ISSN: 2472-3703 DOI: 10.6911/WSRJ.202508_11(8).0002

A Multi-View Feature-Based Retrieval Method for On-Machine Measurement Model Reuse

Zhongbao Liu [1, a, *], Ying Xiang [2, b]

[1]School of Electromechanical Engineering, Shaanxi University of Science and Technology, Xi'an 710016, China

[2] School of Electromechanical Engineering, Shaanxi University of Science and Technology, Xi'an 710016, China

[a]15735715240@163.com, [b]xiangying@sust.edu.cn

[*] Corresponding author

Abstract

In the design of on-machine measurement (OMM) schemes for complex parts, the involvement of diverse geometric features, physical environments, and hardware/software systems makes the process intricate and inefficient. To address the high dependency on manual experience and the low efficiency in current OMM scheme design, this paper proposes a multi-view 3D model retrieval method aimed at facilitating the reuse of existing measurement solutions. Specifically, multiple view images of the same model are processed using a multi-view convolutional neural network (ResNet-18) to extract view-specific features. A self-attention mechanism is introduced to enhance the network's focus on informative views and critical features. These features are then fused into a global feature descriptor through a long short-term memory (LSTM) network. Finally, similarity between the query model and models in the feature database is computed, and retrieval is performed by ranking similarity scores in descending order. Based on the retrieved models, corresponding measurement schemes can be adjusted and reused efficiently. Experimental results demonstrate that the proposed method outperforms the MVRNN approach in terms of Precision and Recall, and improves the mean Average Precision (mAP) by 5%. The method effectively enhances the efficiency of OMM scheme design and shows strong potential for practical engineering applications.

Keywords

On-machine measurement; Model retrieval; Multi-view; Neural networks; Reuse technology.

1. INTRODUCTION

In recent years, with the rapid development of advanced manufacturing technologies such as high-precision CNC machining, additive manufacturing, and complex structural design, the geometry, material composition, and processing of parts have become increasingly complex, placing higher demands on quality control in the manufacturing process[1]. On-machine measurement (OMM), which enables integrated machining and inspection, has emerged as a key technology for improving inspection efficiency and advancing intelligent manufacturing. However, the current design of OMM schemes still heavily relies on manual experience, resulting in low efficiency and limited reusability, which restricts the broader application of OMM[2]. To enhance design efficiency and promote standardization, it is necessary to establish a measurement scheme library based on geometric features and enable rapid reuse through model retrieval. Currently, 3D model retrieval methods mainly include approaches based on geometric descriptors, topological structures, or semantic labels. However, these methods generally rely on manually designed low-level descriptive features, which are highly subjective and heavily dependent on domain expertise. Moreover, they suffer from limited representational capability and low computational efficiency when dealing with complex structural models. In response to the limitations of existing 3D model retrieval methods—particularly their weak representation capabilities for complex structures—this paper proposes a multi-view feature-based retrieval method for on-machine measurement models. The method employs ResNet-18[3] to extract features from multi-view images, incorporates a self-attention mechanism to emphasize key features, and uses a long short-term memory (LSTM)[4] network to fuse spatial relationships between views into a global feature descriptor. For similarity measurement, a combined metric based on Euclidean distance and cosine similarity is introduced to improve retrieval robustness. The proposed method accurately identifies models with similar geometric features, enabling efficient reuse of measurement schemes. Experimental results show that the method outperforms traditional approaches in terms of Precision, Recall, and mean Average Precision (mAP), demonstrating strong practical value and applicability in engineering contexts.

2. THE METHOD PROPOSED IN THIS PAPER

2.1. Model feature extraction

In the multi-view feature extraction process for 3D model retrieval, the core objective of feature extraction is to accurately represent the measurement-related geometric features in each view. Considering that all 2D images from different perspectives originate from the same 3D model, this study designs a convolutional neural network (CNN) architecture with shared weight parameters to extract the desired geometric features. The core of the feature extraction module is based on the ResNet-18 network, whose residual learning mechanism and skip connections effectively mitigate issues such as gradient vanishing and network degradation during the training of deep networks. This allows for more complete preservation of geometric detail in the images. Moreover, ResNet-18 strikes a good balance between network depth and parameter complexity, providing sufficient feature representation capability while maintaining training efficiency. Assuming that the feature representation extracted from each view through the mapping function of the convolutional neural network is denoted as v_{ij}.The formula is given as follows:

$$v_{ij} = f\left(V_{ij};\theta\right). \tag{1}$$

Here, $f\left(\cdot;\theta\right)$ denotes the feature extraction function implemented by ResNet-18, θ represents the parameters of the neural network, and $v_{ij} \in R^d$ refers to the feature vector extracted from the *j*-th view of the *i*-th model. d indicates the dimensionality of the feature vector, which is 512.

For each model *i*, there are multiple view-specific features as $v_i = \left\{v_{i1}, v_{i2}, \cdots, v_{im}\right\}$, and m denotes the total number of views associated with the model. The flow of the ResNet-18 network is shown in Figure 1 below.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202508_11(8).0002

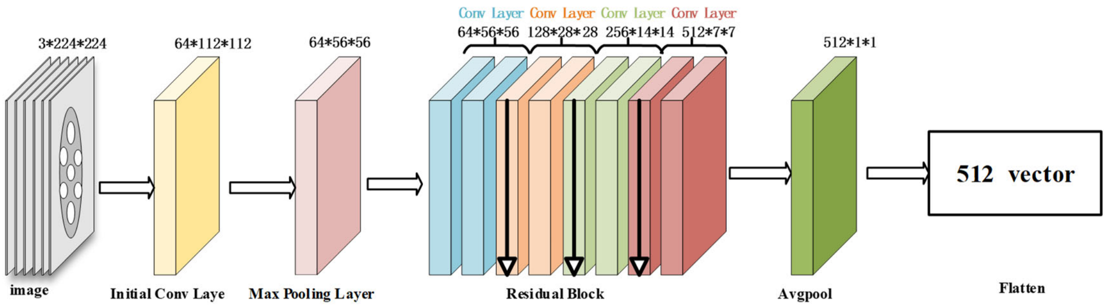

Figure 1. Flowchart of the ResNet-18 network architecture

2.2. Model feature fusion

The features extracted from a single view are insufficient to comprehensively represent the global characteristics of a 3D model. Therefore, it is necessary to fuse the features from multiple views into a unified global feature descriptor. Since different views contribute unequally to the overall representation, the importance of each view should be weighted accordingly during the fusion process. To address this, a self-attention [5]mechanism is introduced to dynamically assign weights, enabling the model to learn which views are more relevant to the current task. By adjusting each view's contribution based on its attention weight, the model can focus more effectively on the most informative perspectives during the feature fusion process.

Each view-specific feature v_{ij} of the model is mapped to $Query$, Key , and $Value$ vectors as follows:

$$Q_{ij} = W_q v_{ij}, K_{ij} = W_k v_{ij}, V_{ij} = W_v v_{ij} \tag{2}$$

Where W_q, W_k, W_v denotes the respective weight matrices, aiming to project the input features into the query, key, and value vector spaces.
The attention score for each view-specific feature is computed as:

$$score\left(v_{ij}, v_{ik}\right) = \frac{\left(Q_{ij}\right)^T K_{ik}}{\sqrt{d_k}} \tag{3}$$

Where d_k is the dimensionality of the key vector, set to 512, and the dot product represents the relevance between the query and key vectors.

The attention weights w_{ij} for each view are then used to compute a weighted sum of the value vectors:

$$w_{ij} = \frac{\exp\left(score\left(v_{ij}, v_{ik}\right)\right)}{\sum_{k=1}^{m} \exp\left(score\left(v_{ij}, v_{ik}\right)\right)} \tag{4}$$

Here, $v_{att,ij}$ represents the aggregated view-specific features after applying the attention-based weighting.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202508_11(8).0002

To better exploit the spatial structural relationships among different views, a recurrent neural network (RNN) is introduced. Assuming the attention-weighted features of model i are denoted as $v_{att,i} = \left\{ v_{att,i1}, v_{att,i2}, \cdots v_{att,im} \right\}$, these features are sequentially fed into the RNN for progressive fusion. The feature representation during the fusion process is given by:

$$h_i^{(t)} = RNN\left(h_i^{(t-1)}, v_{att,it} \right) \tag{5}$$

Where an LSTM network is selected as the RNN unit, $h_i^{(t)}$ represents the fused feature at the t-th step, and $h_i^{(0)}$ is the hidden state at the previous step, typically initialized to zero.
The final global fused feature representation of all views for model i under the RNN framework is expressed as:

$$v_{global,i} = h_i^{(m)} \tag{6}$$

2.3. Similarity retrieval

The same preprocessing and feature extraction procedures are applied to the input query 3D model to generate the query feature v_{query}. This feature is then compared with those in the feature database through similarity computation and matching. Common similarity measures include Euclidean distance and cosine similarity. Euclidean distance:

$$d = d(v_{query}, v_{global,i}) = \left\| v_{query} - v_{global,i} \right\|_2 \tag{7}$$

Cosine similarity:

$$s = s(v_{query}, v_{global,i}) = \frac{v_{query} \cdot v_{global,i}}{\left\| v_{query} \right\| \cdot \left\| v_{global,i} \right\|} \tag{8}$$

Each of these two methods has its respective advantages and disadvantages. To comprehensively assess both absolute differences and directional similarity, a new similarity metric[6] S is proposed by combining Euclidean distance and cosine similarity, defined as follows:

$$S = \alpha \cdot s(v_{query}, v_{global,i}) + (1-\alpha) \cdot \frac{1}{1 + d(v_{query}, v_{global,i})} \tag{9}$$

Here, α is used to balance the weights of the two metrics, $0 < \alpha < 1$.

$$\text{Results} = \{Model_1, Model_2 \cdots\cdots, Model_k\} \tag{10}$$

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202508_11(8).0002

Finally, all results are ranked according to their similarity scores, and the top k matching models are returned. Thus, the corresponding measurement scheme associated with the retrieved model can be invoked.

3. EXPERIMENTAL VALIDATION

To further validate the effectiveness and generalizability of the proposed multi-view feature fusion-based retrieval method for on-machine measurement models, a comprehensive experimental study was conducted on a self-constructed 3D model dataset. The experiments include data preprocessing and network training configuration, evaluation of retrieval performance, comparative analysis with mainstream methods, and ablation studies on key components. These experiments aim to assess the proposed method’s retrieval capability in complex part scenarios from multiple perspectives. The results not only confirm the advantages of the proposed method in terms of retrieval accuracy, robustness, and computational efficiency, but also provide a solid theoretical foundation and practical support for future engineering applications.

3.1. Dataset and Experimental Setup

To further validate the effectiveness and feasibility of the proposed method in the reuse of on-machine measurement schemes, an experimental case study was conducted. Due to the lack of publicly available on-machine measurement model datasets that include parametric design features and machining parameters, the models used in this study were collected from an enterprise design model library and project datasets developed by our research group. A total of 2,200 commonly used CAD models were included. Each model was rendered from multiple virtual viewpoints to generate corresponding 2D images, which were used as multi-view inputs for feature extraction and fusion. These features were then aggregated to form a complete feature descriptor and used to construct a model feature database for subsequent retrieval. Furthermore,a measurement scheme was pre-designed for each part model in the database, providing a foundation for downstream scheme reuse. The implementation environment for this study was based on Python, running on an Intel Core i7-12700H @ 2.30 GHz processor with 16.0 GB of RAM. The retrieval algorithm was developed using the open-source deep learning framework PyTorch.

3.2. Training strategy

The retrieval method proposed in this study consists of three core modules, which are responsible for view-specific feature extraction, dynamic view-weight allocation, and global feature fusion. To enhance the model’s discriminative capability under complex geometric structures, a triplet loss function [7] is introduced during training as a unified optimization objective across the entire pipeline, guiding feature extraction, feature fusion, and similarity measurement in an end-to-end manner. The objective of the triplet loss is to minimize the feature distance between the anchor (query) sample and a positive (same-class) sample, while maximizing the distance between the anchor and a negative (different-class) sample. This mechanism promotes intra-class feature compactness and inter-class feature separability, thereby improving the robustness and retrieval accuracy of the model when dealing with complex geometries. The training parameters are detailed in Table 1.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202508_11(8).0002

Table 1. Training Strategy and Parameters

Stage	Network	Optimizer	Loss function	Learning Rate	Learning Rate Policy	Number of Training Epochs	Batch
Feature Extraction	ResNet-18	Adam	Triplet Loss	0.001	Reduce by a factor of 0.5 every 5 epochs	20	32
Feature Fusion	Self-Attention+LSTM	Adam	Triplet Loss	0.001		20	32
Similarity Measurement	Euclidean Distance + Cosine Distance	Adam	Triplet Loss	0.001		20	32

3.3. Comparison Methods and Evaluation Metrics

To demonstrate the advantages of the proposed method for on-machine measurement model retrieval, several comparative experiments were conducted using different feature extraction and fusion approaches. Specifically, experiments were performed on the self-constructed dataset using LFD (a traditional method), MVCNN (multi-view convolutional neural network), MVRNN (multi-view recurrent neural network), and the proposed method (a network incorporating an attention mechanism).

For the evaluation of retrieval performance, this study primarily employs Recall, Precision, and mean Average Precision (mAP) as key metrics. Precision reflects the quality of the retrieved results, indicating whether the retrieved models are relevant, while Recall measures the completeness of retrieval, assessing whether relevant models have been missed.

3.4. Analysis of Experimental Results

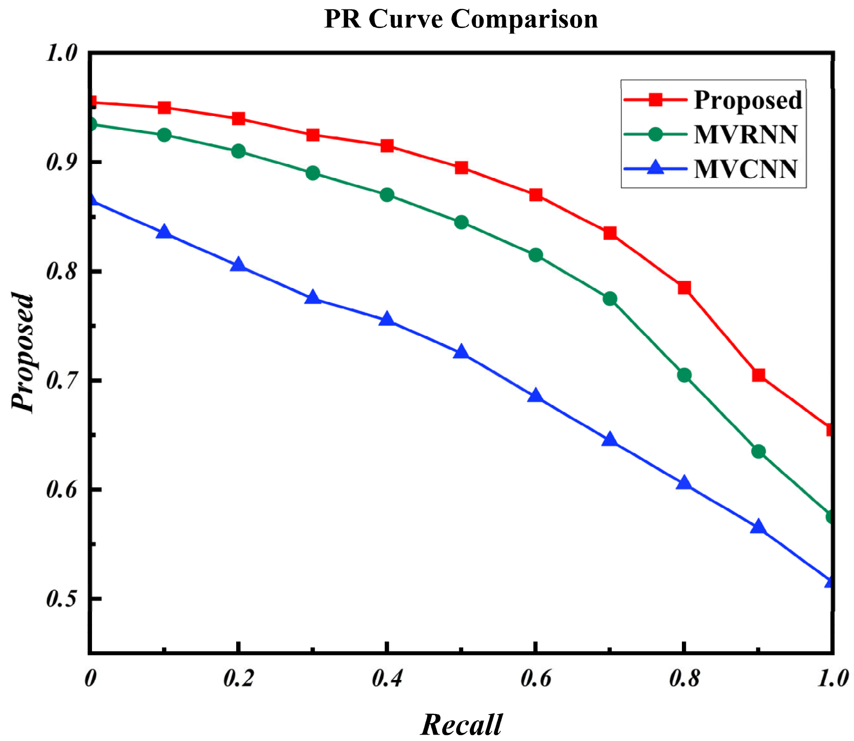

Figure 2. PR Curves of the Three Methods on the Dataset Used in This Paper

Figure 2 presents the Precision-Recall (PR) curves of the three methods on the dataset used in this study. Based on the PR curve data get the mean Average Precision (mAP) values for MVCNN, MVRNN, and the proposed method are calculated as 70.68%, 80.73%, and 85.73%, respectively. The experimental results clearly demonstrate that the proposed method

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202508_11(8).0002

significantly outperforms both MVCNN and MVRNN in terms of Precision, Recall, and mAP. Notably, compared with MVRNN, the proposed method achieves a 5% improvement in mAP. It is also worth noting that the proposed method maintains high precision even in the high-recall region, further validating its superior performance and robustness. In addition, benefiting from the compression of the feature descriptor dimensionality, the proposed method improves retrieval efficiency while enhancing retrieval accuracy.

4. CONCLUSION AND DISCUSSION

This paper proposes a multi-view feature-based retrieval method for on-machine measurement (OMM) model reuse, aiming to improve the efficiency and accuracy of OMM scheme design for mechanical parts. By integrating deep learning techniques with multi-view image representations, the proposed method effectively captures the geometric details of parts from various perspectives, enabling accurate model matching and facilitating the reuse of corresponding measurement schemes. Experimental results show that the multi-view feature-based approach outperforms traditional methods in terms of retrieval accuracy and precision, and demonstrates practical applicability on large-scale datasets.

Nevertheless, challenges remain, such as enhancing sensitivity to local geometric details and improving the efficiency of multi-view feature fusion. Future research may explore the following directions: introducing local geometric descriptors to improve sensitivity to fine structures; integrating CAD semantic information and machining attributes to enrich model representations; and developing an automated measurement scheme generation system tailored to real-world machining equipment, thereby establishing a complete closed-loop workflow from model retrieval to measurement scheme design.

REFERENCES

[1] Q.H. Zhang, From information-assisted production to intelligent manufacturing, Mod. Manuf. 2025 (01) (2025) 54–56.

[2] S. Hong, Application of on-machine measurement technology in CNC machining, Shanghai Light Ind. (06) (2023).

[3] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, 2016, pp. 770–778.

[4] A. Zadeh, P.P. Liang, N. Mazumder, S. Poria, E. Cambria, L.P. Morency, Memory fusion network for multi-view sequential learning, in: Proc. AAAI Conf. Artif. Intell. 32 (1) (2018).

[5] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, 2016, pp. 770–778.

[6] J. Zhu, H. Zeng, S. Liao, Z. Lei, C. Cai, L. Zheng, Deep hybrid similarity learning for person re-identification, arXiv:1702.04858v2, 2017.

[7] F. Schroff, D. Kalenichenko, J. Philbin, FaceNet: A unified embedding for face recognition and clustering, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2015, pp. 815–823.