, the core of which is to use the powerful data computing power of computers and artificial intelligence algorithms to solve

Received on August 20, 2025; Accepted on September 30, 2025

mathematical models of scheduling problems. Meng Leilei et al [3] proposed a hybrid frog-hopping algorithm for the problem of minimizing the maximum completion time in a distributed flexible job shop. Shi Xiaoqiu et al [4] proposed an adaptive variable-level genetic weed algorithm for the flexible job shop scheduling problem with the objective of minimizing the maximum completion time. The artificial intelligence-based approach to scheduling decisions has a certain degree of intelligence and is particularly suitable for solving state exploding shop floor scheduling problems, but the reliability in solving complex scheduling problems is not ideal because the model usually needs to be simplified in the process of mathematical modelling. The simulation approach is very effective in designing and operating manufacturing systems, and it can be used as a support system for real-time scheduling of manufacturing systems. Jin Pengbo et al [5], in solving production rescheduling problems, proposed a decision model that incorporates data simulation, genetic optimization and BP neural networks for rescheduling methods. Cao Yuanchong et al [6] proposed a digital twin-based dynamic scheduling method for discrete assembly plants of complex products, which enable more accurate dynamic scheduling through data interaction between physical space and virtual space. However, there are some problems with the simulation-based real-time support system, for example, the simulation method takes too much time to run in some cases, and the simulation method built from one situation cannot be used in another, i.e. when the environment changes, the model needs to be rebuilt and re-tested to find the appropriate rule set according to the changed environment, which is not ideal for solving shop floor scheduling problems with many dynamic events The adaptability is not ideal when solving shop floor scheduling problems with many dynamic events.

Therefore, in order to give full play to the advantages of artificial intelligence algorithms in solving state explosion shop scheduling problems, as well as to bring into play the decision making ability, fast dynamic response ability and the ability to use knowledge and experience of schedulers, and also to improve the reliability and adaptability of the dynamic scheduling system of intelligent manufacturing shop, this paper focuses on the problem of designing the dynamic scheduling system of the shop based on human-computer interaction. The current research on HCI scheduling is still immature, mainly focusing on scheduling algorithms [7], HCI strategies [8] and HCI scheduling platform design [9]. In this paper, based on the existing domestic and foreign research, we further study the design method of HCI scheduling system and propose HCI scheduling strategies that can ensure efficient, safe and reliable operation of HCI scheduling system.

2 Problem Formulation

As the automation of shop floor system increases and the scale of production elements expands, the frequency of disturbing events such as machine failure, rush order and order cancellation in the production process gradually increases. This puts higher demands on the dynamics, real-time, reliability and stability of shop floor scheduling. At the same time, each disturbance event will consume a lot of energy for schedulers to adapt and modify the existing scheduling plan.

Traditional static shop floor scheduling solves three problems: process start time, machine tool assignment, and machining duration. On the other hand, traditional dynamic shop floor scheduling has to solve two more problems: the problem of scheduling task changes and the problem of resource changes such as machine tools, i.e., dynamic shop floor scheduling is completed by adding processes, deleting defective machines, updating the start and end times of unprocessed processes, and updating machine tool assignments.

In the process of dynamic scheduling of HCI Shop floor, schedulers need to give full play to the role of human in state perception, solution decision and system operation and maintenance, and make full use of the computer-aided generated scheduling solution to complete the shop floor scheduling. Therefore, the key to solve the dynamic scheduling problem of HCI shop floor is the design of the dynamic scheduling system of HCI shop floor.

The design of a dynamic scheduling system for human-computer interaction on the shop floor focuses on the following 3 issues.

(1) Design issues of graphical human-computer interaction interface. The human-computer interaction interface is the foundation for realizing human-computer interaction. The basic problem to be resolved is how to display the scheduling process and results, and how to adapt this interface to the scheduling operation.

(2) Standardization of human-computer interaction. The essence of human-computer interaction is information interaction, and only by adding relevant constraints to the information, the rationality of the scheduling scheme can be ensured, for example the process constraints. In addition, in order to ensure the practicality of the scheduling system, the order of human-computer interaction must be standardized, such as how the preferences of the scheduler are reflected in the priority, and how to arrange the sequence between interactive operations and priority updates.

(3) Priority calculation problem. When calculating the process priority, it is necessary to take into account the equipment selection preference set by the decision maker, the load of the machine and the urgency of the process, etc. This is a key issue to be solved to ensure the safety and reliability of the scheduling system.

3 Human-Computer Interaction Scheduling System Design Method

The scheduling problem is the core problem of HCI scheduling systems. In solving the scheduling problem, it is necessary to consider selecting a graphical tool suitable for HCI first, which can not only represent the scheduling process and results comprehensively and quantitatively, but also have friendly HCI performance. Secondly, efficient and accurate information interaction needs to be designed for the purpose of rapid response to shop floor disturbance events and proactive response to reschedule problems. Finally, it is necessary to propose a calculation method that can calculate priority according to priority rules and preference priorities, which can automatically calculate priorities for schedulers' reference and reduce the burden of schedulers.

3.1 HCI scheduling tool selection

Gantt charts can graphically represent the temporal and spatial relationships between production elements in the shop, i.e., the machine tools, processing sequence and duration of processes are plotted in a certain order by bars of different colors and lengths in a coordinate system consisting of machine tools and time. Gantt charts can express complex production data in a comprehensive and quantitative way, and they can also facilitate schedulers to add, modify and delete processes in a graphical way. Therefore, Gantt charts were chosen as the graphical tool for human-machine interaction.

Gantt charts can be used to represent the before-and-after relationship between processes of a workpiece in a process queue, or to represent the sequential relationship between processes on a machine. Table 1 shows the optimal scheduling results for the shop floor example, and figure 1 shows the queue distribution, where the numbers before and after "/" in Table 1 indicates the candidate machine number and its theoretical machining duration, and "or" indicates the parallel candidate machine number and its theoretical machining duration.

Table 1 6×6 machine tool flexible job shop scheduling problems

	Seq1	Seq2	Seq3	Seq4	Seq5	Seq6
Job 1	1or3/5or4	5or3or2/3or5or1	3or6/4or2	6or2or1/5or6or1	3/1	6or3or4/6or6or3
Job 2	2/6	3/1	1/2	2or4/6or6	6or2or1/5or6or1	
Job 3	2/6	3or6/4or2	6or2or1/ 5or6or1	3or2or6/4or6or6	1or5/1or5	
Job 4	6or2or1/5or6or1	2/6	3/1	5or3or2/3or5or1	3or6/4or2	
Job 5	5or3or2/3or5or1	6or2or1/5or6or1	2/6	1or3/5or4	2or4/6or6	3or2or6/4or6or6
Job 6	3or6/4or2	1/2	3or2or6/4or6or6	2/6	6or2or1/5or6or1	1or4/3or2
Job 7	6/1	1or4/3or2	3or2or6/4or6or6	2or5or1/6or1or6	3/1	
Job 8	3or6/4or2	3or2or6/ 4or6or6	6or2or1/5or6or1	2/6	2or4/6or6	
Job 9	6/1	1or5/1or5	6or3or4/6or6or3	1/2	3or2or6/4or6or6	2or4/6or6
Job 10	3or6/4or2	3or2or6/ 4or6or6	5or3or2/3or5or1	6/1	2or4/6or6	1or4/3or2

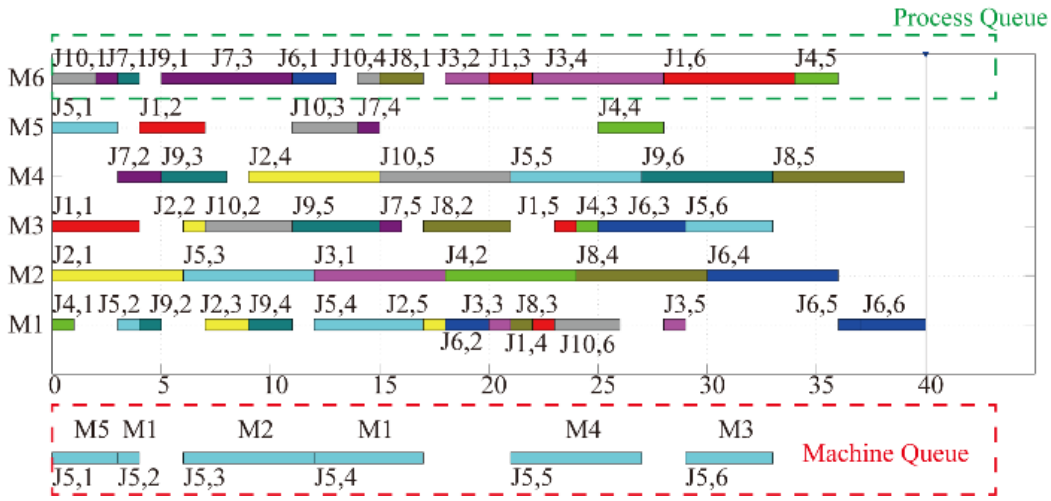


Figure 1 Scheduling results and queue representation

3.2 Data transfer and interaction function module

In the human-computer interaction shop floor dynamic scheduling system, the data transmission and interaction can be divided into 4 stages, as shown in Figure 2.

(1) When the scheduler encounters a disturbance event, he first interrupts the production process and obtains specific information about the disturbance time from the shop floor production monitoring module, as shown in ① and ② of Figure 2.

(2) The scheduler calculates and updates the priority of all unscheduled processes by setting preference coefficients in the process configuration area given the candidate machining machines for the process and their corresponding machining durations, as shown in③, ④, ⑤ and ⑥ of Figure 2.

(3) The scheduler refreshes the Gantt chart blocks of the processes in the priority queue and drags the Gantt chart process blocks in the priority queue area into the scheduling window, as shown in ⑦, ⑧, ⑨, and⑩ of Figure 2.

(4) After all processes are dragged into the scheduling window, the scheduler locks the scheduling results and downloads them to the shop floor controller to complete the scheduling, as shown in ⑪, ⑫, ⑬ and ⑭ of Figure 2.

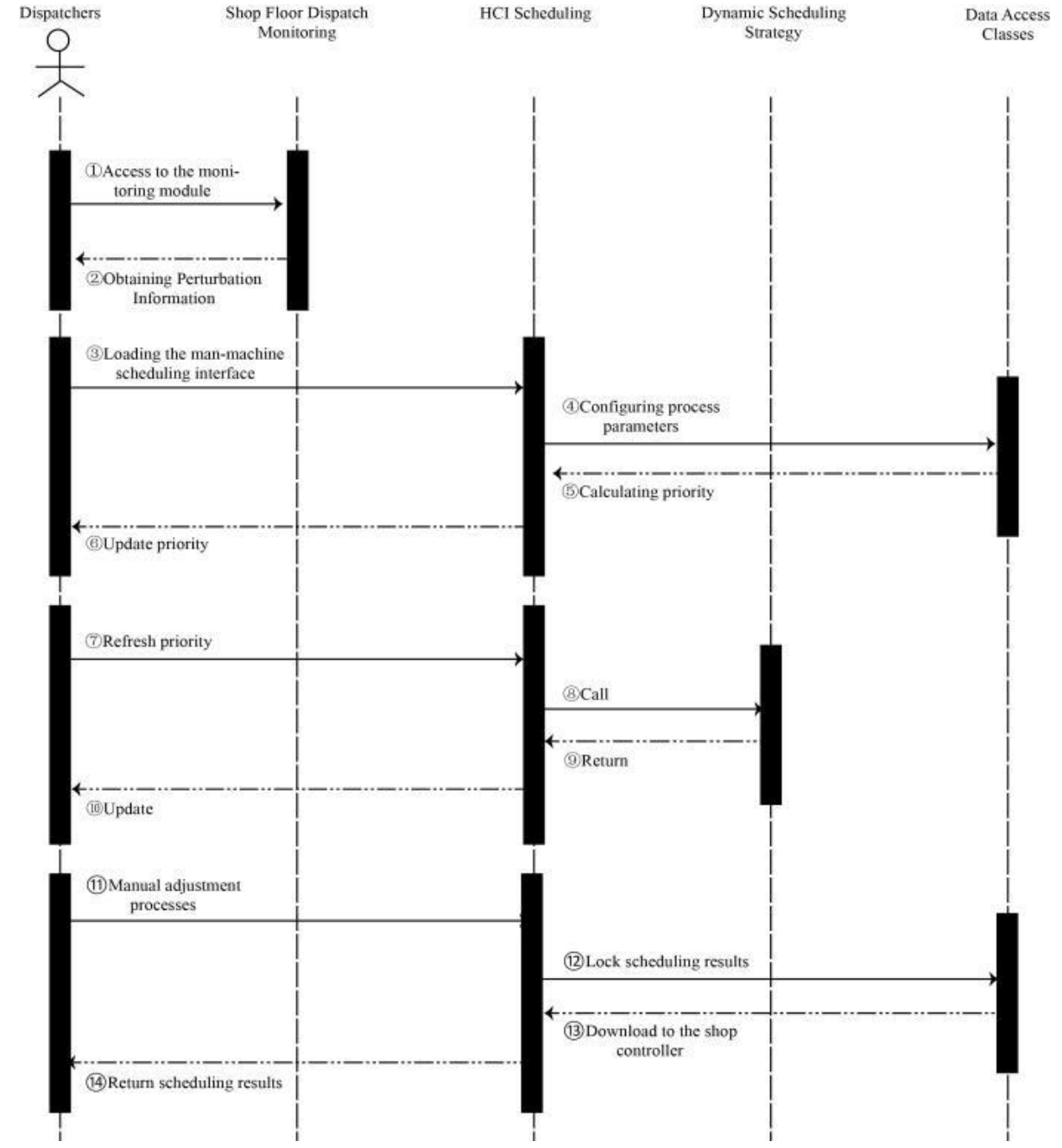


Figure 2 Data transmission and interaction process

3.3 Work order priority algorithm

The priority rule scheduling method is a traditional method for solving shop floor scheduling problems. Due to its simplicity, it is especially suitable for solving dynamic shop scheduling problems and real-time reproducible shop scheduling problems. The priority rule scheduling strategy is actually a greedy strategy, which mainly targets local or short-term objectives with high computational efficiency. Because of the large uncertainty of the HCI shop floor, it is the preferred strategy in many dynamic and real-time scheduling systems because of its feature of seeking relatively more desirable scheduling results with lower computational cost, although it cannot find the global optimal solution.

Shop floor scheduling needs not only to determine the process sequence problem of the process, but also to solve the problem of selecting processing equipment for the process. The priority rule scheduling strategy usually has the ability to solve both of these problems, and common shop scheduling priority rules are shown in Table 2.

Table 2 Workshop scheduling priority rules.

Rule	Full Name
SPT	The machine with the shortest processing time is preferred for the process
ECP	Priority is given to the machine that can complete the process at the earliest
MWR	Priority is given to workpieces with more remaining processes
LWR	Priority is given to workpieces with fewer remaining processes
MPR	Priority is given to workpieces with more processing time remaining
LPR	Priority is given to workpieces with little remaining machining time

Usually, it is difficult to satisfy the requirements of the shop floor scheduling with one scheduling rule. Therefore, priority rules can often be combined when solving shop floor scheduling problems. In the previous research results of this paper, a heuristic algorithm based on the combination of ECP and MWR rules was proposed, and the performance of the solution was better [10]. Therefore, the heuristic algorithm built on the combination of ECP and MWR rules is chosen as the basis of the process prior calculation algorithm in this paper.

In solving the actual dynamic scheduling problem, in order to give full play to the role of human in state sensing, solution decision making and system operation and maintenance, the balance between the decision maker's equipment selection preference and task completion needs to be considered. In order to weigh the two factors, the weights of them need to be given to the calculation of the integrated priority, and the formula for the priority.

$$S = \omega_1 \times S_S + \omega_2 \times S_P \tag{1}$$

Where S is the composite priority value, SS and SP are the priority rule score and personal preference score, respectively, ω1, ω2 are the priority rule score weight and personal preference score weight, respectively, and all are subject to normalization.

4 Dynamic Scheduling Strategy for Human-Computer Interaction Shop Floor

The operational state of shop floor can be modeled by two virtual queues. One is the process queue, which is generally generated by priority rules; the other is the machine queue, which is generally generated based on the part machining process flow. The process queue is used to simulate the processes of each workpiece, which are arranged in order of priority in the cache of the respective assigned machine, with the process with the highest priority at the top of the queue and the first to be processed, called the head of the queue. And the machine queue is used to simulate the machine tools assigned to each process of a workpiece arranged in accordance with the process flow on the processing line of the workpiece, thus forming a queue. Therefore, the essence of smooth shop operation is that the two virtual queues are executed according to the established arrangement, the essence of shop operation status monitoring is whether the execution is the same as the established arrangement, and when the shop encounters an abnormal event disturbance, the members of the whole queue and its order can be adjusted to ensure that the shop operation can be restored to another steady state.

The basic principle of using virtual queues to achieve dynamic scheduling of HMIs is illustrated in figure 3. ①The scheduler identifies the processes affected by abnormal events and rearranges the process and machine queues according to the automatically calculated priority values.② The scheduler arranges the process at the head of the queue at the earliest of the scheduling reference time, the completion time of the previous process on the same workpiece, and the completion time of the aforementioned process on the same machine. ③ The scheduler drags the process at the head of the queue into the scheduling window and removes the process from the process queue, and the process with the highest priority in the process queue is automatically added as the process at the head of the queue. ④ After all processes have been foreseen, the scheduling plan is locked and downloaded to the shop floor system controller.

Take the case shown in figure 3 as an example, according to the base time of the scheduling window, at the current moment when the workpiece 9 is transferred to the workpiece's next process machine 3, according to the shop machine queue information machine 3 is in a fault state, at this time the workpiece 9 is in a waiting

state at machine 3; when the workpiece 8 is transferred to the workpiece's next process machine 6, according to the machine queue information machine 3 is in a fault state, at this time the workpiece 8 is in a processing state at machine 6; when the workpiece 4 is transferred to the workpiece's next process machine 2, according to the machine queue information machine 2 is in a processing state, at this time the workpiece 4 is in a waiting state at machine 2.

A demonstration of the process of implementing a dynamic strategy for human-machine interaction is illustrated in figure 4.

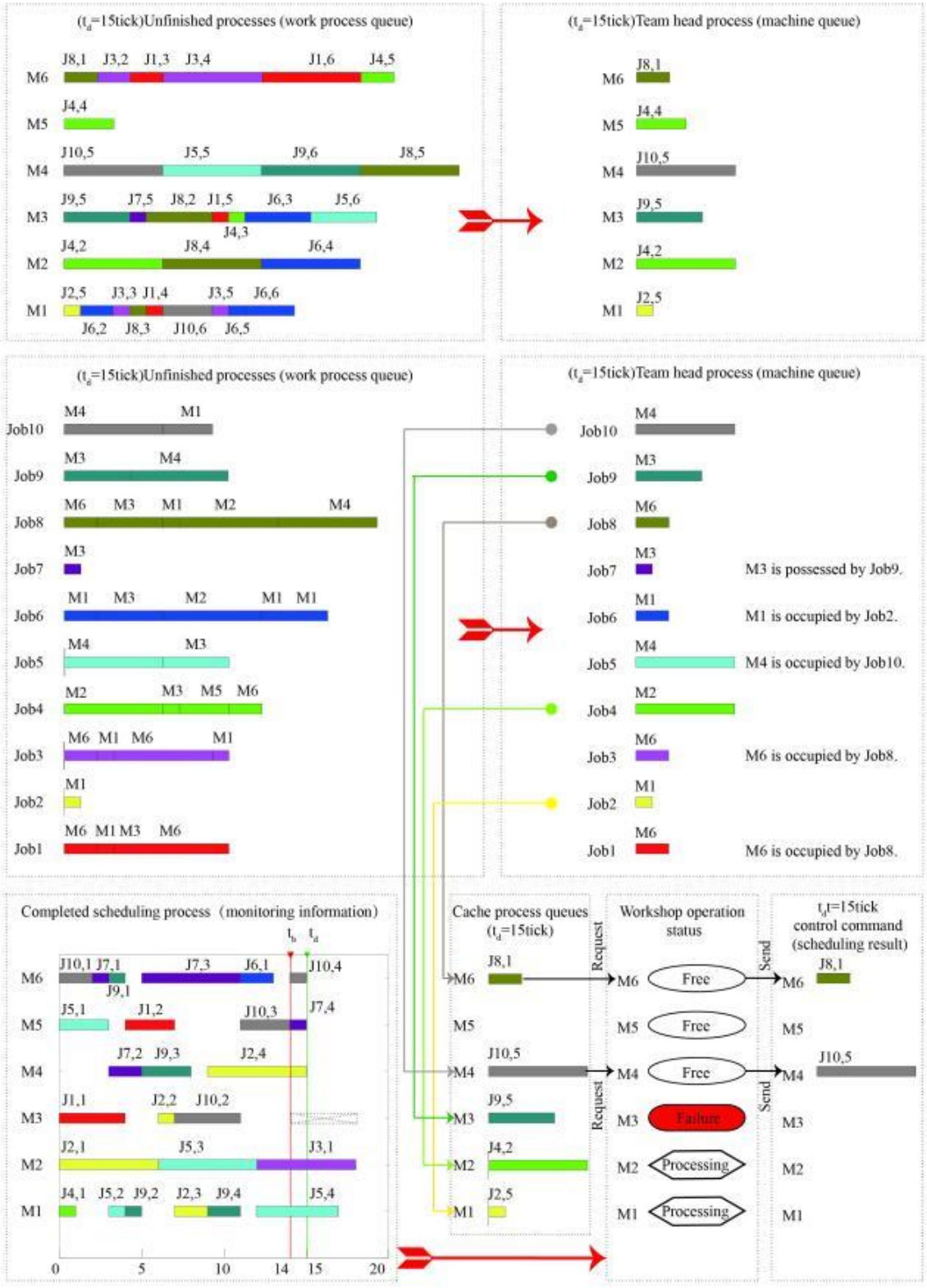
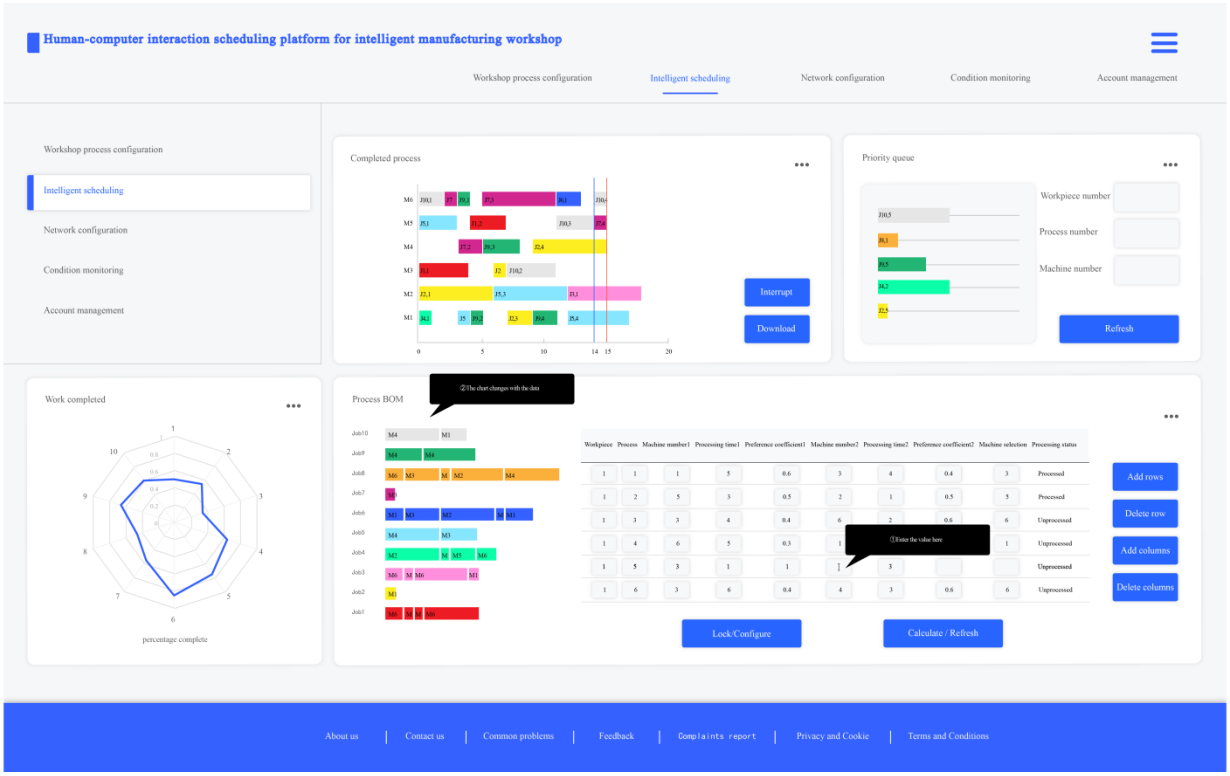
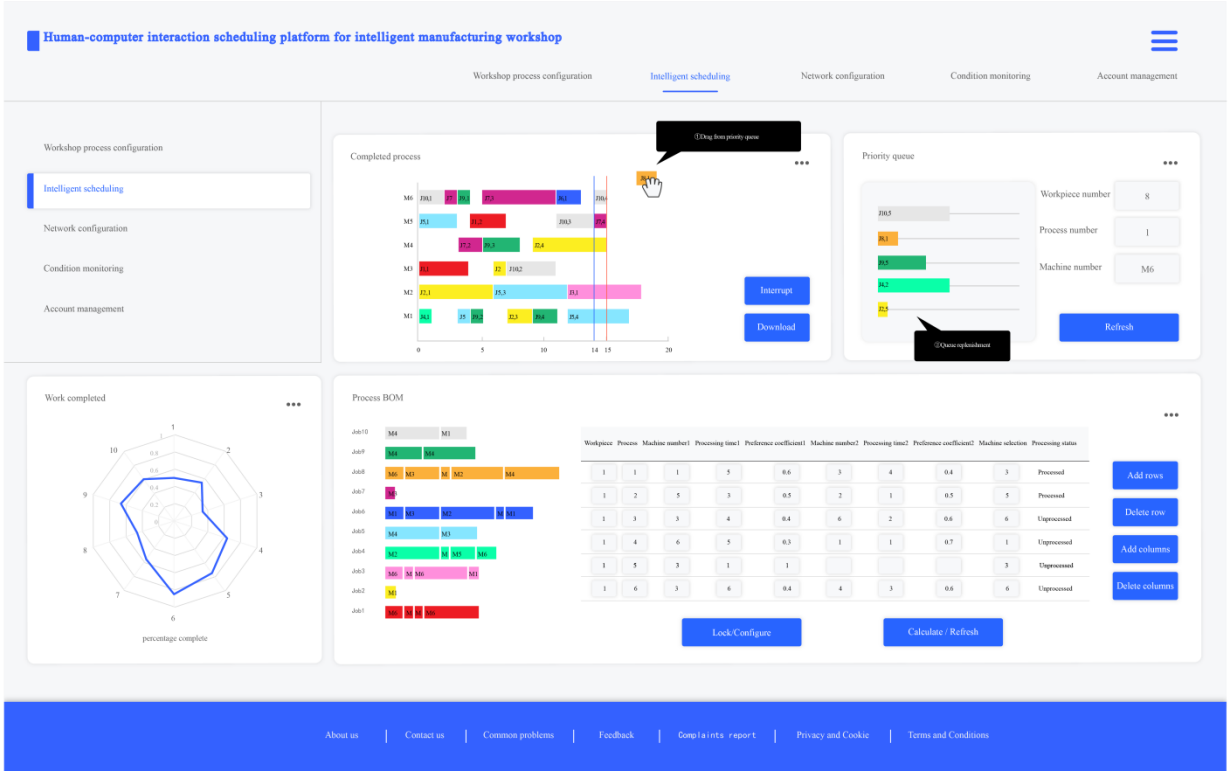


Figure 3 Shop floor information interaction



(a)



(b)

Figure 4 The process of implementing dynamic human-computer interaction

The scheduler first interrupts the production process when a disturbance event is raised. Candidate machines for the process and their corresponding processing duration are given in the process configuration area, and the priorities of all unscheduled processes are calculated and updated by setting preference factors, as showed in figure 4(a). Then the Gantt chart processes blocks of the processes in the priority queue are refreshed. Next, the scheduler drags the Gantt process blocks in the priority queue into the scheduling window,

as showed in figure 4(b). Finally, after all processes are dragged into the scheduling window, the scheduling results are locked and downloaded to the shop floor controller to complete scheduling.

5 Case and Application Analysis

In order to verify the feasibility of the dynamic scheduling strategy of the HCI shop floor and to test the performance of the HCI shop floor dynamic scheduling system in solving scheduling problems, the computer CPU main frequency is 2GHz and memory is 2.0GB under Windows 7 environment, and MATLAB programming is used to choose to simulate machine failure and expediting order on the standard flexible job shop scheduling arithmetic MK06 [11] scenarios. The origin scheduling scheme is shown in figure 5, and the meaning of the elements in the graph is the same as in Section 2.1.

To measure the results of scheduling, the origin moment metric is chosen. By examining the origin moments for robustness and stability metrics, values closer to zero indicate ideal results for dynamic scheduling [12]. The origin moment is calculated as following:

$$MID = \sqrt{Robustness^2 + Stability^2} \tag{2}$$

In the formula, Robustness and Stability denote robustness and stability, respectively.

The robustness metric is the relative robustness metric proposed by Kouvelis and Yu [13]:

$$RM = \frac{|MS\_R - MS\_P|}{MS\_P} \times 100\% \tag{3}$$

The stability indexes proposed by Al-Hinai and Elmekkawy [14] were used:

$$SM = \frac{\sum_{i=1}^{n}\sum_{j=1}^{q_i}|CO_{ijP} - CO_{ijR}|}{(\sum_{i=1}^{n} O_i) \times MS\_P} \times 100\% \tag{4}$$

Where *MS_P* is the planned completion time for pre-scheduling, qi is the total number of processes for workpiece i, COijP is the planned machining completion time for the jth process of workpiece i, COijP is the actual machining completion time for the jth process of workpiece i, and Oi is the number of all processes that do not change machining machines (excluding faulty machines).

The solution obtained for the dynamic scheduling problem caused by machine fault disturbance is shown in figure 6. By comparing the schemes of figure 5 and figure 6, the RM index is calculated to be 0.1167, the SM index is calculated to be 0.4025, and the origin moment is calculated to be 0.4190 using Eq. (2).

For the dynamic scheduling problem caused by the expedited single disturbance, the solution obtained is shown in figure 7. By comparing the schemes in figure 5 and figure 7, the RM index is calculated to be 0.1167, the SM index is calculated to be 0.3514, and the origin moment is calculated to be 0.3703 using Eq. (2).

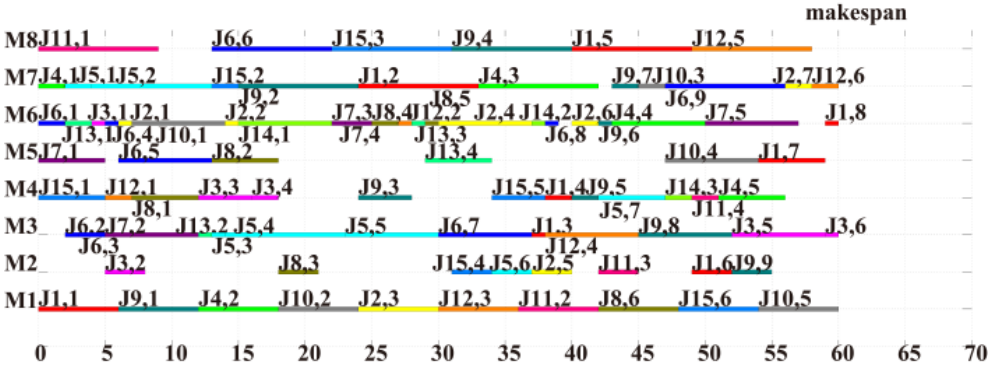


Figure 5 Original scheduling scheme

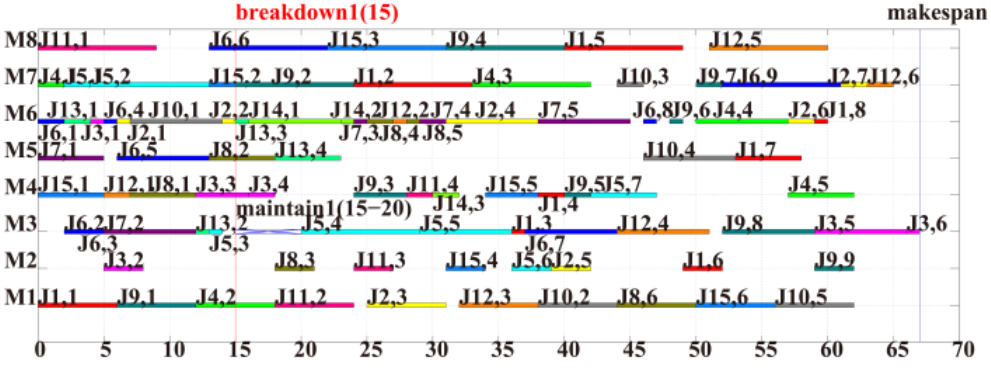


Figure 6 Scheduling scheme for dynamic scheduling problems caused by machine fault disturbances

Figure 7 Scheduling scheme for dynamic scheduling problems caused by expedited order disturbances

Therefore, it is clear from the origin moment index that the HMI-based shop floor dynamic scheduling system is efficient and reliable in solving dynamic scheduling problems, and its related strategies are feasible and stable.

6 Conclusions

The shop floor dynamic scheduling system with human-computer interaction implements the shop floor dynamic scheduling through computer-aided decision making and human-computer interaction. In order to solve the design problem of graphical human-computer interaction interface, the standardization problem of human-computer interaction and the priority calculation problem in the system design, a human-computer interaction interface based on Gantt chart is designed, a data transmission and interaction architecture is constructed, and a priority algorithm that integrates priority rules and scheduler preferences is proposed. The human-computer interaction shop floor dynamic scheduling strategy is proposed, and the origin moments obtained by using it in solving the flexible dynamic operation shop floor scheduling problem caused by machine failure and expediting order are 0.4190 and 0.3703, respectively. From the origin moment index, it can be seen that the human-computer interaction-based shop floor dynamic scheduling system is efficient and reliable in solving the dynamic scheduling problem, and its related strategy is feasible and stable.

Further research will continue to focus on the integration of reality and information interconnection in the process of human-computer interaction scheduling.

Acknowledgments: This work was supported by the Tianjin Enterprise Science and Technology Commissioner Project (Grant No. 23YDTPJC00740, Grant No. 24YDTPJC00610) and the Tianjin Tiankai Higher Education Science and Technology Innovation Park Enterprise R&D Special Project (Grant No. 23YFZXYC00027).

Competing Interests: The authors declare no competing interests.

References

[1] Fan, C., Wang, W., & Tian. Flexible job shop scheduling with stochastic machine breakdowns by an improved tuna swarm optimization algorithm [J]. Journal of manufacturing systems, 2024 (74):180-197.

[2] Zheng, T., Zhou, Y., Hu, M., & Zhang. Dynamic scheduling for large-scale flexible job shop based on noisy DDQN [J].International Journal of Network Dynamics and Intelligence, 2023 (45):100015-100015.

[3] Cai, J., Lei, D., Wang, J, et al.. A novel shuffled frog-leaping algorithm with reinforcement learning for distributed assembly hybrid flow shop scheduling [J].International Journal of Production Research, 2023,61(4):1233-1251.

[4] Li, L., Deng, L., Tang, B., et al. Bi-objective flexible job shop scheduling on machines considering condition-based maintenance activities [J]. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 2024,238(6):1244-1255.

[5] Wu, R., Zheng, J., & Yin, X. Dynamic Scheduling for Multi-Objective Flexible Job Shops with Machine Breakdown by Deep Reinforcement Learning [J]. Processes,2025,13(4):1246.

[6] Nisar, M. U., Ma'ruf, A., & Halim, A. H. In Proceedings of the International Manufacturing Engineering Conference & The Asia Pacific Conference on Manufacturing Systems [C]. Singapore: Springer Nature Singapore,2022.

[7] Luo, W., Huang, K., Liang, X., et al. Process manufacturing intelligence empowered by industrial metaverse: A survey [J]. IEEE Transactions on Cybernetics, 2024(42):89-91.

[8] Subramanian, K., Thomas, L., Sahin, M., et al. Supporting human–robot interaction in manufacturing with augmented reality and effective human–computer interaction: a review and framework [J]. Machines, 2024,12(10):706.

[9] Agostinelli, T., Generosi, A., Ceccacci, S., et al. Validation of computer vision-based ergonomic risk assessment tools for real manufacturing environments [J]. Scientific Reports,2024,14(1):27785.

[10] Tian, S., Wang, T., Zhang, L., et al. The Internet of Things enabled manufacturing enterprise information system design and shop floor dynamic scheduling optimisation [J]. Enterprise Information Systems,2020,14(9-10):1238-1263.

[11] Zambrano-Rey, G. M., González-Neira, et al. Minimizing the expected maximum lateness for a job shop subject to stochastic machine breakdowns [J]. Annals of Operations Research,2024,338(1):801-833.
[12] Azim, N. H., & Subki, A. Abiotic stresses induce total phenolic, total flavonoid and antioxidant properties in Malaysian indigenous microalgae and cyanobacterium [J]. Malaysian Journal of Microbiology, 2018,(11):25-33.
[13] Kouvelis, P., & Yu, G. Robust discrete optimization and its applications [D]: Springer Science & Business Media,2013.
[14] Fan, C., Wang, W., & Tian, J. Flexible job shop scheduling with stochastic machine breakdowns by an improved tuna swarm optimization algorithm [J]. Journal of manufacturing systems,2024,(74):180-197.

Viser Technology Pte. Ltd.
Mechanical Engineering Science
DOI: 10.33142/mes.v7i2.18370

Research Article Open Access

The Research on CAD Design System of Shaper Cutter Based on VB and Matlab

Xiaoqiang WU[1*], Rui XUE[1], Kan XING[2], Tiegang WANG[3], Peng WANG[3], Fenghe WU[4], Yabin GUAN[4], S. ZHANG[1]

1. Tianjin TANHAS Technology Co., Ltd., Tianjin, 301600, China
2. Tianjin No.1Machine Tool Co., Ltd., Tianjin, 300305,China
3. Tianjin University of Technology and Education, Tianjin, 300222, China
4. Yanshan University, Qinhuangdao, Hebei, 066000, China

[*]***Corresponding Author:*** *Xiaoqiang WU, E-mail: wangzai8402@163.com*

Abstract

To solve the problem of low efficiency in the design of shaper cutter, a design and calculation system of shaper cutter is developed by using Matlab and VB mixed programming. The structure parameters of shaper cutter are calculated by using Matlab, and the system interface is developed by VB. The example shows the feasibility of the design method.

Keywords: *Shaper cutter, Matlab; VB*

1 Introduction

Gear as the basic parts in the entire machinery industry are widely used in aerospace, automotive industry, machine tools, engineering machinery and other fields. At present the main gear processing method is the generating method, this method is based on the principle of gear meshing to process the gear, including the gear shaping, hobbing, planing, shaving, grinding and other methods. As the most commonly used involute cylindrical gear machining method, the gear shaping method ranks only second to the hobbing method in the entire gear processing system, and is widely used for machining internal gears, racks, herringbone gears, gears with shoulder and Dual or multiple gears with slot cutter groove References are cited in the text just by square brackets With the improvement of productivity, the shortcomings of the traditional design method of shaper cutter are more prominent. But the development of CAD / CAM technology provides the conditions for rapid design. So in recent years, more and more researchers have begun to pay attention to the fast design system for shaper cutter. With the help of the computer's powerful computing ability, the shaper cutter design method become more flexible, and the design efficiency is greatly improved, but the manual calculation error is greatly reduced. In this paper References are cited in the text just by square brackets [1], a practical method of involute shaper cutter CAD is proposed, which realizes the professional design of involute shaper cutter. Paper[2] starts from the design theory of shaper cutter and hobbing cutter, proposed a overall design of gear cutter CAD system, based on the Visual Basic as a development platform, developed a gear cutter design system, and an example is given to demonstrate the correctness and feasibility of the design method.

To design a shaper cutter, the most difficult part is to calculate maximum and the minimum displacement coefficient of the shaper cutter. Matlab as a good mathematical software, with efficient numerical calculation and symbolic computing functions, as well as the complete graphics processing functions, so the calculation of the displacement coefficient of this part is mainly completed by Matlab References are cited in the text just by square brackets [3]. But Matlab also has its own limitations, in general, the program finished by Matlab can not run apart from the software alone, and its interface production function is relatively weak. Visual Basic as a good visual programming software, can easily create a friendly user interface, with an open object-oriented architecture, and can easily use other applications to provide the function References are cited in the text just by square brackets [4]. So this paper combines the two software, using Visual Basic to

Received on June 5, 2025; Accepted on July 28, 2025

develop the interface of the design system, the use of Matlab as the core of the data processing, and finally developed a user-friendly, powerful gear cutter CAD design software References are cited in the text just by square brackets [5].

2 Principle of Shaper Cutter

2.1 Principle of gear shaping

It is well know that the gear shaping process is achieved through the generating principle, as shown in figure 1, in the generating process, the main movement of the gear shaping is the shaper cutter reciprocating along the tool axis. The machined gear is engaged with the shaper cutter to generate the involute tooth profile.The engagement moment can be decomposed into a circular feed movement and a dividing movement. the transmission ratio equals with the gear engagement.

When shaping the internal teeth, the cutter and the workpiece rotating in the same direction of rotation, and the direction is opposite when shaping the external teeth. The shaper cutter should be relieved during the escape motion to reduce the friction between the cutter and the tooth surface. Finally, the workpiece should have the radial feed movement until the workpiece is cut into the desired depth.

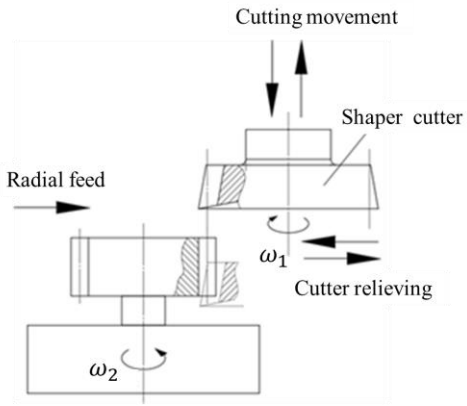


Figure 1 Principle of gear shaping

2.2 The structure of shaper cutter

The geometric structure has almost never changed since the shaper cutter was invented. Here use the helical gear shaper cutter as an example to illustrate the main structure of the shaper cutter. As shown in figure 2, a cutter tooth is composed by two side blades, a top edge, a rake face, two side flanks, and a top flank face.

Due to the presence of the relief angle, the radius of the top edge of and the tooth thickness at the reference circle are continuously reduced in the end profile from the rank face. Currently, when designing a shaper cutter, the tooth profile at the original section is first constructed. The top edge radius and the tooth thickness at the reference circle in the tooth profile are the standard values.

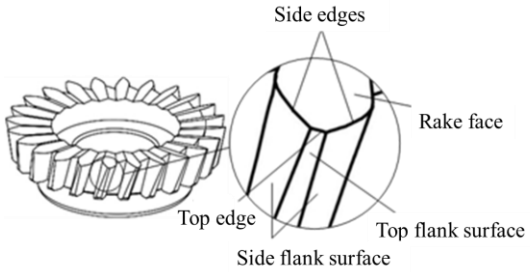


Figure 2 The Structure Of Shaper Cutter

The tooth profile in each end profile from the original section to the rank face is equivalent to the tooth profile of the gear with positive correction. And the modification coefficient gradually increases from the original section to the initial rank face. The tooth profile on the cross section from the original section to the last sharpening end face is equivalent to the tooth profile of the gear with negative correction. And the modification coefficient gradually decreases from the original section to the end face. as shown in figure 3.

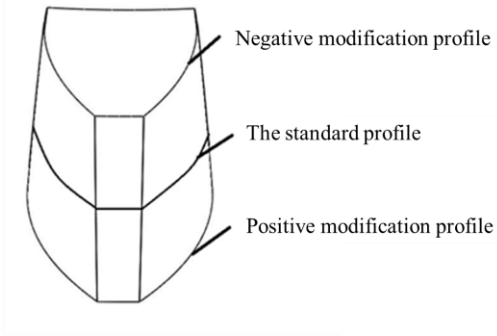


Figure 3 The modification diagram of the shaper cutter

3 Design Steps of Shaper Cutter

When designing the shaper cutter, the first step is to calculate the design parameters. The design parameters are composed of the basic parameters and structure parameters. The basic parameters are determined according to the workpiece parameters. The workpiece parameters are shown in Table 1:

Table 1 The Basic Parameters

	Workpiece	Match gear	Cutter gear
Modulus	m1	m2	m
Pressure angle on pitch circle	α_1	α_2	α
The original modification coefficient	x1	x2	x
Aaddendum coefficient	h_{a1}^*	h_{a2}^*	h_a^*
Tip clearance coefficient	c_1^*	c_2^*	

According to the basic design requirements of shaper cutter, m=m1=m2, α=α1=α2,h *α1=h* α2,c*

1=c* 2. The original modification coefficient shown in Table 1 refers to the modification coefficient of the workpiece. If the workpiece is a gear with addendum modification, the original modification coefficient of the shaper cutter is equal to the original modification coefficient of the cutter gear according to the principle of gear shaping,which is x=x2=-x1. As the shaper cutter teeth should cut out the workpiece tooth profile and the tooth top clearance, so the addendum of the shaper cutter should be higher than the standard involute gear tooth addendum, then the addendum coefficient of the cutter gear is:

$$h * \alpha=,h *\alpha 1+c* 1 \quad (1)$$

The main structural parameters of the design of the shaper cutter are shown in Table 2:

Table 2 The Structural Parameters

Number Of Teeth	Rake angle of the top edge[°]	Relief angle of the top edge[°]	The max modification coefficient	The min modification coefficient
Z	γ	α_e	x_{max}	x_{min}

The parameters shown in the table are the basis for the design of the shaper cutter, and are all indispensable. They must be accurately calculated in order to ensure the correctness of the final shaper cutter.

3.1 The cutter teeth

Currently, to determine the number of cutter teeth on the shaper cutter, the first step is to select the nominal diameter of the reference circle, China's standard GB / T6081-2001 regulate the nominal diameter of 75mm, 100mm, 125mm, 160mm, 200mm. The most commonly used nominal diameter are 75mm, 100mm and 125mm. The main reason for selecting the standard nominal reference circle diameter is to standardize the shaper.

In general, the diameter of the shaper cutter should be as large as possible under the allowable conditions of the gear shaping machine and the gear grinding machine, which can effectively improve the cutter life and reduce the interference of the transition curve. But this will increase the cutter manufacturing cost. In addition, due to the current shaper cutter manufacturing process mainly rely on the generating gear grinding method, with the restriction of the involute cam plate on the grinding machine, extra checking should be involved according to involute cam plate parameters. After choose nominal reference circular diameter, the number of teeth on the shaper cutter:

$$Z=d' 0 /m_1 \quad (2)$$

Z should be rounded, m1 means the modulus of the machined gear, and d' 0 means the diameter of the reference circle.

3.2 The rake angle and relief angle of the top edge

The reasonable rake angle and relief angle is necessary to ensure the normal cutting of the shaper cutter. According to the structure of the shaper cutter, the rake angle on the side edge will increase when the rake angle on the top edge is increased, and the increase of the rake angles could effectively improve the surface quality of the workpiece.

There is a relief angle on the top edge of the shaper cutter, and two side flank faces on each tooth are actually two involute spiral surfaces which in the same angle, but the opposite direction. And this structure makes the cutter has the side edge relief angle. Studies have shown that increasing the relief angle of the shaper can effectively increase the durability of the shaper cutter.

However, for the current shaper cutter, increasing the rake and relief angle of the shaper cutter will lead to increased tooth profile error, reduce the number of sharpening. Therefore, the current design method stipulates that the rake angle on the top edge is 5 °and the relief angle is 6 °.

3.3 The modification coefficient

The maximum and minimum modification coefficient of the shaper are limited by a number of factors, so it is necessary to consider these factors to determine the maximum and minimum modification coefficient of the shaper cutter.

The maximum modification coefficient xmax. Increasing the maximum modification coefficient of the shaper cutter, the number of sharpening of the shaper cutter can be increased, and the tool life can be improved effectively. At the same time, the addendum of the shaper cutter under the large modification coefficient is higher, the curvature radius of the cutting edge is larger, so that the processed gear has the small surface residual and high surface quality. However, the excessive increase in the maximum modification factor of the shaper cutter will cause the addendum to become sharp and reduce the tool life, and the gear processed by the large modification shaper is easy to have the transition curve interference problem. So it is necessary to check these two issues when choosing the maximum modification coefficient.

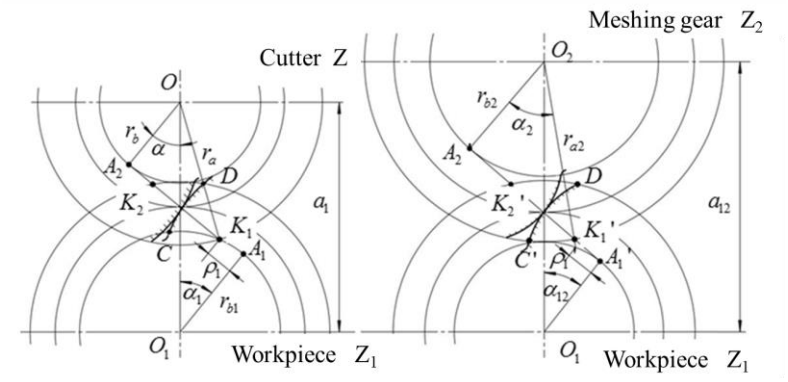
According to the current shaper cutter design theory, the actual addendum width Sa of the shaper should not be less than the minimum allowable width [Sa], but [Sa] is determined according to the shaper cutter modulus m, the tip strength, the cooling capacity and the tool durability, here is the empirical equation:

$$[Sa]=-0.0107m^2+0.2643m+0.3381 \quad (3)$$

The actual addendum width:

$$s_a = \left[\frac{\pi + 4x_{max1} \tan\alpha}{z} + 2\left(inv\alpha - inv\alpha_a\right) \right] r_a \quad (4)$$

αa means the pressure angle on the addendum circle. According the equation (4), the maximum modification coefficient xmax1 could be figure out.



(a)the shaper cutter and the workpiece (b)the workpiece engaging with the gear

Figure 4 Checking Of The Transition Curve Interference When Machining The Workpiece

The transition curve interference may occur between gears machined with positive modification shaper cutter. As shown in figure 4, when the tooth is machined, only the CD segment on the tooth profile of the workpiece is involute tooth, and the line between point C and dedendum is not the involute but the transition curve. The profile of the transition curve is actually the trajectory of the top edge and the side edge of the cutting blade relative to the workpiece. The transition curve is tangent to the involute tooth profile at point C, and the point C is the starting point of the involute on the tooth profile, but also the starting point of the transition curve.

The curvature radius of the involute at the point C ρ1 equals to the line A1K1, and can be calculated by the equation:

$$\rho_1 = A_1A_2 - K_1A_2 = a_1 \sin\alpha_1 - \sqrt{r_a^{\,2} - r_b^{\,2}} \tag{5}$$

After careful analysis, the greater of radius curvature is, the higher starting point of the transition curve has.

When the finished workpiece z1 engaging with the gear z2, as shown in figure 4 (b), the effective meshing line is K' 1K' 2, and the working part of the tooth profile on z1 is C'D, that is, above the point C', the tooth profile should be the involute to ensure the normal engagement. If the starting point C is lower than the point C, the gear meshing can be carried out normally, otherwise the transition curve interference problem occurs.

It is necessary to ensure that CD ⩾ C'D to avoid the interference between the tooth root transition curve of the workpiece with the mating gear, which is the curvature radius ρ1' larger than ρ1. To express this relation in a different form:

$$a_{12}\sin\alpha_{12} - \sqrt{r_{a2}^{\,2} - r_{b2}^{\,2}} \geq a_1\sin\alpha_1 - \sqrt{r_a^{\,2} - r_b^{\,2}} \tag{6}$$

α1 is the engagement angle during the gear shaping, a1 the center distance when the final radial feed is completed during gear shaping, α_{12} and a_{12} is the engagement angle and the matching center distance between the worlpiece and the engaging gear. α1 and a1 could be figure out by the following equations:

$$inv\alpha_1 = \frac{2\left(x_1 + x_{\max 2}\right)\tan\alpha}{z_1 + z} + inv\alpha \tag{7}$$

$$a_1 = \frac{m\left(z_1 + z\right)}{2} \cdot \frac{\cos\alpha}{\cos\alpha_1} \tag{8}$$

In the equation (7), x1 is the original modification coefficient of the workpiece.

According to equation (7), a maximum modification coefficient xmax2 can be obtained under the condition that no transition curve interference occurs.

Finally, to compare xmax1 with xmax2, the small one is the final modification coefficient the shaper cutter should have.

The minimum modification coefficient xmin. With the grinding of the shaper cutter, the modification coefficient could be gradually reduced, in general, when the tooth thickness is ground to the final allowance, the ideal state is the modification coefficient is the minimum permissible coefficient xmin. In this way, the shaper cutter can be ground most times and have the longest life. But when the shaper cutter is in the negative modification, the tooth thickness become thinner, the tooth strength can not fulfill the requirement, may also lead to the undercutting or end cutting.

In the process of gear shaping, if the modification coefficient of the cutter is too small, as shown in figure 5, the tooth tip of the shaper may be cut into the involute profile of the workpiece, this phenomenon is called undercutting, so when designing the shaper cutter, the checking of the minimum modification coefficient to prevent from undercutting is very important.

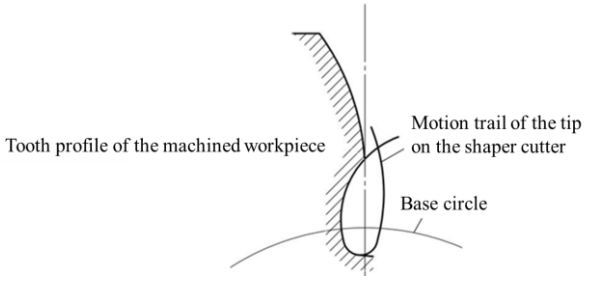


Figure 5 Diagram Of The Undercutting

According to the requirements mentioned above, ρ1 ⩾0,which is:

$$a_1 \sin\alpha_1 - \sqrt{r_a^{\,2} - r_b^{\,2}} \geq 0 \tag{9}$$

$$inv\alpha_1 = \frac{2\left(x_1 + x_{\min 1}\right)\tan\alpha}{z_1 + z} + inv\alpha \tag{10}$$

$$a_1 = \frac{m\left(z_1 + z\right)}{2} \cdot \frac{\cos\alpha}{\cos\alpha_1} \tag{11}$$

According to the equation (9) and the parameters of the shaper cutter and the workpiece, the undercutting phenomenon can be checked with the selected minimum modification coefficient. If the phenomenon still exist, increase the modification coefficient and check it again until the phenomenon do not happens anymore, in this way, the minimum modification coefficient xmin1 can be determined.

When the modification coefficient of the shaper cutter is small, the tooth height of the shaper cutter is small too, in the gear shaping process, the tooth root of the cutter may be interfere with the addendum of the machined workpiece, and the tip of the workpiece could be cut down, and this phenomenon is called end cutting.

As is shown in figure 6, to prevent the end cutting phenomenon, one condition need to be achieved,O1K1⩽O1A2, which is:

$$r_{a1} \leq \sqrt{\left(a_1 \sin\alpha_1\right)^2 + r_{b1}^{\,2}} \tag{12}$$

In the case of no end cutting:

$$\tan\left(\alpha_1\right)_{\min} \geq \frac{2\sqrt{r_{a1}^{\,2} - r_{b1}^{\,2}}}{m\left(z_1 + z\right)\cos\alpha} \tag{13}$$

According to the gear meshing principle:

$$inv\alpha_1 = \frac{2\left(x_1 + x_{\min 2}\right)\tan\alpha}{z_1 + z} + inv\alpha \tag{14}$$

In the case of no gear end cutting:

$$x_{\min 2} = \left[inv\left(\alpha_1\right)_{\min} - inv\alpha\right] \cdot \frac{z_1 + z}{2\tan\alpha} - x_1 \tag{15}$$

According to the equation (15), the minimum allowable modification coefficient xmin2 of the shaper cutter can be obtained without the end cutting.

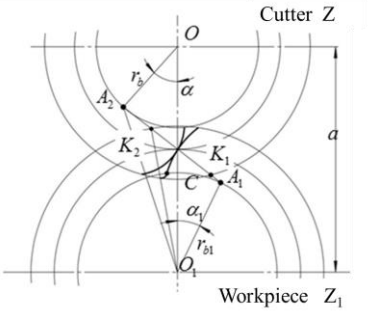


Figure 6 Diagram Of The End Cutting Checking

Finally, compare the xmin1 with xmin2, and choose the bigger one as the minimum modification coefficient xmin.

In summary, the main structure of the shaper cutter can be obtained after the calculation of all the parameters.

4 CAD System of Shaper Cutter

From the above, the shaper cutter design needs to calculate a series of parameters, and the parameter calculation process is quite complex, the manual calculation prone to errors, resulting in the final design failure. So in order to ensure the correctness and efficiency of the design results, the application of Matlab programming to finish the basic parameters calculation. At the same time, to develop the shaper cutter design system, Matlab interface production function is weak and can not run away from the Matlab environment, which is pretty inconvenient, so this article seeks a reliable way to solve this problem, that is, using Matlab and VB mixed programming way to achieve the function of the software.

4.1 System function module

Since the functional structure of the design system must be clear and reliable, this design of the system using modular design ideas, the design system is divided into three modules, human-computer interaction interface, design and calculation and the results display.

Human-computer interaction interface. For a design software, a good human-computer interaction interface is necessary. A good user interface should be simple and clear, so most of the man-machine interaction modules described in this chapter use a concise graphical user interface, as is shown in figure 7:

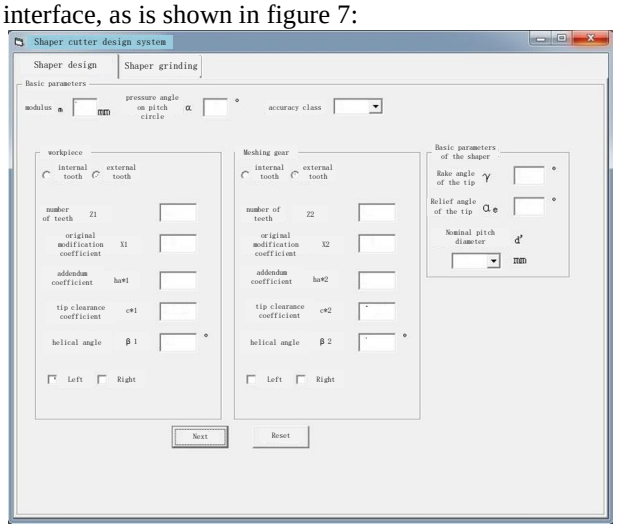


Figure 7 The Interface Of The Shaper Cutter Design System

Design and calculation module. The design calculation module is the core module of the whole shaper cutter design system, bearing the most important calculation task in the system. In the design process, this module gain the design parameters from the human-computer interaction interface module, and finish the main structure data calculation task according to the

design theory and algorithm of the shaper cutter. The calculation result of this module directly determines the correctness of the shaper cutter.

Results display module. The results display module is designed to show the design results visually after the end of the design, so that the user can directly observe the design results. The module is actually integrated in the human-computer interaction interface, will be demonstrated in the follow-up design examples.

4.2 The realization of the main functional module[6]

This shaper cutter design system is the use of Matlab and VB mixed programming to achieve its function. The system uses Matlab as the core of calculation, with VB to achieve human-computer interaction and the results display function. Actually, in this system the interface developed by VB as the main program, the Matlab algorithm as a subroutine, the subroutine is called through VB to achieve a combination of the two mixed programming.

In general, VB calls the Matlab files in the forms of intermediate file transfer, dynamic data exchange (DDE), ActiveX automation services, DLL dynamic link library and COM components. But the first three methods are too tedious and can not be separated from the Matlab environment to run. DLL dynamic link library requires third-party compiler software to compile the Matlab file into a DLL file, and now the software has been stopped. COM components as a common object interface can be directly called by VB, and the Matlab version used in this design system integrated with deploytool, the use of the tool can compile Matlab files into COM components for VB calls. And the most important thing is the use of COM component technology helps the mixed program does not need to run in the Matlab environment, and no need to switch the environment which is greatly improving the system speed. So this chapter mainly uses COM component technology to realize the mixed programming system development and the main process is as follows:

Write the Matlab program. The COM component creation requires that each M file is written in the form of a function to facilitate the transfer of parameters between VB and Matlab.

According to the design principles described above, the M function file of the parameter calculation sub-module in the design calculation module is written first, named Parameter.m, and the calculation flow of this module is shown in figure 8:

The creation and calls of the COM components. Use the deploytool tool in Matlab to create the COM components containing the above M function file. Firstly, create the COM component project and add a new class in the project, then add the M function file that runs through to the class, finally compile and generate the dynamic link DLL file. After the three steps mentioned above, the M function file are written into the COM components.

When VB calls COM components, there should be a reference, to add the COM component to the VB project. When writing the main program, the use of CALL statement to call the M function packaged in the COM component. The call form is "Call class name. Function name (parameter)" which can achieve the function of the shaper cutter design system.

4.3 The design example

In order to verify the degree of functional realization of the design system, the design parameters shown in Table 3 are taken as an example, and the design system is used to design the shaper cutter.

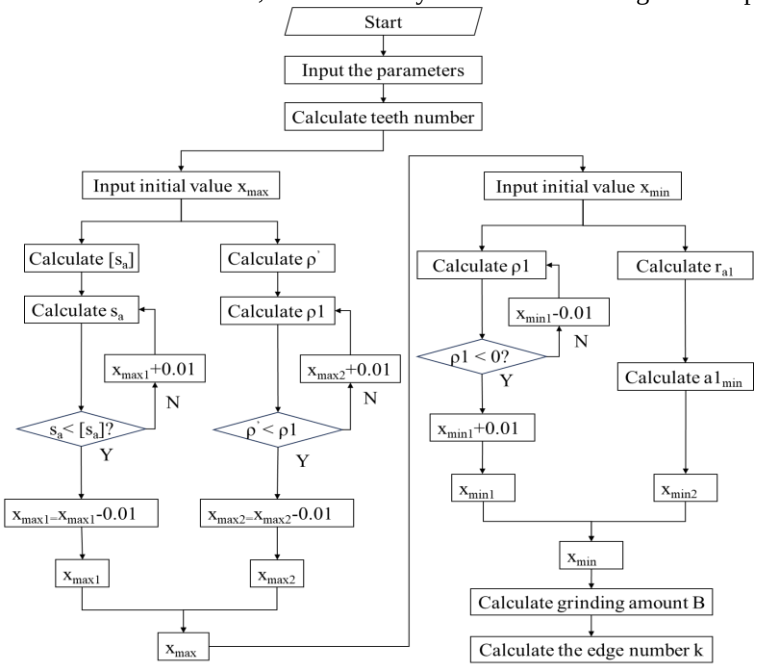


Figure 8 Calculation Process Of All The Parameters

Table 3 The Parameters Of The Workpiece

Basic parameters	Workpiece	Meshing gear
Modulus[mm]	2	2
Number of teeth	30	30
Pressure angle on pitch circle[°]	20	20
addendum coefficient	1	1
tip clearance coefficient	0.25	0.25
The modification coefficient	0	0
Helical angle[°]	15(left)	15(right)

Launch the system, enter the workpiece parameters in the initial operation interface, select 5 ° as the rake angle on the tip edge, and 6 °as the relief angle on the top edge. The nominal pitch circle diameter is 75mm. Start the calculation and the results shown in figure 9:

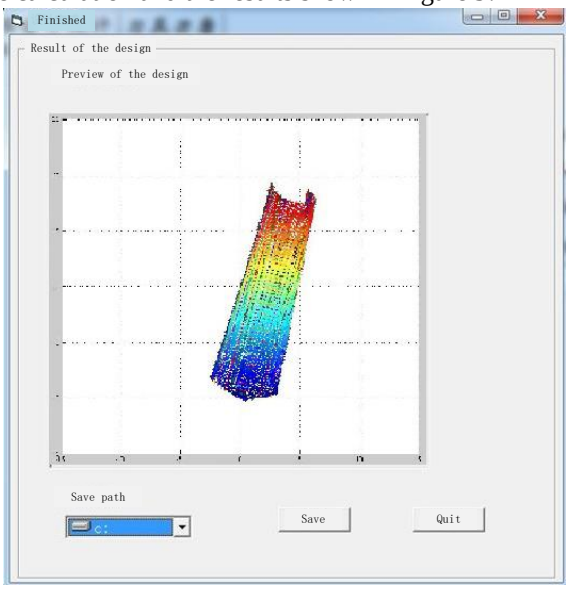


Figure 9 The design result

5 Summary

Based on the basic design theory of the shaper cutter, this paper developed a CAD system of shaper cutter design by using VB and Matlab mixed programming. This design system can finish the shaper cutter design based on the parameters of the workpiece, and the feasibility of the system is verified through a design example.

Acknowledgement: This work was supported by the project of the special fund on the National Natural Science Foundation of China (52165060).

References

[1] WANG Mei, WANG Jie, JIA Zhixin. Development of CAD expert system for involute shaper Cutter [J]. Journal of Sichuan University, 2002,34(2):85-88.

[2] XI zujiang. Research and Development of Object - oriented Gear Tool CAD System [D]. Yichang:Sanxia university, 2004.

[3] ZENG Mengxiong, ZHU Dalin, DU Xuan. CAD system of shaper cutter based on VB Environment [J]. Tool technology, 2002,36(1):33-36.

[4] ZHANG Shengwen, WU Chunqiao, ZHU Yulong. Research on computer aided tool and process parameter optimization system based on VB and Matlab [J]. Manufacturing automation, 2012, 34(1):64-68.

[5] YUAN Zhejun, Gear tool design [M]. Beijing: National Defense Industry Press, 2014.

[6] HUANG Xiquan, Interface programming between VB and Matlab based on COM component [J]. Programming language, 2004 (8):19-21.

Viser Technology Pte. Ltd.
Mechanical Engineering Science
DOI: 10.33142/mes.v7i2.18371

Research Article Open Access

A Simple and Reliable Eccentric Locking Mechanism

Mengjiao NIU, Yong ZHAO, Yongliang YUAN*

He'nan Polytechnic University, Jiaozuo, He'nan, 454000, China

***Corresponding Author:** *Yongliang YUAN, E-mail: yuan-yong-liang@163.com*

Abstract

In view of the time-consuming and unreliable deficiencies of the cross-axis work piece in the clamping process, combined with the working characteristics of the eccentric mechanism, a simple and fast eccentric locking mechanism is designed. The push rod is quickly driven by the combined action of the handle and the drum, so that the cross shaft work piece can be quickly locked in the axial direction. The eccentric locking mechanism not only has simple operation and convenient maintenance, but also has the characteristics of low manufacturing cost and high life, and has certain reference value for future special fixture design.

Keywords: *cross shaft, eccentric mechanism, locking mechanism, special fixture*

1 Introduction

The cross shaft is a fundamental component extensively utilized in various mechanical transmission systems [1]. Its primary function is to endure bending moments and torques during power transmission, thereby ensuring the efficiency and stability of the system. Due to its critical role, the machining and heat treatment processes of the cross shaft are paramount to its overall performance and durability [2].

In conventional manufacturing processes, securely and efficiently clamping the cross shaft workpiece presents significant challenges [3]. Traditional clamping methods often involve the use of large nuts for axial positioning, a technique that can be both time-consuming and potentially unreliable during high-volume production runs [4]. This approach can lead to inefficiencies in the manufacturing process, and in some cases, even accidents if the workpiece is not adequately secured. Traditional methods may fail to provide sufficient locking force, resulting in misalignment or movement during machining, which adversely affects the precision and quality of the finished component [5-7]. Hence, it is imperative to develop an improved clamping mechanism that ensures quick, reliable, and robust fixation of the cross shaft workpiece [8].

The reliance on large nuts and similar conventional locking methods can be cumbersome and inefficient, especially in high-volume manufacturing settings. The manual effort required to secure and release these nuts not only slows down the production process but also introduces variability in the clamping force applied [9]. This variability can lead to inconsistent machining results, where the precision and quality of the cross shaft are compromised. Moreover, in scenarios where the locking force is insufficient, there is a heightened risk of the workpiece shifting during the machining process, which can result in defective products and potential safety hazards [10].

To address these challenges, it is essential to explore alternative clamping mechanisms that can offer more consistent, reliable, and efficient performance [11]. The eccentric mechanism, known for its simplicity and effectiveness in various mechanical applications, presents a promising solution [12]. An eccentric mechanism operates on the principle of using a driving and a rotational locking mechanism that are not coaxial [13-17]. This unique arrangement allows for significant motion amplification from minimal input movement, making it particularly suitable for applications that require rapid and secure locking.

Integrating the eccentric mechanism into the design of a specialized fixture for the cross shaft can significantly enhance the clamping process [18]. The proposed design leverages the inherent advantages of the eccentric mechanism to develop a straightforward, fast, and reliable locking system. This system comprises a handle, roller, and support structure that collaborate to drive a push rod, thereby securing the cross shaft axially. The innovative approach aims to reduce setup times, improve locking force, and ensure consistent precision during machining operations, particularly when drilling

Received on August 4, 2025; Accepted on September 18, 2025

the Φ8 hole in the cross shaft.

Through detailed analysis and rigorous testing, this paper demonstrates that the eccentric locking mechanism simplifies operation and maintenance. Additionally, it reduces manufacturing costs and extends the fixture's lifespan. The findings underscore the potential of this mechanism to serve as a valuable reference for future fixture designs in similar applications. This advancement paves the way for enhanced efficiency and reliability in mass production environments, marking a significant improvement over traditional clamping methods.

By implementing this novel clamping mechanism, manufacturers can achieve more reliable and efficient production processes, ultimately contributing to higher quality components and reduced operational risks. This paper highlights the critical aspects of the design and testing phases, providing comprehensive insights into the practical benefits and potential applications of the eccentric mechanism in mechanical transmission systems. The success of this innovative approach underscores the importance of continuous development and optimization in manufacturing technologies to meet the evolving demands of modern engineering and production.

2 Cross Shaft Workpiece Analysis

Considering the process requirements of the cross shaft, a special fixture needs to be designed for the machining of the Φ8 hole [19]. The part drawing of the cross shaft is shown in Figure 1.

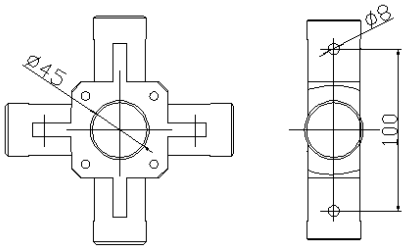


Figure 1 Part Drawing of the Cross Shaft Workpiece

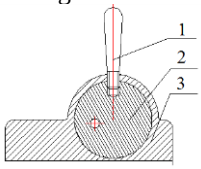
According to traditional clamping methods, a Φ45 hole is typically used for axial positioning, and the workpiece is rotationally constrained using a limit groove [20-22]. After positioning the workpiece, the simplest way to lock it is by using a large nut that fits with the Φ45 positioning shaft; rotating the nut secures the workpiece. However, cross shaft parts are widely used and usually need to be produced in batches. If the aforementioned method is used to lock cross shaft workpieces, the loading and unloading process becomes cumbersome, and there is a risk that the locking force may be insufficient, potentially leading to accidents.

In traditional methods, manually rotating the large nut during loading and unloading is time-consuming, affecting production efficiency [23]. This method is particularly inconvenient in batch production [24]. Additionally, the manual operation may result in an unstable locking force, which could loosen during production, causing the workpiece to shift or even fall, leading to safety hazards and quality issues.

To address these problems, this paper designs a simple and fast locking mechanism based on the principle of the eccentric mechanism. The eccentric mechanism generates locking force through changes in the eccentric distance, enabling quick and reliable locking. Compared to the traditional method of locking with a large nut, the eccentric mechanism offers advantages such as ease of operation, stable locking force, and strong adaptability. The use of an eccentric mechanism significantly reduces the loading and unloading time of the workpiece, improves production efficiency, and ensures the reliability and safety of the locking process.

3 Principle of the Eccentric Mechanism

The eccentric mechanism primarily operates by using a driving mechanism and a rotational locking mechanism that are not coaxial [25]. The structural principle is shown in Figure 2.



1-Handle; 2-Roller; 3-Support

Figure 2 Principle of the Eccentric Mechanism

From Figure 2, it can be seen that the eccentric mechanism is designed to be simple yet effective, consisting of core components such as the handle, roller, and support. The handle is the driving component in the eccentric mechanism, used to apply external force [26]. When force is applied to the handle and it begins to rotate, it will drive the movement of the eccentric axis. The roller serves as the rotational locking component. There is typically an eccentric axis hole on the roller, used to connect other structural components of the mechanism and achieve eccentric motion. The base acts as the fixed component of the eccentric mechanism, providing support for the roller and other moving parts to ensure stability and reliability of the mechanism. During machining and assembly, it is common to thread the handle and roller together to ensure their coordinated operation. Once external force is applied to the handle, it begins to rotate, activating the eccentric shaft hole on the roller. This action initiates the movement of the mechanism connected to the roller, thereby enabling eccentric motion of the entire mechanism [27].

One of the advantages of the eccentric mechanism

is its simple and easy-to-manufacture structure, along with rapid operation. Even when the handle is rotated at a relatively small angle, it can generate a large range of motion within the mechanism connected to the roller. This characteristic allows the eccentric mechanism to quickly and effectively tighten and secure the workpiece, thereby improving work efficiency.

Furthermore, due to its relatively small space occupation, the eccentric mechanism offers greater flexibility in design and layout. This versatility enables the eccentric mechanism to be widely used in various engineering and mechanical applications, making it an ideal choice for solving complex workpiece fixation and positioning problems.

4 Application of the Eccentric Mechanism in Fixture Design

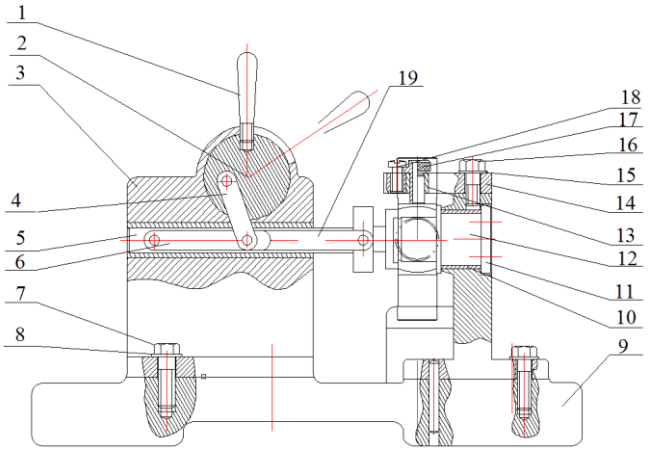
Through the analysis of the cross shaft workpiece, when machining the Φ8 hole on the cross shaft, in order to achieve rapid and efficient positioning of the cross shaft, the following strategy is adopted based on the characteristics of the eccentric mechanism for positioning and tightening the cross shaft workpiece [28-30]:

(1) Use the Φ45 hole and the frame to provide axial positioning and axial movement restriction for the cross shaft.

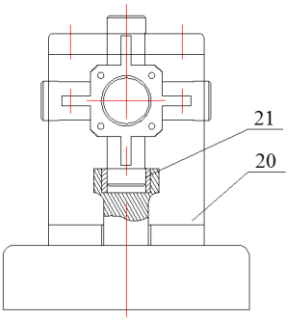
(2) Utilize a limit groove to restrict the rotation of the cross shaft around the axis of the Φ45 hole. To ensure proper positioning, the limit groove needs to be of a through-hole type.

(3) Incorporate the features of the eccentric mechanism by adding a connecting rod on the roller of the eccentric mechanism. With the action of the handle and roller, the push rod can move along the axis of the Φ45, thereby locking the cross shaft.

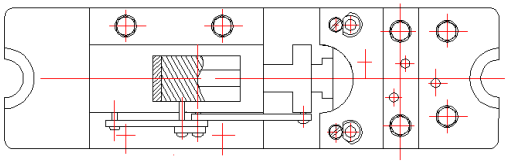
The application of the eccentric mechanism in the specialized fixture for the cross shaft is shown in Figure 3.



(a) Front View



(b) Left View



(c) Top View

1- Handle; 2- Roller; 3- Support; 4- Connecting Rod; 5- Shaft; 6- Push Rod; 7- Bolt; 8- Washer; 9- Base; 10- Liner; 11- Cover Plate; 12- Positioning Pin; 13- Drill Sleeve Liner; 14- Drill Template; 15- Thick Washer; 16- Bolt; 19- Connecting Rod; 20- Support Plate; 21- Groove Liner

Figure 3 Application of the Eccentric Mechanism in the Fixture

To efficiently achieve the machining of the Φ8 hole while ensuring accuracy, it is necessary to design a quick-change drill sleeve structure. Considering that this part is typically produced in large quantities, the fixture design needs to incorporate a quick-change drill sleeve on the drill template. The structure is shown in Figure 4.

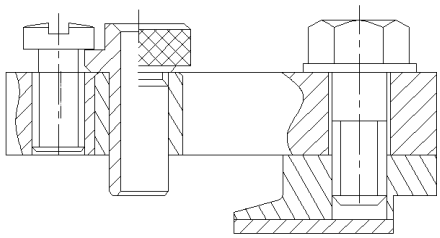


Figure 4 Schematic Diagram of Quick-Change Drill Sleeve

Through the manufacturing and testing of this specialized fixture, it has been shown to effectively improve machining efficiency while enhancing the reliability of the clamping process. This fixture design serves as a valuable reference for future similar designs.

4 Conclusion

Initially, the study identified the inefficiencies of the conventional machining process for cross shafts, highlighting its time-consuming nature and low reliability. Drawing upon the principles of the eccentric mechanism and recognizing the mass production requirements of cross shafts, the research team developed a straightforward, swift, and dependable eccentric mechanism. Following this, extensive analysis and practical experimentation were conducted to validate the efficacy of this clamping technique. The results not only confirmed the simplicity and ease of operation of the proposed method but also underscored its robustness and reliability. Additionally, the study emphasized the potential applicability of the eccentric mechanism in similar fixture designs, offering a framework for future optimization endeavors in this domain.

References

[1] Sun Heng, Chen Zuomo. Mechanical Principles [M]. Beijing: Higher Education Press, 2000.

[2] Wang Shaonan, Liu Zhan, Gao Fanbo. Improvement of Cross Shaft Support Process Plan [J]. Practical Technology of Automobiles, 2020.

[3] Xie Fuguai, Mei Bin, Liu Xinjun, et al. Discussion on New Mode and Equipment for Processing Large and Complex Components [J]. Journal of Mechanical Engineering, 2020, 56(19): 70-78.

[4] Goh Y M, Micheler S, Sanchez-Salas A, et al. A Variability Taxonomy to Support Automation Decision-making for Manufacturing Processes [J]. Production Planning & Control, 2020, 31(5): 383-399.

[5] Dharmaraj K. Automated Freeform Assembly of Threaded Fasteners[D]. Loughborough University, 2015.

[6] Jia Z, Bhatia A, Aronson R M, et al. A Survey of Automated Threaded Fastening [J]. IEEE Transactions on Automation Science and Engineering, 2018, 16(1): 298-310.

[7] Kharlamov Y A, Sokolov V I, Krol O S, et al. Assurance of Cutting Tools Reliability [J]. 2020.

[8] Qin Guohua, Zhang Weihong. Modern Design Methods of Machine Tool Fixtures [M]. Beijing: Aviation Industry Press, 2006.

[9] Liu Jing, Zhu Hua, Chang Junran. Comprehensive Practice of Mechanical Design [M]. Chongqing University Electronic Audiovisual Publishing Co., Ltd., 2020.

[10] An Jiansheng, Li Peng. Discussion on Mechanical Design and Safety Design [J]. Engineering Research and Practice, 2023, 4(10): 147-149.

[11] Liu Dawei, Li Bingbing, Fu Zhanglei, et al. Configuration Principle and Dynamic Anti-Slip Mechanism of Non-Circular Gear Differential [J]. *Journal of Mechanical Engineering*, 2023, 59(5): 67-76.

[12] Gameros A, Lowth S, Axinte D, et al. State-of-the-Art in Fixture Systems for the Manufacture and Assembly of Rigid Components: A Review[J]. International Journal of Machine Tools and Manufacture, 2017, 123: 1-21.

[13] Teixeira Carvalho D J, Moroni L, Giselbrecht S. Clamping Strategies for Organ-on-a-Chip Devices[J]. Nature Reviews Materials, 2023, 8(3): 147-164.

[14] Zhang Hongwei, Wu Zhiqiang, Gong Yubin. Principles and Maintenance of Automotive Automatic Transmissions [M]. Huazhong University of Science and Technology Press Co., Ltd., 2019.

[15] Yang Y, Wang J, Zhou S, et al. Design of a Novel Coaxial Eccentric Indexing Cam Mechanism[J]. Mechanism and Machine Theory, 2019, 132: 1-12.

[16] Maritano M. Design of Electro-Mechanical Height Adjustment System for Multi-link Suspension[D]. Politecnico di Torino, 2023.

[17] Du L, Yuan J, Bao S, et al. Robotic Replacement for Disc Cutters in Tunnel Boring Machines[J]. Automation in Construction, 2022, 140: 104369.

[18] Luo M, Luo H, Axinte D, et al. A Wireless Instrumented Milling Cutter System with Embedded PVDF Sensors[J]. Mechanical Systems and Signal Processing, 2018, 110: 556-568. [19] Udoinyang H N. Technological manufacturing process of a part “Shaft housing”[J]. 2022.

[19] Cao Yongjie. Research on the Processing Technology and Fixture Design of Baffle Shaft[J]. Machine Tools and Hydraulics, 2017, 45(8): 19-21

[20] Jiao, Yuanyan, Wu, Jing. Elbow Lathe Processing Thread Fixture Design[J]. Machine Tool & Hydraulics, 45(14): 179-180.

[21] Lei Mingwei, Shi Wenpu, Ren Pingchuan. New Lathe Fixture for Clamping Double Eccentric Shaft with Phase Requirements[J]. Machine Tool & Hydraulics, 46(21): 105-106,111.

[22] Jinfeng W, Guanbao G A O, Wuxiong W, et al. Design and

Experiment of Key Components of Side Deep Fertilization Device for Paddy Field[J]. Nongye Jixie Xuebao/Transactions of the Chinese Society of Agricultural Machinery, 2018, 49(6).

[23] Zhang Kaixue, Xu Baojun, Fang Zhongqiu. Research on Turning Processing Technology of Thin-walled Long Shaft[J]. Machine Tools and Hydraulics, 2017, 45(8): 46-48.

[24] Song Pengju, He Yaohua, Wu Yafei. Stability Analysis of Handbrake Pulling Force with Double Nut Locking Process[J]. Journal of Jiangsu University (Natural Science Edition), 2017, 38(3).

[25] Tinwala F, Cronin J, Haemmerle E, et al. Eccentric Strength Training: A Review of the Available Technology[J]. Strength & Conditioning Journal, 2017, 39(1): 32-47.

[26] Si Guangju, Zhong Kangmin. Design and Mechanical Calculation of Increased Force Self-Locking Impact Pneumatic Fixture [J]. Mechanical Design and Research, 2008(1): 96-99.

[27] Zhang Yamin, Li Zhanfeng, Ding Junjian. Design and Application of Special Fixtures for Turning Thin Disc-shaped Workpieces [J]. Machine Tools and Hydraulics, 2013(22): 60-61.

[28] Xiao Jide, Chen Ningping. Machine Tool Fixture Design [M]. Beijing: Machinery Industry Press, 1998: 46-47.

[29] Thomas R G, Deepak Lawrence K, Manu R. Step AP 242 Managed Model-Based 3D Engineering: An Application towards the Automation of Fixture Planning[J]. International Journal of Automation and Computing, 2021, 18(5): 731-746.

[30] Boyle I, Rong Y, Brown D C. A Review and Analysis of Current Computer-Aided Fixture Design Approaches[J]. Robotics and Computer-Integrated Manufacturing, 2011, 27(1): 1-12.