



Study on Feature Point Matching for Diverse Workpieces Based on an Improved Speeded Up Robust Features(SURF) Algorithm

Wenjing Liu^{1,3} , Yue Ma^{1,2} , Chongwei Tan³ , Ben Niu⁴ , Petrishin Grigory⁵ , Shaofeng Wang^{1,2,*} , Yanjie Xu³ 

¹ Inner Mongolia University of Science & Technology, School of Mechanical Engineering, No.7 Arden Ave, Baotou, Inner Mongolia Province 014010, China; nkd_lwj@imust.edu.cn(WL)

² Inner Mongolia Key Laboratory of Intelligent Diagnosis and Control of Mechatronic Systems, No.7 Arden Ave, Baotou, Inner Mongolia Province 014010, China; prongs@foxmail.com(YM)

³ State Key Laboratory of Special Vehicle Design and Manufacturing Integration Technology, Democracy Road, Qingshan District, Baotou, Inner Mongolia Province 014060, China; a1234565552025@outlook.com(CT); m7894561231998@outlook.com(YX)

⁴ Shanxi Aerospace Tsinghua Equipment Co, LTD, No. 266, Luyangmen South Road, Yan 'an South Road Sub-district, Luzhou District, Changzhi City, Shanxi Province 046012, China; 2845550624@qq.com

⁵ Sukhoi State Technical University of Gomel: Homyel, BY, Republic of Belarus, petrishingrig@gmail.com

*Correspondence: Shaofengwang2025@qq.com

Article history

Received 06.02.2026

Accepted 01.06.2026

Available online 30.06.2026

Keywords

Machine Vision,
Identification of diverse
workpieces,
Salient region,
Feature matching,
SURF.

Abstract

Owing to variations in the size, shape, moving speed, and direction of different workpieces, robotic arm vision systems exhibit insufficient capability in identifying feature points across diverse workpieces, resulting in limited flexibility of visual recognition. To address this issue, this study focuses on feature matching methods for multi-workpiece identification, introducing a feature point matching algorithm into the visual module. Specifically, A multi-workpiece recognition feature matching method based on the improved Speeded Up Robust Features (SURF) algorithm is proposed. The method incorporates the extraction of salient regions to reduce the number of mismatched points and adopts the principal component analysis (PCA) dimensionality reduction principle to reduce the dimension of feature point description vectors, thereby shortening the matching time. Furthermore, comparative experiments were conducted against the traditional SURF algorithm. Experimental results demonstrate that, compared to the traditional algorithm, the improved SURF algorithm enhances recognition accuracy by over 95% and improves matching speed by nearly 50%, contributing to the further optimization of the SURF algorithm's matching performance. Workpiece recognition experiments were conducted using a KUKA robotic arm for validation, which also confirms that the proposed improved SURF algorithm can meet the requirements of workpiece recognition in practical working environments.

DOI: 10.30657/pea.2026.32.31

1. Introduction

In recent years, with the acceleration of industrial production automation, workpiece recognition has become an indispensable component in production lines, strongly promoting the development of industrial intelligence (Wei G, 2023). Workpiece detection technology is widely applied in manufacturing, with its core lying in precise target positioning and classification of diverse workpieces (Chen C et al., 2024).

Modern industry imposes increasingly stringent demands on workpiece precision (Anshi M et al., 2024), as high-precision workpieces are essential for ensuring the performance and stability of mechanical equipment.

Over the past few years, machine vision systems have played an increasingly crucial role in industrial production. A machine vision system captures workpiece images via a high-resolution camera. Subsequently, it extracts feature information such as edges, contours, and colors of the workpieces



by leveraging the deep-learning and image - processing algorithms. This enables rapid and accurate recognition of workpieces.

For instance, an industrial AI vision system can support the simultaneous identification of multiple workpieces and precisely handle randomly arranged workpieces. In the industrial realm, machine vision finds applications in image recognition (Oztemel E et al., 2020; Wenjian C et al., 2026), defect detection (Kan L et al., 2025; Jide Q et al., 2018), and size measurement (Agnieszka S, 2026; Zhisheng D et al., 2016), among other areas.

With the advent of intelligent manufacturing, the integration of vision technology has made the application of robotic arms in industry more intelligent (Marfo R et al., 2025). As production requirements continue to rise, industrial production is steadily moving towards greater intelligence. This trend not only presents higher challenges to workpiece recognition technology but also spurs the continuous advancement of related technologies, as well.

On one hand, the workpiece recognition needs to continuously enhance its accuracy to meet the strict demands of high-precision production. On the other hand, the recognition system must be compatible with a wide variety of workpieces to achieve accurate identification and classification.

In actual production scenarios, due to diverse production environments, workpiece recognition can be affected by various external factors. These factors can undermine recognition accuracy, and different recognition algorithms may also lose their precision under such influences. Consequently, higher requirements are placed on the design and optimization of recognition algorithms.

To address the aforementioned issues, a variety of tracking algorithms have emerged. Leveraging these algorithms, foreign robot companies have allocated substantial resources to the development and research of dynamic tracking systems. For instance, the V-STARS system developed by GSI Company in the United States is an industrial digital photography - based coordinate measurement system (Likun W et al., 2024).

The second-generation model IRB360 of the Swiss ABB FlexPicker series robot, recognized as the world's fastest robot, boasts a processing speed of up to 180 cycles per minute. In scenarios with high processing-speed requirements, it can achieve twice the efficiency. The robot's vision module can accurately identify the target object. In practical applications, it can efficiently sort different workpieces.

Meanwhile, domestic scholars have also conducted extensive research on tracking algorithms. Zhao (Zhiyuan Z, 2022) analyzed the deficiencies of the original tracking algorithm. They proposed improvements to it, utilized the Kalman filter for prediction, and finally analyzed the entire tracking process. Subsequently, they developed an automatic tracking system for the robotic arm based on visual assistance, which enhanced the real-time performance of the tracking algorithm. Zheng (Yuhui Z et al., 2019) proposed a bilinear structured support vector machine algorithm. Through early classification training, this algorithm significantly improved the computing speed, thereby achieving better real-time performance in the subsequent target - tracking process.

Currently, the traditional industrial robotic arms employ an interceptive tracking strategy for dynamic workpiece tracking. Machine vision systems capture workpiece images, and trajectory prediction algorithms are used to forecast the workpiece's motion trajectory and velocity (Jie P et al., 2026). Upon receiving control commands, the robotic arm prepositions itself to the predicted interception position to grasp the workpiece. However, if the workpiece's velocity or direction changes during operation and the robotic arm cannot react in real time, tracking failure occurs. In practical engineering applications, dynamic tracking systems are typically designed for specific target workpieces. When dealing with different workpieces, the entire system requires reconfiguration, which lacks generalizable workpiece recognition capability and results in limited flexibility (Peng G et al., 2025).

In light of the aforementioned problems, this paper is structured as follows. In the subsequent sections, pre - processing of the original image is performed, which includes grayscale conversion, enhancement, and smoothing. Section 2 presents the implementation process of the traditional SURF algorithm and proposes improvements to address its limitations. Firstly, the salient region is extracted. Secondly, PCA (Principal Component Analysis) dimensionality reduction is conducted. This method reduces high - dimensional data to low - dimensional data, while retaining most of the information from the original dataset, thus ensuring the accuracy of the training process. In Section 4, through experimental verification and analysis, it is demonstrated that the matching accuracy, real - time performance, and robustness of the improved SURF algorithm can meet the real - time requirements of dynamic tracking. Finally, the conclusions of this study are summarized in Section 5.

2. Improvements to the SURF algorithm

2.1. Introduction to the SURF algorithm

Compared to the Scale-Invariant Feature Transform (SIFT) matching algorithm, the SURF algorithm improves computational efficiency to a certain extent, while retaining key advantages of SIFT, such as robustness to variations in illumination, occlusion, and other external conditions.

Owing to these properties, SURF is widely employed for feature matching in complex environments. In machine vision applications, it is commonly used for target object recognition, image registration, and mosaicing. Additionally, the SURF-based feature descriptors facilitate 3D reconstruction by enabling multi-view, comprehensive feature integration for holistic object modeling. Detailed comparisons between the two algorithms are presented in Table 1 below (Yang S, 2021).

Table 1. Comparison of the SURF algorithm and the SIFT algorithm

Project	SURF algorithm	SIFT algorithm
image processing mode	Hessian matrix	Gaussian difference
construction of the scale integral image		multiple convolution
space		
Reference direction	Haar wavelet sign	pixel gradient
stats describing the subset	Haar wavelet sum	modes and direction
the dimension of the feature descriptor	64 Dimensions	128 Dimensions

2.2. Implementation process of the traditional SURF algorithm

The matching flowchart of the traditional SURF algorithm is shown in Figure 1, and the steps are as follows:



Fig. 1. Flowchart of the traditional SURF algorithm matching

Step 1: Construct the Hessian matrix and construct the scale space

The SURF algorithm, built upon the SIFT framework, enhances image matching efficiency by constructing the Hessian matrix for each pixel to identify image keypoints more effectively (Hauagge D C et al., 2012). By computing the determinant and trace of the Hessian matrix, the local intensity variations in pixel neighborhoods can be analyzed to detect potential interest points. The Hessian matrix at a point $f = (x, y)$ is defined by the following equation:

$$H(f(x, y)) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}. \quad (1)$$

After deriving the Hessian matrix using the above formula, the Gaussian kernel size is adjusted to construct a scale space, generating multi-scale images. Thereafter, for each octave in the scale space, three adjacent layers of images are selected. Taking the second-layer image as the reference, each pixel in this layer is treated as a candidate point and compared with to $3 \times 3 \times 3$ neighborhood — including corresponding points in the adjacent upper and lower layers — to identify local extrema of the Hessian determinant.

Step 2: Main orientation of the feature points is determined

The SURF algorithm achieves rotational invariance by assigning each keypoint a dominant orientation, which is determined by the most significant local intensity gradient direction (Zhaozhen J et al., 2020). After generating the multi-scale images, the feature orientation is computed by statistically analyzing Haar wavelet responses within a 60 sector around the keypoint. Specifically, the x - and y -direction gradients are aggregated to construct a gradient orientation histogram, from which the peak value defines the dominant orientation, ensuring descriptor invariance to rotational changes.

Step 3: Description of feature points

Taking the keypoints selected in Step 1 as the center, the local image patch is divided into 4×4 subblocks aligned with the dominant orientation determined in Step 2. For each subblock, Haar wavelet responses (in x and y directions, denoted as dx and dy) are computed using a $2S \times 2S$ kernel, where S represents the scale parameter. The sums of dx , dy , $|dx|$ and $|dy|$ are then aggregated within each subblock to form a local feature vector. This process results in a $4 \times 4 \times 4 = 64$ -dimensional feature descriptor, as the 4 statistical values (sum of dx , dy , $|dx|$, $|dy|$) are calculated for each of the 16 subblocks. The sub-regions, used for feature descriptor construction, are illustrated in Figure 2.

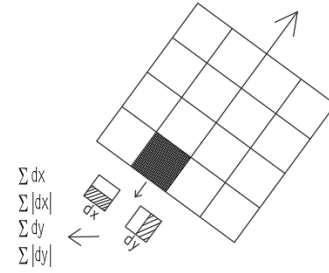


Fig. 2. Characteristic points describe the subregions

Step 4: Feature point matching

Feature matching plays a critical role in target recognition, image registration, and 3D reconstruction, where it is used to locate corresponding keypoints across images (Xin L et al., 2020; Tao F et al., 2025; Linxiong W et al., 2024). The matching process involves comparing the Euclidean distance between the feature vectors of two candidate keypoints. A smaller distance indicates higher similarity between the features, while a larger distance suggests dissimilarity. To determine successful matching, a predefined distance threshold is employed: if the distance between two feature vectors falls below this threshold, the keypoints are considered a valid match; otherwise, they are rejected.

The traditional SURF algorithm assesses the feature point correspondence by comparing the Euclidean distance between feature vectors during image matching (Qiang Z et al., 2018). However, relying solely on this metric often leads to a significant number of false matches. To address this, a salient region extraction is introduced in this paper to filter out non-salient correspondences that are unlikely to be valid. Additionally, the traditional SURF algorithm employs 64-dimensional feature descriptors (Kunming X, 2020), which incur substantial computational overhead during matching. To meet the real-time requirements of dynamic tracking systems, the authors propose dimensionality reduction of feature descriptors using principal component analysis (PCA). This approach reduces descriptor dimensionality while preserving the majority of discriminative information from the original feature space, enabling faster and more efficient matching. The implementation process of the improved SURF algorithm is shown in Figure 3.

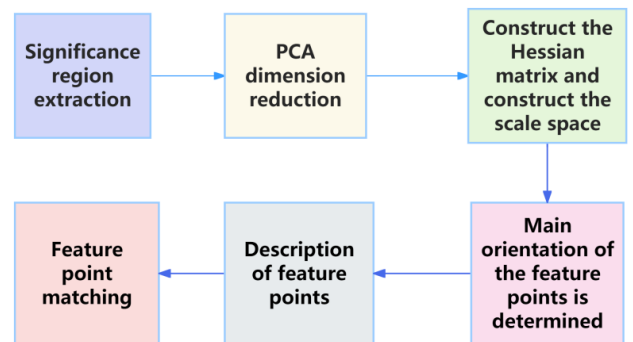


Fig. 3. Flowchart of the improved SURF algorithm

2.3. Increased Salient region extraction

The global contrast-based image saliency detection (LC) (Zhuoqun X et al., 2019) operates by computing the global contrast of each pixel within the entire image. The saliency value $S(I_k)$ of a pixel I_k in image I , is defined as follows :

$$SalS(I_k) = \sum_{I_i \in I} \|I_k - I_i\|. \quad (2)$$

In equation(2), the variable I_i represents a grayscale value with a range of [0,255], which is the standard grayscale intensity interval for 8-bit images. For a given grayscale image, the grayscale value of each pixel I_k is denoted as a_m . Substituting $I_k = a_m$ into the equation, it can be further rewritten as equations (3) and (4):

$$SalS(I_k) = \|a_m - a_0\| + \|a_m - a_1\| + \dots + \|a_m - a_n\|, \quad (3)$$

$$SalS(I_k) = \sum_{n=0}^{255} f_n \|a_m - a_n\|. \quad (4)$$

In formula(4), f_n denotes the frequency of the n -th pixel value in the image. The visualization result of salient region detection is presented in Figure 4.

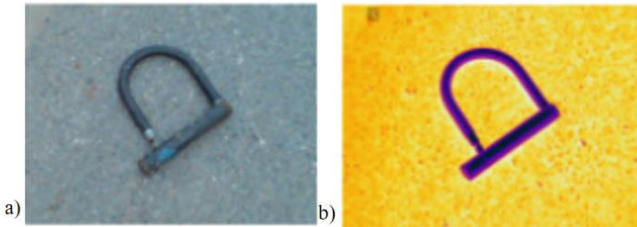


Fig. 4. Bicycle lock significant area detection effect; a) original image; b) salience map

Figure 4 illustrates that salient region detection effectively suppresses complex background noise in images, identifying the bicycle lock as the salient region to eliminate substantial background interference. Experimental results demonstrate that this approach mitigates the effects of color, illumination, and scale variations during feature matching to a certain extent, enhancing the robustness of the matching process.

2.4. Bring in PCA dimension reduction

Principal component analysis (PCA) is one of the most widely used dimensionality-reduction and multivariate statistical techniques (Roque OA et al., 2022). High-dimensional data inherently incur substantial computational complexity; however, simply removing dimensions arbitrarily may lead to significant information loss. To address this, dimensionality reduction methods aim to project the high-dimensional data onto a lower-dimensional subspace, while retaining the majority of discriminative information from the original dataset to ensure analytical accuracy (Lateef J et al., 2026; Cai E et al., 2026) Compared to the Linear Discriminant Analysis (LDA) and T-distributed Stochastic Neighbor Embedding (t-SNE), the core advantages of Principal Component Analysis (PCA) reside in its superior computational efficiency, strong adaptability to large-scale datasets, capacity to maximize the preservation of global data variance, and its inherent nature as an

unsupervised general-purpose dimensionality reduction technique. The core differences among these three dimensionality reduction methods are elaborated in detail in Table 2.

Table 2. Core differences table between PCA, LDA and t-SNE

Project	PCA	LDA	t-SNE
Core principle	Maximization of projection variance	Maximization of inter-class separation	Preservation of local structure
Supervisory	Unsupervised	Supervised	Unsupervised
Main objective	Dimensionality reduction and data visualization	Feature dimensionality reduction prior to classification tasks	High-dimensional data visualization
Computational complexity	Relatively low	Relatively low	Extremely high
Retained structure	Global structure	Class-discriminative information	Local structure
Interpretability of results	Exhibits a certain degree of interpretability	Projection directions are not intended to discriminate between classes and are irrelevant to class discrimination	The low-dimensional embedding results lack interpretability

As a cornerstone in data analysis, the PCA operates by transforming linearly correlated variables into a set of linearly independent principal components via orthogonal transformation. This process maximizes the retention of variance in the original data (Xin C et al., 2023), thereby preserving critical structural information. In this paper, the input vector for PCA dimensionality reduction is the complete 64-dimensional SURF feature descriptor extracted from each feature key point. This descriptor is constructed based on Haar wavelet responses: after performing a Haar wavelet transform on the pixels in the neighborhood of each feature point, the response values and response sums in the horizontal and vertical directions are counted to form a 64-dimensional feature vector. It can fully characterize the grayscale, texture and direction information of feature points, and is the core feature representation form of the SURF algorithm. In the experiment, the original 64-dimensional SURF descriptors of all valid feature points filtered by salient region extraction are retained as the input data for PCA dimensionality reduction, with only normalization preprocessing performed and no additional cropping of feature vectors to ensure the integrity of feature information.

For the selection of the target dimension k of PCA dimensionality reduction, this paper adopts the double criterion of "maximization of variance contribution rate+industrial real-time constraint", combined with the parametric verification of fixed principal component number. Finally, $k=32$ is selected as the optimal dimensionality reduction dimension, and the specific selection process and criteria are as follows:

(1)Variance contribution rate screening: PCA analysis is performed on the sample set of 64-dimensional SURF feature descriptors, and the variance contribution rate and cumulative variance contribution rate of each principal component are calculated. The principal component dimensions with a cumulative variance contribution rate $\geq 95\%$ are retained, and three candidate dimensions of $k=28, 32$, and 36 are initially screened. The cumulative variance contribution rate reaches 96.23% at $k=32$ and 97.15% at $k=36$, with an increase of only 0.92% . This ensures that most of the discriminative information of the original features is retained after dimensionality reduction, avoiding the loss of feature information caused by dimensional compression.

(2)Industrial real-time constraint: Based on the candidate dimensions, the algorithm running time of each dimension is tested under the same experimental environment. It is required that the matching time of the algorithm after dimensionality reduction meets the real-time requirement of industrial dynamic tracking, excluding the situation where the running time exceeds the limit due to excessively high dimensions.

(3)Parametric verification and trade-off: Parametric verification of double indicators is carried out for the candidate dimensions. The results show that at $k=32$, the algorithm maintains a high cumulative variance contribution rate of 96.23% , with a matching accuracy of 97.46% and a running time of only $0.32s$. Compared with $k=28$ (92.15% accuracy), the accuracy is significantly improved; compared with $k=36$ ($0.41s$ running time), it better meets the real-time requirement, which is the optimal balance point of variance retention, matching accuracy and running efficiency.

In the experiment, the cumulative variance contribution rate was calculated based on 1200 valid feature points after salient region extraction of the mixed dataset of Type I and Type II workpieces. The sample set covers the workpiece features under different illumination and angles to ensure the representativeness of variance statistics; the running time test of all dimensions was repeated 10 times to take the average value, excluding the influence of accidental factors on the timing results.

Suppose the dataset contains m data points with the original feature dimension $n=64$ for each data point, and the target dimension after dimensionality reduction is determined as k according to the above criteria. The procedural steps for feature descriptor dimensionality reduction based on the PCA are as follows:

Step 1: Construct a data matrix X , where each row corresponds to the Haar wavelet response (Paula U et al., 2020) vector of a pixel, and each column represents a specific feature dimension in the sample vector. The data matrix with m rows (samples) and n columns (features) is defined as follows :

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix}. \quad (5)$$

Step 2: Centralize the data matrix X by subtracting the column-wise mean from each element within its respective column, yielding a zero-mean matrix X' with the same dimension

of $m \times n$. The column mean and centralized matrix are defined as follows: equation (6) and equation (7).

$$\bar{x}_i = \frac{1}{m} \sum_{j=1}^m x_{ji} (i = 1, 2, \dots, n), \quad (6)$$

$$X' = \begin{bmatrix} x_{11} - \bar{x}_1 & x_{12} - \bar{x}_2 & \cdots & x_{1n} - \bar{x}_n \\ x_{21} - \bar{x}_1 & x_{22} - \bar{x}_2 & \cdots & x_{2n} - \bar{x}_n \\ \vdots & \vdots & \vdots & \vdots \\ x_{m1} - \bar{x}_1 & x_{m2} - \bar{x}_2 & \cdots & x_{mn} - \bar{x}_n \end{bmatrix}. \quad (7)$$

Step 3: Calculate the covariance matrix E of the centralized matrix X' , where the element e_{ij} represents the covariance between feature x_i and feature x_j . The covariance matrix is defined as follows:

$$E = \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1n} \\ e_{21} & e_{22} & \cdots & e_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ e_{n1} & e_{n2} & \cdots & e_{nn} \end{bmatrix}. \quad (8)$$

Step 4: Perform eigenvalue decomposition on the covariance (Zongqing H, 2024; Salton A T et al., 2023; Shushan Z et al., 2013) matrix to compute its eigenvalues λ_k and the corresponding eigenvectors v_k associated with each eigenvalue.

Step 5: Sort the eigenvectors in descending order from left to right by column according to their corresponding eigenvalues to form an eigenvector matrix (Das D et al., 2024), and take the first k columns to form a matrix M , which is an $n \times k$ order matrix.

Step 6: Calculate the sample feature matrix Y reduced to k dimensions by the formula $Y = X'M$, where Y is an $m \times k$ order matrix, i.e., the feature descriptor matrix after PCA dimensionality reduction.

After dimensionality reduction via Principal Component Analysis (PCA), the extracted feature descriptors are projected into a lower-dimensional feature space while preserving the maximal variance of the original features, leading to more efficient and robust matching of image features. A comparative analysis of the feature distribution between the PCA-reduced features and traditional Speeded-Up Robust Features (SURF) is presented in Figure 5.

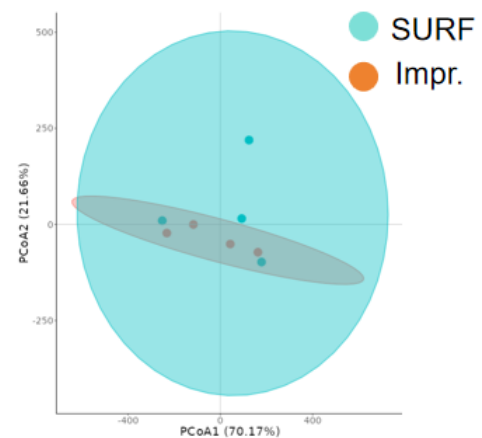


Fig. 5. Comparative Analysis of Feature Descriptors: Projection in Principal Component Space (PCA) versus Classical SURF Representation

The proposed method is compared with the PCA-SURF algorithm proposed by Pancham (Pancham A et al., 2015) for clear methodological positioning. The PCA-SURF in the literature directly applies PCA to reduce the dimension of the 64-dimensional SURF descriptor without any preprocessing of feature points. In contrast, the improved algorithm introduces salient region extraction as a front-end filtering step before PCA dimensionality reduction. Furthermore, the dimension selection follows a double-criterion strategy: maximal cumulative variance contribution rate + industrial real-time constraint, which determines the optimal dimension $k=32$. The accuracy-runtime trade-off curves under different PCA dimensions are given in Figure 6 to validate this selection. The proposed method is specifically designed for dynamic multi-workpiece recognition in industrial scenarios, while conventional PCA-SURF is a general-purpose matching approach. These differences lead to higher accuracy, fewer mismatches, and better real-time performance for robotic vision-based workpiece tracking.

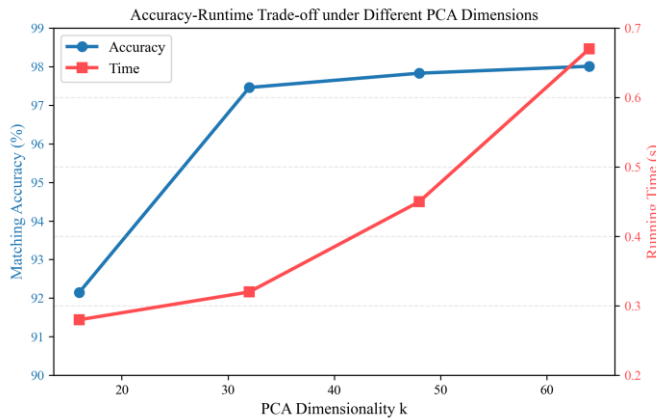


Fig. 6. Accuracy-runtime trade-off curves under different PCA dimensions

3. Experiment and discussion

3.1. Experimental environment

The experimental setup is detailed in Table 3. To evaluate the matching accuracy, real-time performance, and robustness of the improved SURF algorithm, in this section, Type I and Type II workpieces are employed as case studies for experimental validation and analysis. The two distinct workpiece datasets, utilized in the experiments, are visualized in Figure 7. All timing tests are repeated 10 times with a warm start to avoid system fluctuation. A ground truth match is defined as a correctly matched feature pair whose geometric reprojection error is no larger than 2 pixels, which serves as the unified criterion for calculating precision, recall, and F1-score. Random Sample Consensus (RANSAC) is employed to estimate the homography matrix for mismatch removal. In this work, RANSAC is used to estimate the homography matrix between image pairs to eliminate mismatches. The detailed RANSAC parameters are set as follows:

- (1) Reprojection error threshold: 3.0 pixels;

- (2) Maximum number of iterations: 2000;

- (3) Stopping criterion: confidence level of 0.99 or convergence of the inlier set.

Table 3. Experimental environment configuration

Name	Environment configuration
Operating system	Windows 10-64-bit
CPU	AMDR7-5800H, 3.2GHz
Programming environment	Python 3.7
Library function	OpenCV 3.4.2.16
RAM	16 GB
GPU acceleration	No

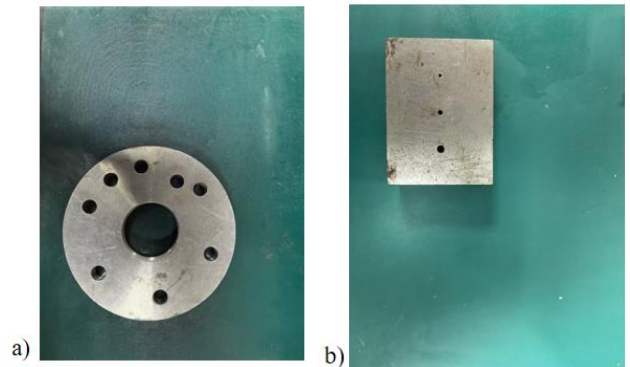


Fig. 7. Collected images of Type I and Type II workpieces; a) Type I; b) Type II

To ensure the reproducibility and comparability of all experimental results, a unified set of fixed parameters is adopted throughout the algorithm implementation and testing. These parameters include key thresholds for feature detection, dimensionality reduction, feature matching, and outlier rejection. The complete list of core algorithm parameters is summarized in the following table 4.

Table 4. Key Fixed Parameters of the Proposed Algorithm

Parameter	Value
Hessian matrix threshold	400
Salient region detection method	Global contrast (LC)
PCA reduced dimension k	32
Feature matching distance ratio	0.75
RANSAC reprojection threshold	3.0 pixels
RANSAC maximum iterations	2000
Ground truth match criterion	Reprojection error ≤ 2 pixels

3.2. Case 1: Comparative analysis of image matching for type I part

The objective of the image matching experiments in this section is to evaluate the performance of the improved SURF algorithm. Using a Type I workpiece in the laboratory as a case study, in this section, the matching effectiveness and success rate of the improved algorithm are analysed through the comparative image matching with the traditional SURF method are analysed. The visualization result of salient region detection for the Type I workpiece is presented in Figure 8. The matching results of the traditional SURF algorithm and

the improved algorithm proposed in this paper are shown in Figures 9 and 10, respectively.

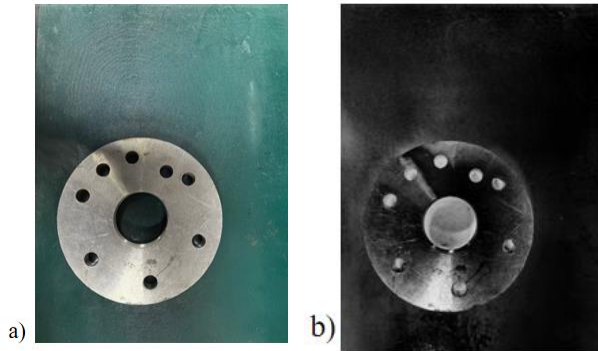


Fig. 8. Result of significant area detection for I-type workpiece; a) original image; b) salience map

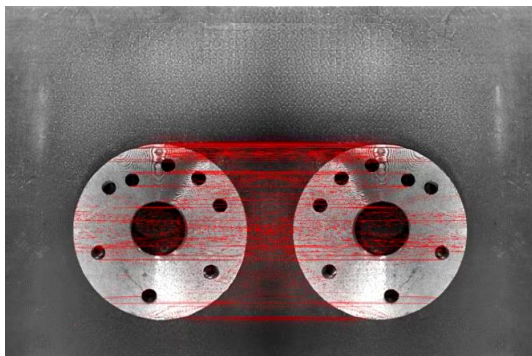


Fig. 9. Traditional SURF feature matching results for type I artifact

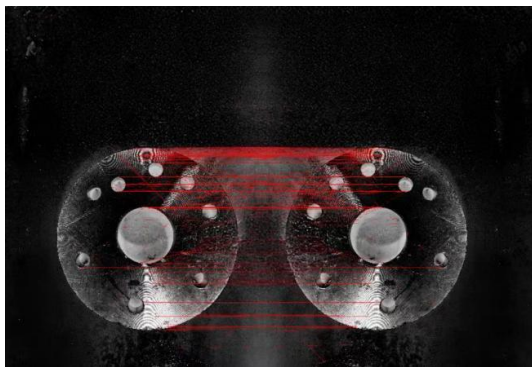


Fig. 10. Match results of the improved SURF algorithm for type I artefact

As demonstrated in Figures 9 and 10, the improved SURF algorithm exhibits a significant enhancement in matching precision for Type I workpiece images, compared to the traditional SURF method. Through timing measurements of the host computer algorithm, the improved approach is shown to have effectively reduced the runtime, enabling the real-time operation in the dynamic tracking system proposed in this paper. Quantitative results of the image matching experiments — including accuracy metrics and computational efficiency — are summarized in Table 5 of the Appendix and Figure 11.

As shown in Table 5, the improved algorithm achieves a matching accuracy exceeding 95% and a 56% increase in processing speed compared to the baseline SURF method.

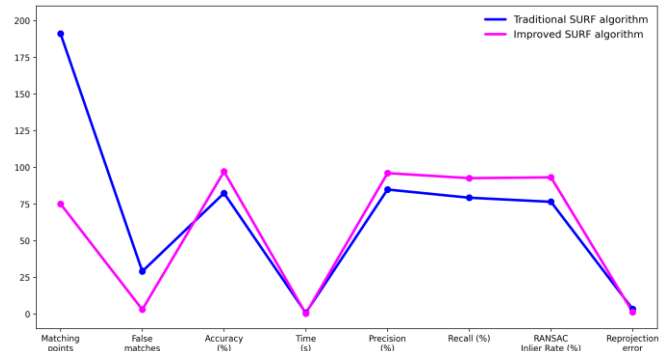


Fig. 11. Image Comparison of Pre- and Post-Algorithm Performance Improvement for Type-I Part

These results demonstrate that the proposed modifications significantly enhance both matching precision and computational efficiency, thereby improving the overall performance of the SURF algorithm. The substantial gains in accuracy and speed enable the improved algorithm to meet the real-time requirements of dynamic tracking systems, showcasing its practical utility in resource-constrained applications.

3.3. Case 2: Comparative analysis of image matching for type II part

To validate the robustness of the improved SURF algorithm, image matching experiments on a Type II workpiece are conducted using the methods described above. The salient region detection visualization for the Type II workpiece is shown in Figure 12. The matching results of the traditional SURF algorithm and the proposed improved SURF algorithm are presented in Figures 13 and 14, respectively.

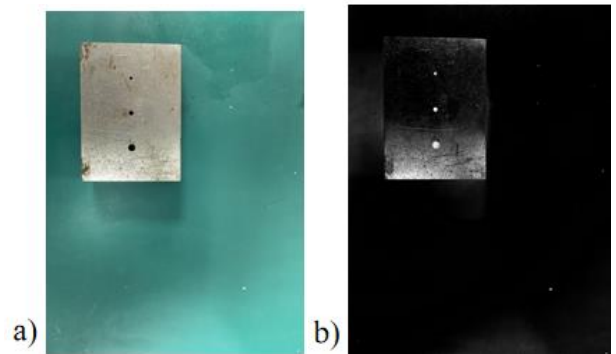


Fig. 12. II type workpiece salient region detection result; a) original image; b) salience map

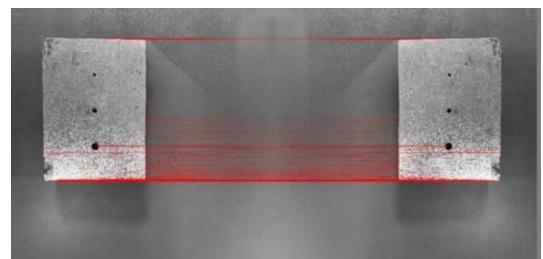


Fig. 13. Traditional SURF feature matching results for Type II workpiece

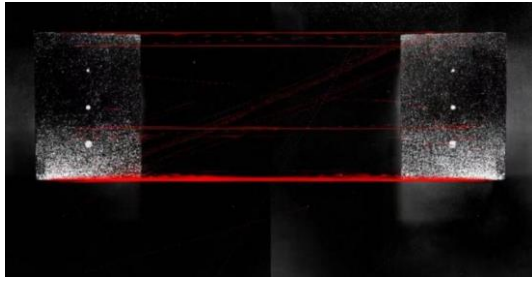


Fig. 14. Improved SURF algorithm matching results for Type II workpiece

As demonstrated in Figures 13 and 14, the improved image matching algorithm maintains effective correspondence for different Type II workpieces under varying conditions, exhibiting high matching precision. Quantitative results of the image matching experiments, including accuracy metrics and robustness evaluations, are summarized in Table 6 of the Appendix and Figure 15.

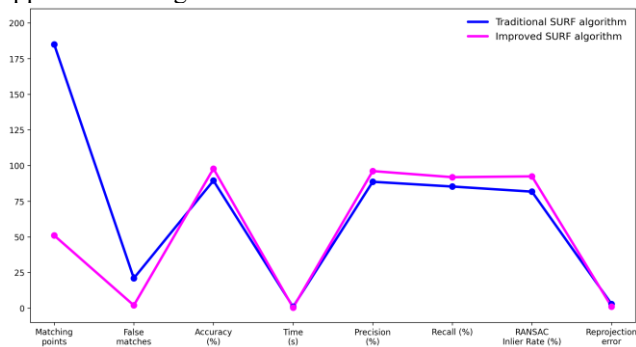


Fig. 15. Image Comparison of Pre- and Post-Algorithm Performance Improvement for Type-II Part

As shown in Table 6, the proposed algorithm achieves a matching precision of 97% for Type II workpiece and a 55% improvement in processing speed compared to the baseline method. These results demonstrate that the improved algorithm effectively generalizes to different workpiece types and exhibits strong robustness across varying operational conditions.

To verify the anti-interference performance of the proposed improved SURF algorithm in complex industrial scenarios, experimental images of Type I and Type II workpieces were collected for three typical interference conditions: illumination variation, image blur and surface occlusion. Ten repeated matching experiments were conducted for each interference type, and the mean and standard deviation of the experimental results were calculated. Thus, the algorithm performance indicators of different workpiece types under various interferences were obtained, as shown in Table 7 of the Appendix.

It can be seen from the experimental results in the table that under three typical interference conditions, including illumination variation, image blur and surface occlusion, the improved SURF algorithm maintains an average matching accuracy of over 93% for both Type I and Type II workpieces, with the average reprojection error all below 2 pixels. In addition, the standard deviation of each index is small, indicating that the algorithm is less affected by external interferences and exhibits excellent stability. Meanwhile, although there are slight

differences in the performance indicators between Type I and Type II workpieces, the overall fluctuation range is extremely small, which further verifies that the algorithm has good anti-interference adaptability for different types of workpieces and can meet the practical application requirements in complex industrial scenarios.

To validate the superiority and competitiveness of the improved SURF algorithm presented in this work, a thorough performance comparison is conducted against three representative feature extraction and matching methods, including the traditional SURF, classic SIFT (Lowe G D, 2004) and Oriented FAST and Rotated BRIEF (ORB) (Rublee E et al., 2011) algorithms. All experiments are implemented on the identical workpiece dataset, experimental platform, parameter settings, and evaluation metrics to guarantee the fairness and standardization of the comparative results, as quantitatively exhibited in Table 8.

Table 8. Performance comparison between the proposed algorithm and SIFT/ORB baselines

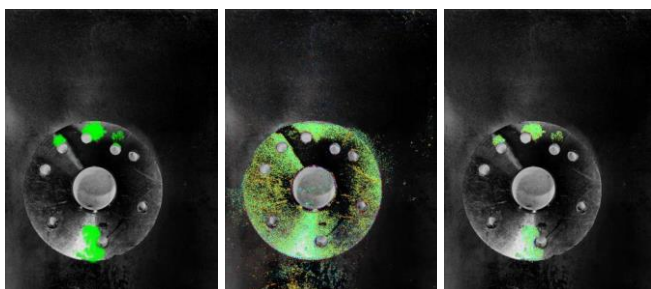
Algo-rithm	Accu-racy (%)	Time (s)	Preci-sion (%)	Recall (%)	F1-score	RANSAC Inlier Ratio (%)	Avg. Reprojection Error (pixels)
SIFT	83.54	1.21	85.12	80.47	0.827	74.68	3.41
ORB	88.71	0.41	89.36	86.52	0.879	82.31	2.75
Tradi-tional SURF	85.82	0.84	87.15	84.36	0.857	79.08	3.07
Proposed Improved SURF	97.46	0.32	96.82	94.17	0.962	93.28	1.11

In contrast to SIFT, ORB, and conventional SURF, the proposed method shows distinct advantages in both efficiency and accuracy. It achieves significantly higher accuracy, precision, recall, F1-score, and RANSAC inlier ratio while consuming less computation time. Moreover, the average reprojection error of the improved algorithm is remarkably lower than that of the three contrast algorithms, indicating more stable and precise feature matching performance. Experimental results show that the proposed algorithm achieves the highest matching accuracy and F1-score, while maintaining the shortest running time. It outperforms SIFT, ORB and traditional SURF in comprehensive performance, which proves its effectiveness for industrial multi-workpiece visual recognition tasks.

3.4. Ablation Experiment Analysis

To quantitatively verify the independent performance improvement effects of two core optimization modules, salient region extraction and PCA dimensionality reduction, in the improved SURF algorithm, as well as their synergistic effect. After the combination, a series of comparative ablation experiments were designed in this study. The dataset, hardware operating environment and core parameters of the SURF algorithm were kept consistent throughout the experiments, and only the combination and adjustment of optimization modules

were carried out. Four groups of control experiments were set up: the traditional SURF algorithm, the SURF algorithm with only salient region extraction, the SURF algorithm with only PCA dimensionality reduction ($k=32$), and the improved SURF algorithm integrating salient region extraction and PCA dimensionality reduction. The experimental algorithm results are shown in Figure 16. The experiments were conducted on a mixed feature dataset of Type I and Type II workpieces, and matching accuracy, algorithm running time, number of mismatched points and F1-score were uniformly selected as the performance evaluation indicators. The F1-score was calculated with the geometric matching error of feature points ≤ 2 pixels as the criterion for true matching, and the statistical results of each group of experiments are shown in Table 9 and Figure 17.



(a)SURF+Salient (b)SURF+PCA (c)SURF+Salient Region+PCA

Fig. 16. Visualization of ablation experiment results

Table 9. Ablation Experiment Results Comparison

Algorithm Type	Matching Accuracy (%)	Running Time (s)	Number of Mismatched Points	F1-Score
Traditional SURF Algorithm	85.82	0.84	25	0.836
SURF+Salient Region Extraction	94.15	0.57	8	0.928
SURF+PCA Dimensionality Reduction ($k=32$)	92.07	0.39	11	0.905
SURF+Salient Region+PCA	97.46	0.32	4	0.962

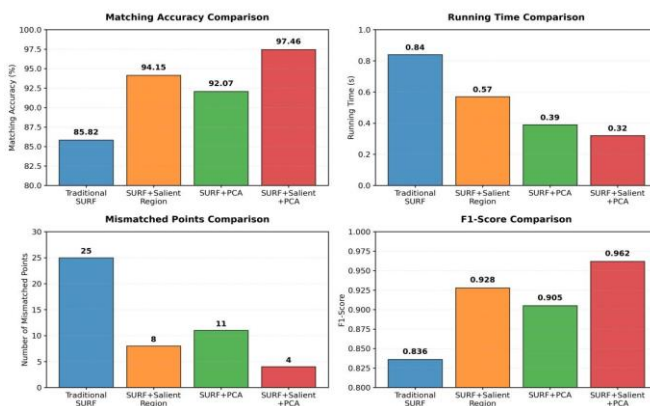


Fig. 17. Ablation Experiment Performance Comparison of the Improved SURF Algorithm

It can be concluded from the ablation experiment results in Table 7 that the introduction of any single optimization module can achieve targeted improvement in the performance of the traditional SURF algorithm. After adding only the salient region extraction module to the SURF algorithm, the matching accuracy is increased by 8.33% compared with the traditional algorithm, the number of mismatched points is reduced by 68%, and the running time is also shortened by 32.14%. This result indicates that salient region extraction can effectively filter out the background noise interference of images in industrial scenarios, eliminate meaningless redundant feature points, reduce the probability of mismatches from the feature extraction stage, and at the same time reduce the amount of computation for subsequent feature matching, realizing the dual preliminary optimization of accuracy and efficiency. The SURF algorithm with only the PCA dimensionality reduction module has a running time shortened by 53.57% and a matching accuracy increased by 6.25% compared with the traditional algorithm, which verifies the dimensional compression effect of PCA dimensionality reduction on the 64-dimensional high-dimensional feature descriptor. On the premise of retaining 96.23% of the cumulative variance of the original features, the computational complexity of the feature matching process is greatly reduced, and the running efficiency of the algorithm is significantly improved.

The improved algorithm proposed in this paper, which integrates the two modules of salient region extraction and PCA dimensionality reduction, realizes the synergistic effect of the two optimization strategies and achieves an optimization effect of $1+1>2$. Compared with the traditional SURF algorithm, the matching accuracy is increased by 11.64%, the running time is shortened by 61.90%, the number of mismatched points is reduced by 84%, and the F1-score is increased by 0.126. Compared with the algorithms with a single optimization module, it not only retains the advantage of salient region extraction in improving matching accuracy, but also inherits the optimization effect of PCA dimensionality reduction in running efficiency without any performance loss. This result fully proves the rationality and effectiveness of the two-step optimization strategy of "salient region extraction + PCA dimensionality reduction" proposed in this paper: the feature points are first screened through salient region extraction to ensure the purity of the feature data participating in the subsequent matching, and then PCA dimensionality reduction is performed on the high-dimensional feature descriptors to avoid information redundancy caused by the dimensional compression of invalid feature points, so that the algorithm achieves the optimal performance in two core dimensions of matching accuracy and running efficiency.

This part of the ablation experiments quantitatively verifies the effectiveness of each core optimization strategy and the rationality of the module combination of the improved SURF algorithm from the perspective of algorithm modules, providing solid experimental support for the improvement of the overall performance of the algorithm. On this basis, this study further carries out physical robot arm recognition and dynamic tracking experiments with the KUKA industrial robotic arm to verify the adaptability and comprehensive performance of the

proposed improved algorithm in actual industrial application scenarios.

4. Robotic Arm Recognition and Tracking Experiment

4.1. Camera calibration

The camera imaging model describes the projection from a 3D world space to a 2D image plane through four coordinate systems, namely the world coordinate system ($O_w - X_w Y_w Z_w$), the camera coordinate system ($O_c - X_c Y_c Z_c$), the image coordinate system ($o - xy$), and the pixel coordinate system ($oi - uv$). The geometric transformations among these systems allow the pixel coordinates of an image point to be converted into world coordinates. Such a transformation relies on two matrix parameters: the intrinsic parameter matrix, which describes the internal optical properties of the camera, and the extrinsic parameter matrix, which represents the pose of the camera frame relative to the world frame. The camera's internal parameters calibration is shown in Table 10.

Table 10. Camera calibration results

Parameter name	Parameter value
Camera internal parameters	$\begin{bmatrix} 5694.246 & 0.000 & 0 \\ -26.523 & 5126.652 & 0 \\ 850.413 & 501.362 & 1 \end{bmatrix}$
Distortion coefficient	Radial : [-0.0258, 1.5216, -103.1125] Tangential : [-0.0021, -0.0047]

In a vision-guided robotic arm system, the motion planning and control of the manipulator are implemented under its base coordinate system. Consequently, the target pose detected by the vision system in the camera frame must be transformed into the robot base frame, which is accomplished by hand-eye calibration. The essential objective of hand-eye calibration is to determine the relative pose between the camera and the robot end-effector or base, and to calculate the corresponding hand-eye transformation matrix.

Using this transformation matrix, the target pose expressed in the camera coordinate system can be uniformly converted into the robot base coordinate system. This enables the robotic arm to acquire the precise 3D position and orientation of the target, and thus achieves stable dynamic tracking and control in practical industrial scenarios.

Considering the practical application requirements and high-precision positioning demands of the proposed system, the Eye-to-Hand configuration was employed for hand-eye calibration in this work. The conversion formula is as follows:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = R \begin{bmatrix} x \\ y \end{bmatrix} + M. \quad (9)$$

In the formula, R represents the rotation matrix in the hand-eye transformation, and M represents the translation matrix in the hand-eye transformation. The above formula can be expressed as:

$$\begin{cases} x' = ax + by + c \\ y' = a'x + b'y + c' \end{cases} \quad (10)$$

Collect several points:

$$\begin{cases} x_0' = ax_0 + by_0 + C \\ x_1' = ax_1 + by_1 + C \\ x_2' = ax_2 + by_2 + C \end{cases}, \quad (11)$$

$$\begin{cases} y_0' = a'x_0 + b'y_0 + C' \\ y_1' = a'x_1 + b'y_1 + C' \\ y_2' = a'x_2 + b'y_2 + C' \end{cases}. \quad (12)$$

The transformation in matrix form is:

$$\begin{bmatrix} x_0 & y_0 & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} x_0' \\ x_1' \\ x_2' \end{bmatrix}, \quad (13)$$

$$\begin{bmatrix} x_0 & y_0 & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{bmatrix} \begin{bmatrix} a' \\ b' \\ c' \end{bmatrix} = \begin{bmatrix} y_0' \\ y_1' \\ y_2' \end{bmatrix}. \quad (14)$$

The main coordinate systems and their conversion relationships are shown in Figure 18.

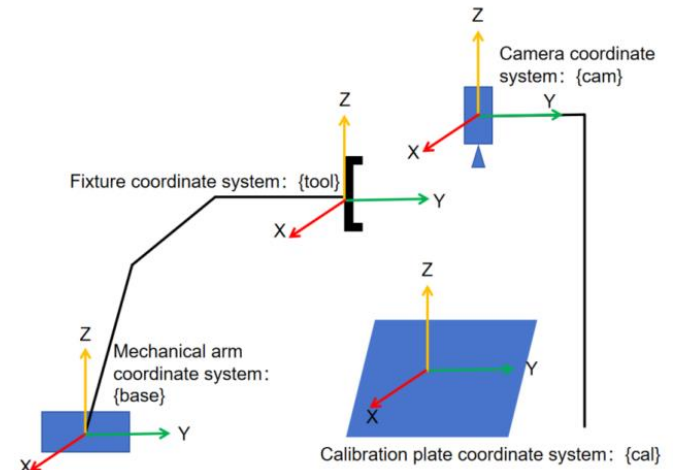


Fig. 18. Hand-eye calibration coordinate system transformation relationship

In this configuration, the experimental setup consists of a KUKA KR20R180 six-degree-of-freedom industrial robotic arm, an XMINI industrial camera, and a standard calibration plate. Calibration is performed using the nine-point calibration method within the VisionMaster software platform. The camera is mounted on a fixed support, and the calibration pin attached to the robot end-effector is required to vertically align with each of the nine calibration points along a predefined path. The two-dimensional visual calibration platform is constructed as illustrated in Figure 19. The pixel coordinate diagram of the calibration plate is shown in Figure 20. The hand-eye transformation matrix is shown in Equation 15.

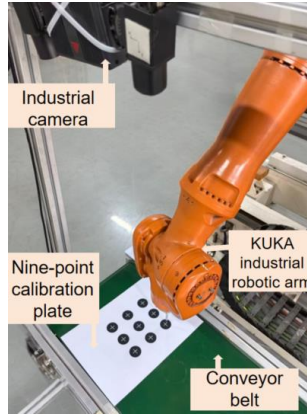


Fig. 19. Hand-eye calibration test bench

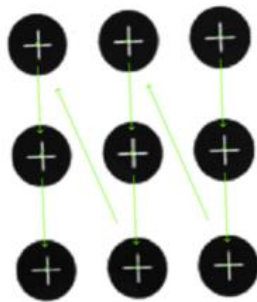


Fig. 20. Calibration plate pixel coordinate diagram

$$\begin{bmatrix} 0.0064813 & -0.187551 \\ -0.0526745 & -0.162554 \\ 0 & 0 \end{bmatrix} \quad (15)$$

4.2. Experiment

The experimental platform, established in a laboratory setting, comprises core equipment, including a KUKA KR20 six-degree-of-freedom industrial robotic arm, a X mini high-speed wide-field industrial camera, and a conveyor unit. The experimental steps are as follows:

Step 1: Activate the conveyor belt for workpiece conveyance.

Step 2: Operate the camera to capture images of the target workpiece.

Step 3: Carry out recognition and localization on the acquired image, and output the pixel coordinates of the target position.

Step 4: Transform the pixel coordinates into the base coordinate system of the robotic arm to accomplish target localization.

Step 5: Manipulate the robotic arm to move to the specified position, and compare this actual position to the identified target position to validate the positioning accuracy.

In the experiment, dynamic path planning was implemented by the robotic arm according to the target pose data delivered by the visual system, while a PID control strategy was adopted to achieve high-precision tracking of the moving workpiece. To validate the stability and system response time performance of the system, repeated experiments were conducted on

Type I and Type II workpieces under different operating speeds and lighting conditions.

During the experiment, tracking tests were conducted under multi-factor variable conditions: the workpiece was set to move at varying speeds on the conveyor belt, slight oil contamination or local occlusions were present on its surface, and the light intensity of the environment was adjusted. For each tracking process, three key performance indicators were recorded: average positioning error, maximum tracking error, and system response time. Specifically, under the standard operating conditions: Type I workpiece, conveyor belt speed of 0.5 m/s, and environmental light intensity of 1000 lux, six consecutive tracking experiments were carried out, the results of which are presented in Figure 21.

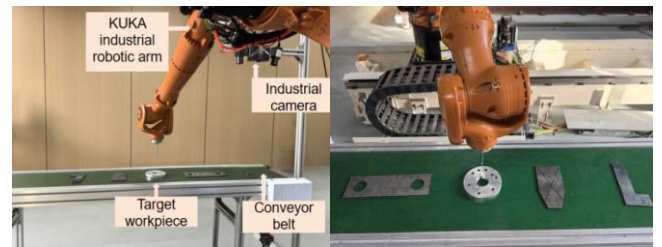


Fig. 21. Recognition and Tracking of Type I Workpiece by Robotic Arm

Experimental data indicate that the improved SURF algorithm-guided robotic arm achieves an average positioning error of 2.42 mm, a maximum tracking error of no more than 4.85 mm, and a stable single-tracking system response time within 0.4 seconds. In contrast, the traditional SURF algorithm-guided system yields an average error of 7 mm, a maximum error of 11.79 mm, and a system response time of 0.65 seconds, as summarized in Table 11.

Table 11. Actual Coordinates versus Identified Coordinates of Type I Workpiece

Actual coordinates X/mm	Actual coordinates Y/mm	Conventional algorithms calculated value of X/mm	Conventional algorithms calculated value of Y/mm	Optimized algorithm to calculate the accurate coordinates of X/mm	Optimized algorithm to calculate the accurate coordinates of Y/mm
656.73	-936.57	652.13	-934.14	655.87	-935.77
684.39	-913.40	680.41	-917.45	683.24	-914.12
521.01	-845.39	510.46	-841.24	524.03	-844.52
589.36	-824.26	578.14	-820.64	584.56	-823.58
623.48	-901.35	620.17	-904.31	622.15	-902.36
648.36	-941.02	645.36	-939.15	649.95	-942.75

For each disturbance condition (illumination variation, image blur, surface occlusion), 10 independent repeated matching experiments are carried out. The results are presented as mean $\pm$ standard deviation for quantitative stability analysis.

For clarity, time-related metrics in this paper are defined as follows:

- (1) Matching time: The duration of only feature extraction and matching. In this experiment, it ranges from 0.32 s to 0.42 s.
- (2) System response time: The total latency from image acquisition to output of the target 2D pixel coordinates.
- (3) Full robot control cycle: The complete period from image capture to the end-effector reaching the target pose. In this experiment, it was approximately 5 seconds. All time measurements are repeated 10 times and averaged to ensure stability.

Under the identical standard operating conditions applied to Type I workpieces, six consecutive tracking experiments were performed on Type II workpieces, with the results presented in Figure 22. The experimental data showed that the improved algorithm achieved an average positioning error of 2.49 mm, a maximum tracking error of 4.07 mm, and a system response time of 0.42s. By comparison, the traditional SURF algorithm produced an average error of 6.75 mm, a maximum error of 7.81mm, and a response time of 0.76 s, as summarized in Table 12. These findings confirm that the improved SURF algorithm possesses excellent tracking accuracy and real-time performance. Even when the conveyor belt speed was elevated to 1.5 m/s, the average error of the improved algorithm increased marginally, and it still met the strict positioning accuracy criteria for practical industrial assembly scenarios.

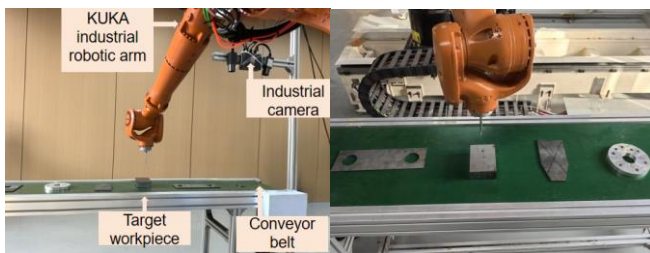


Fig. 22. Recognition and Tracking of Type II Workpiece by Robotic Arm

Table 12. Actual Coordinates versus Identified Coordinates of Type II Workpiece

Actual coordinates of X/mm	Actual coordinates of Y/mm	Conventional algorithms calculated value of X/mm	Conventional algorithms calculated value of Y/mm	Optimized algorithm to calculate the accurate coordinates of X/mm	Optimized algorithm to calculate the accurate coordinates of Y/mm
614.64	-943.84	619.13	-939.14	615.14	-945.50
685.64	-918.28	680.41	-917.45	684.48	-918.41
562.24	-994.83	558.46	-989.24	566.12	-992.74
565.05	-911.59	561.14	-917.64	566.55	-914.31
643.38	-935.36	647.17	-929.31	642.59	-937.51
651.50	-948.50	648.36	-956.15	653.40	-951.34

When 5% of the workpiece surface area is randomly occluded, the traditional algorithm suffers three instances of tracking failure. In contrast, the improved algorithm only produces a slight positioning deviation in one test case, and it rapidly resumes stable tracking following simple adjustments — an outcome that highlights its robust anti-interference performance. Under different lighting environments, the fluctuation range of positioning errors for the improved algorithm is controlled within ± 0.3 mm, whereas the traditional algorithm yields an error fluctuation range of up to ± 0.7 mm. This experimental result further verifies that the improved algorithm maintains high stability even under the complex operating conditions.

The experimental results show that the robotic arm completes a single recognition and positioning cycle with an average system response time of 5 seconds, and the tracking error of the end effector is constrained within ± 2.5 mm, which meets the high-precision operation criteria for industrial field applications. Notably, in the case of random speed variations of the workpiece, the system remains capable of stable recognition and continuous tracking. This not only confirms the reliability of the system but also further verifies the synergistic effectiveness and robustness of the integrated algorithm and control system. In conclusion, the robotic arm recognition and tracking system, combined with the improved SURF algorithm, can effectively cope with diverse interference factors in actual industrial environments, thereby exhibiting substantial practical application value.

5. Summary and conclusion

In this paper, an enhanced SURF algorithm is presented for robust and efficient feature point matching of diverse workpieces in industrial vision systems. The proposed method integrates salient region extraction and PCA dimensionality reduction to suppress background interference, reduce mismatches, and accelerate feature matching.

Experimental results illustrate that the proposed algorithm yields higher matching accuracy and shorter computational time than the conventional SURF algorithm, while exhibiting strong robustness against illumination changes, blur, and partial occlusion. Comparative evaluations with SIFT and ORB further validate its superior overall performance and practical competitiveness. Physical experiments on a KUKA robotic arm verify that the proposed method meets the accuracy and real-time demands of dynamic workpiece recognition and tracking.

In future work, the algorithm will be further optimized and compared with state-of-the-art learning-based matching methods to improve its adaptability in more complex industrial scenarios.

Acknowledgements

This work is funded by the National Natural Science Foundation of China (U24A20135), Supported by Inner Mongolia Natural Science Foundation major project (2023ZD12), University basic research business expenses (2023RCTD012), University basic research busi-

ness expenses (2023QNJS075) , Postgraduate Research Innovation Program and of Inner Mongolia Autonomous Region (KC2024053B), University basic research business expenses (2024YXXS012) , State Key Laboratory of Special Vehicle Design and Manufacturing Integration Technology Open Research Fund Project (GZ2023KF012) , Special scientific research project of first-class discipline(YLXKZX-NKD-023), Special scientific research project of first-class discipline(YLXKZX-NKD-005) and Science and Technology Plan Project of Inner Mongolia Autonomous Region(2026YFHH0156).

Reference

- Agnieszka S, Jacek K, Tomasz B, Piotr O, Krzysztof F, Gabriel K, Grzegorz G, Wojciech J. 2026. Novel Robotic Test Rig for Camshaft Geometry Measurement with a Collaborative Robot. *Sensors*, 26(7):2206. DOI:10.3390/S26072206.
- Anshi M, Guofu Z, Lingwei M. 2024. On-line Detection Method of Dimensional Deviation for Machined Workpieces Based on Laser Sensor. *Laser Journal*, (10):215-220. DOI:10.14016/j.cnki.JGZZ.2024.10.215.
- Cai E, Wu Z, Todd D M, Zhen H. 2026. Underwater image enhancement for vision-based structural dynamic monitoring. *Automation in Construction*, 187:106931. DOI:10.1016/J.AUTCON.2026.106931.
- Chen C, Azman A. 2024. Improved Deep Learning Model for Workpieces of Rectangular Pipeline Surface Defect Detection. *Computers*, 13(1):30. DOI:10.3390/COMPUTERS13010030.
- Das D, Naskar R. 2024. Image splicing detection using low-dimensional feature vector of texture features and Haralick features based on Gray Level Co-occurrence Matrix. *Signal Processing: Image Communication*, 125. DOI:10.1016/j.image.2024.117134.
- Hauagge D C, Snavely N. 2012. Image matching using local symmetry features. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. DOI:10.1109/cvpr.2012.6247677.
- Jide Q, Bin C, Jiye Q, Hengjun Z, Gang C. 2018. Mura defect detection method for liquid crystal screen based on Machine vision. *Computer Science*, 45(6). DOI:10.11896/j.issn.1002-137X.2018.06.052.
- Jie P, Zhengze Z, Qingsong F, Ranfei X, Zheng Y. 2026. Trajectory Prediction-Enabled Self-Decision-Making for Autonomous Cleaning Robots in Semi-Structured Dynamic Campus Environments. *Sensors*, 26(7):2258. DOI:10.3390/S26072258.
- Kan L, Chuanshuai J, Yu C, Lu Y, Baoshan P, Chaobing L, Wu W, Huiyong Z. 2025. An optimized YOLOv8n based model for real time defect detection in taro strip production. *Scientific Reports*, 15(1):44685. DOI:10.1038/S41598-025-28216-X.
- Kunming X. 2020. The dimension reduction algorithm of PCA and CFS based on MIC. *Hubei University of Technology*. DOI:10.27131/d.cnki.ghugc.2020.000549.
- Lateef J, Yu X. 2026. Novel computer vision algorithm for accurate multi-label classification of concrete bridge defects. *AI in Civil Engineering*, 5(1):10. DOI:10.1007/S43503-026-00095-9.
- Likun W, Zi W, Peter K, Gumma K, Turner A, Ratchev S. 2024. Digital-twin deep dynamic camera position optimisation for the V-STARS photogrammetry system based on 3D reconstruction. *International Journal of Production Research*, 62(11):3932-3951. DOI:10.1080/00207543.2023.2252108.
- Linxiong W, Yi T, Ping Y, Wang Z. 2024. Simulation analysis of correlated Hartmann wavefront detection based on feature matching. *Chinese Journal of Lasers*, 51(22):193-204. DOI:10.3788/CJL240482.
- Lowe G D. 2004. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2):91-110. DOI:10.1023/B:VISI.0000029664.99615.94.
- Marfo R, Vesely A. 2025. Human-centered perspectives on trust, usability, and ethical concerns in machine vision applications. *AI and Ethics*, 6(1):84. DOI:10.1007/S43681-025-00950-Y.
- Oztemel E, Gursev S. 2020. Literature review of Industry 4.0 and related technologies. *Journal of Intelligent Manufacturing*, 31(6):127-182. DOI:10.1007/s10845-018-1433-8.
- Pancham A, Withey D, Bright G. 2015. Tracking image features with PCA-SURF descriptors. *Proceedings of the International Conference on Machine Vision Applications*. DOI:10.1109/mva.2015.7153206.
- Paula U, Robinson J, Javier B M. 2020. Algorithm of detection, classification and gripping of occluded objects by CNN techniques and Haar classifiers. *International Journal of Electrical and Computer Engineering*, 10(5):4712-4720. DOI:10.11591/IJECE.V10I5.PP4712-4720.
- Peng G, Renliang L, Zehao F, Jiaying C. 2025. Research on recognition and localization of workpieces on rotating body surfaces based on vision recognition algorithm. *Journal of Physics: Conference Series*, 2951(1):012074. DOI:10.1088/1742-6596/2951/1/012074.
- Qiang Z, Songqi H, Weibo Y. 2018. Workpiece image feature matching based on improved SURF Algorithm. *Journal of Changchun University of Technology*, 39(4):373-377. DOI:10.15923/j.cnki.cn22-1382/t.2018.4.10.
- Roque O A, Arturo J Y, Alvaro A I. 2022. Fault detection and classification in kinematic chains by means of PCA extraction-reduction of features from thermographic images. *Measurement*, 197. DOI:10.1016/J.MEASUREMENT.2022.111340.
- Rublee E, Rabaud V, Konolige K, Bradski G. 2011. ORB: An efficient alternative to SIFT or SURF. *Proceedings of the International Conference on Computer Vision*. DOI:10.1109/iccv.2011.6126544.
- Salton A T, Pimentel G A, Melo J V, Castro R S, Benfica J. 2023. Data-driven Covariance Tuning of the Extended Kalman Filter for Visual-based Pose Estimation of the Stewart Platform. *Journal of Control, Automation and Electrical Systems*, 34(4):720-730. DOI:10.1007/S40313-023-01006-4.
- Shushan Z, Juan P W. 2013. The Improvement and Application of Intelligence Tracking Algorithm for Moving Logistics Objects Based on Machine Vision Sensor. *Sensor Letters*, 11(5):862-869. DOI:10.1166/sl.2013.2658.
- Tao F, Huiqin W, Ke W, Rui L, Zhan W. 2025. A Multimodal Bone Stick Matching Approach Based on Large-Scale Pre-Trained Models and Dynamic Cross-Modal Feature Fusion. *Applied Sciences*, 15(15):8681. DOI:10.3390/AP15158681.
- Wei G. 2023. Research on Workpiece Identification and Positioning System Based on Machine Vision. *Nanjing Information Engineering University*. DOI:10.27248/d.cnki.GNJQC.2023.000278.
- Wenjian C, Yan Z, Liu L, Jin Q, Chunchi M, Tianbin L, Liuyuan Z, Ziyi Z. 2026. High-precision identification of groundwater states in tunnel: An image recognition-driven two-stage neural network approach. *Results in Engineering*, 29:109336. DOI:10.1016/J.RINENG.2026.109336.
- Xin C, Yiwei H, Peng X. 2023. Parameter estimation of the structured illumination pattern based on principal component analysis (PCA): PCA-SIM. *Light: Science & Applications*, 12(2):150-152. DOI:10.1038/s41377-022-01043-9.
- Xin L, Yang W, Zhou F, Qi X, Haibo Y, Huizhen Y. 2020. Adaptive optical image post-processing based on multi-channel blind recognition. *Acta Photonica Sinica*, 49(2):71-79. DOI:10.3788/gzxb20204902.0201003.
- Yang S. 2021. Research on the key technology of robot carton palletizing based on machine vision. *Shandong University of Technology*. DOI:10.27276/d.cnki.gsdgc.2021.000484.
- Yuhui Z, Le S, Shunfeng W, Jianwei Z, Jifeng N. 2019. Spatially Regularized Structural Support Vector Machine for Robust Visual Tracking. *IEEE Transactions on Neural Networks and Learning Systems*, 30(10):3024-3034. DOI:10.1109/TNNLS.2018.2855686.
- Zhaozhen J, Yanhou Z, Hao Z. 2020. Image registration and change detection method based on wavelet transform and SURF algorithm. *MIPPR: Automatic Target Recognition and Navigation*, 11429. DOI:10.1117/12.2542897.
- Zhisheng D, Qing P, Xiaoyun Z. 2016. Measurement of Workpiece Size and Angle based on Machine Vision. *Computer Measurement and Control*, 24(2):27-29,41. DOI:10.16526/j.cnki.11-4762/tp.2016.02.007.
- Zhiyuan Z. 2022. Research on the dynamic object detection and tracking algorithm of the workpiece sorting mechanical arm. *Lanzhou Jiaotong University*. DOI:10.27205/d.cnki.gltcc.2022.000206.
- Zhuoqun X, Zhihong X. 2019. Over-exposure area detection algorithm based on modified saturation features. *Electronic Technology and Software Engineering*, (11):68-70. DOI:10.20109/j.cnki.ets.2019.11.050.
- Zongqing H. 2024. Based on covariance block structure of the tensor learning. *University of Science and Technology of China*. DOI:10.27517/d.cnki.gzjkj.2024.002678.

Appendix

Table 5. Systematic Performance Comparison of Algorithms for Type-I Part

Algorithm	Matching points	False matches	Accuracy[%]	Time[s]	Precision(%)	Recall(%)	RANSAC Inlier Rate (%)	Average re-projection error (pixels)
Traditional SURF algorithm	191	29	82.36	0.75	84.81	79.23	76.44	3.26
Improved SURF algorithm	75	3	97.05	0.23	96.00	92.56	93.12	1.18

Table 6. Systematic Performance Comparison of Algorithms for Type-II Part

Algorithm	Matching points	False matches	Accuracy[%]	Time[s]	Precision(%)	Recall(%)	RANSAC Inlier Rate (%)	Average re-projection error (pixels)
Traditional SURF algorithm	185	21	89.47	0.94	88.65	85.32	81.73	2.89
Improved SURF algorithm	51	2	97.86	0.42	96.15	91.87	92.45	1.05

Table 7. Anti-Interference Stability Analysis Table

Type of interference	Type-I Parts			Type-II Parts		
	Number of experiments (N)	Average matching accuracy (%) $\pm$ Standard deviation	Average reprojection error (pixels) $\pm$ standard deviation	Number of experiments (N)	Average matching accuracy (%) $\pm$ Standard deviation	Average reprojection error (pixels) $\pm$ standard deviation
Illumination variation	10	95.32 $\pm$ 1.28	1.32 $\pm$ 0.26	10	94.97 $\pm$ 1.35	1.38 $\pm$ 0.29
Image blur	10	94.76 $\pm$ 1.53	1.48 $\pm$ 0.31	10	94.31 $\pm$ 1.61	1.55 $\pm$ 0.33
Surface occlusion	10	93.85 $\pm$ 1.87	1.65 $\pm$ 0.38	10	93.42 $\pm$ 1.95	1.72 $\pm$ 0.41