Robotica (2016) volume 34, pp. 925–941. © Cambridge University Press 2014
doi:10.1017/S0263574714001970

Target-tools recognition method based on an image feature library for space station cabin service robots

Lingbo Cheng†‡§, Zhihong Jiang†‡§, Hui Li†‡§*, Bo Wei†‡§ and Qiang Huang†‡§

†*IRI, School of Mechatronic Engineering, Beijing Institute of Technology, Beijing, China*
‡*Key Laboratory of Biomimetic Robots and Systems, Ministry of Education, China*
§*Key Laboratory of Intelligent Control and Decision of Complex System, China*

(Accepted Jun 19, 2014. First published online: July 28, 2014)

SUMMARY
This paper presents a method to improve the speed and accuracy rate for space robot visual target recognition based on illumination and affine invariant feature extraction. The method takes illumination changes, strong nonlinear light due to refraction and reflection, target affine transformation and occlusion into consideration, all of which occur on the cabin target surface and affect the target recognition accuracy seriously. In this paper, a method is proposed to capture a same target at multi-viewpoints to establish feature library for high recognition accuracy and speed at any viewpoint. By using an analysis of the light intensity and gray level transformation, we obtain the corrected image which reduce the influence of illumination change. Then the affine moment invariants features of the correction images at multi-viewpoints were extracted and the average feature datum were stored in the library. To verify the validity of the method, a robot vision system provided images, followed by image preprocessing, dynamic local threshold segmentation and feature extraction. These methods were verified on a target recognition system of space robot built for this research. The experimental results showed that the methods were feasible and effective.

KEYWORDS: Feature extraction; Object recognition; Feature library; Robot vision; Space station.

1. Introduction

Astronauts use video and instrumentation to monitor a space station cabin. Cabin maintenance and operation involve various movements, such as touching, plugging, and pressing.[1,2] However, the space environment is so harsh, such as microgravity and narrow space, and current technological limitations on a space station, that the astronauts cannot stand a long-term retention in the space station and they often bear a heavy burden of operation tasks and physical challenges. Therefore, there is an urgent need to develop a service robot capable of providing assistance or even replacing astronauts to conduct monitoring, maintenance and operations on a space station. Although this task presents significant technical challenges, it also offers the strong possibility of increased safety for astronauts as well as economic benefits.[3,4]

Intelligent vision systems are essential both for monitoring and also for the autonomous intelligent operations of a space robot. Vision systems allow space station cabin service robots to complete maintenance tasks while under the monitor and control of astronauts or ground controllers.[5] Maintenance tasks include compartment status monitoring, circuit board replacement, instrumentation control, repair tool application and computer operation. Although robot vision technology is considered relatively mature, factors in a space station cabin, such as lighting conditions, illumination changes near the porthole and microgravity still pose significant challenges for target recognition based on robot visual systems.[6] For example, strong reflection and refraction of light in the space station cabin can occur, which may lead to strongly nonlinear of the images collected by the robot

* Corresponding author. E-mail: lihui2011@bit.edu.cn

https://doi.org/10.1017/S0263574714001970 Published online by Cambridge University Press

visual system. Specifically, the changing of illumination near the porthole could cause serious interference to the recognition system. Also, because of the effect of microgravity, recognizable objects may undergo affine transformation,[5] which can cause the target recognition task to fail and the accuracy rate to decline. Therefore, the development of a new target recognition method that is capable of overcoming illumination and target affine transformation is a top priority.

This paper studies target recognition of service robots for common tools used in a space station. Specifically, we are focused on solving the target recognition problem under different illumination and affine transformation. There are two factors to be considered.

The first one is the changing illumination. As the object recognition based on global features (such as: color, sharp, context, etc.) is sensitive to lighting geometry and illumination, so our purpose is to extract the illumination invariance or to reduce the sensitivity under different illuminations. In ref. [7] Funt and Finlayson extended Swain's color indexing method[8] to be illumination independent by indexing on an illumination-invariant set of color descriptors based on color ratios among neighboring pixels. They assumed the neighboring pixels have similar surface normal and achieved the satisfactory results. But the experimental conditions did not include different object viewpoints. Later, Nayar and Bolle proposed the illumination-independent color ratios which might be negatively affected by the geometry and pose of the object.[9] In ref. [10] Gevers and Smeulders proposed a derivative-based color invariant which is both robust to variations of illuminant color and lighting geometry. However, a small perturbation of sensor values will cause a large jump in the transformed values. The above illumination invariant features are global invariance entirely. In ref. [11] Gevrekci and Gunturk researched the image brightness invariant based on local properties and proposed a method of illumination robust feature extraction transform (IRFET). IRFET has a great robust illumination changes, while it does not obtain affine and photometric invariant feature detectors.

The second factor is affine transformation. M. K. Hu first proposed the concept of 'moments' based on the rules of the Cartesian coordinate system and obtained seven moment invariants, having invariants of scale, translation and rotation.[12] However, the Fundamental theorem of affine invariants in Hu's paper was stated incorrectly. After that, many researchers extended the definition of moments to the boundary of objects. In 1993, Reiss[13] and Flusser and Suk[14,15] independently published new sets of the affine moment invariants (AMI) which has already have a good application in character recognition and scene matching.[16,17] The AMI's can be derived in several ways that differ from each other by mathematical tools used.[18] Suk and Flusser[19] presented a general method how to automatically generate the AMI's of any weights and orders. This method is based on representation of the AMI's by graphs. The major contribution of the paper is the possibility of identifying and eliminating dependent invariants. However, the proper invariants selection is always data-dependent and cannot be generalized to other data sets.

Besides, there are several researches studied the invariants under both varying illumination and affine transformations. For instance, D. Lowe extracted target interest points based on local invariants which was called scale invariant feature transform (SIFT).[20,21] The SIFT features improve on previous approaches by being largely invariant to changes in scale, illumination and affine distortions. However, the number of the threshold is excessive and they are difficult to determine. Furthermore, the high dimension of feature descriptor will lead to complicated calculation. Although many experts have made improvements based on SIFT,[22] the calculations involved are extremely large.

A method is proposed that takes into account the illumination instability and microgravity of the space environment. An analysis of light intensity and gray level transformation for image processing is combined with affine moment invariant feature extraction. This process provides the target features with illumination and affine invariants. In order to improve the speed and accuracy rate for space robot visual target recognition, this paper presents research on image acquisition, illumination changes and affine invariant feature extraction, as well as a feature matching algorithm and a recognition method.

The paper is organized as follows: Section 2 describes a space station cabin robot system and target recognition system scheme. Section 3 establishes the image feature library. Section 4 introduces the target recognition method. Section 5 presents a series of experiments used to establish the proposed method and feature library. Finally, Section 6 presents the conclusions and outlines future research.

https://doi.org/10.1017/S0263574714001970 Published online by Cambridge University Press

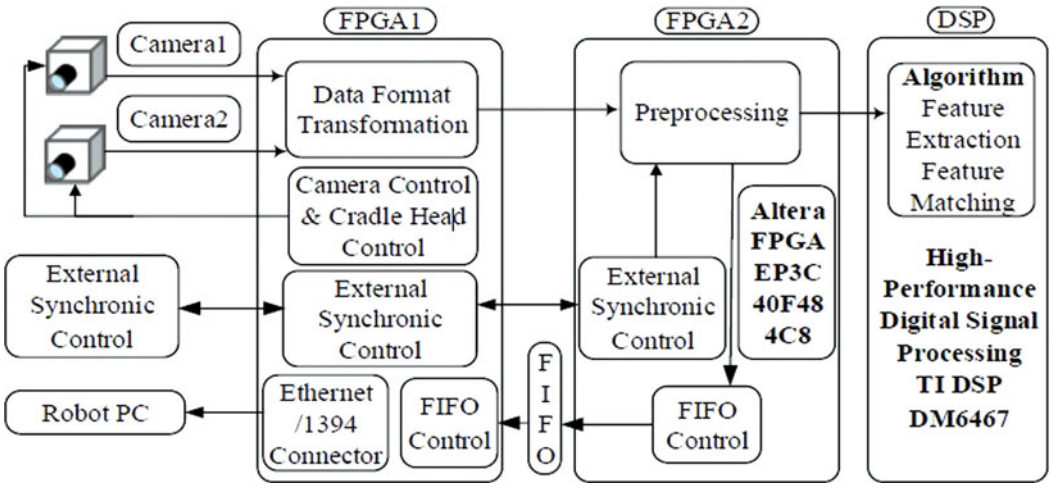


Fig. 1. Experimental system for robot object recognition.

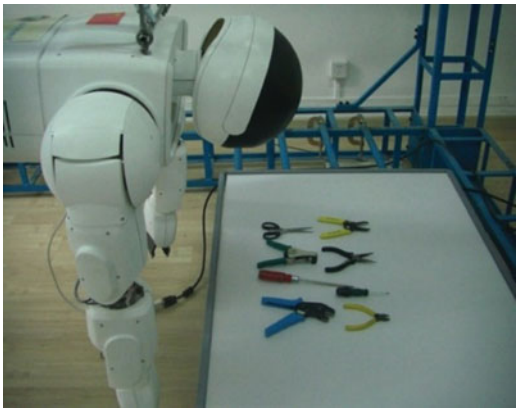


Fig. 2. Toolbox for robot object recognition.

2. System Scheme

2.1. Robot system

Fast and accurate target recognition is fundamental to the success of a spacecraft cabin service robot when acting as an assistant or replacement. This paper proposes a target recognition method based on an image feature library in order to better accommodate the confined enclosure of the space station environment and also the limited number of recognizable targets.

Common repair tools were chosen as the targets for recognition and were used to establish the feature library. A target recognition system for a space station cabin service robot was built to test target recognition (Fig. 1).

The system was based on an existing service robot (Fig. 2) and used an MER-125-30UC/UM-L 1/3”CCD binocular digital camera and a image processor that combined DSP and FPGA. The signal processing board included two FPGA chips and one digital signal processor. The chip models were Altera FPGA EP3C40F484C8 and TI DSP TMS320 DM6467 Da Vinci.

The system signal processing-board received the input image signals via Ethernet, with a 1394 standard interface for communication with the outside world. It was also equipped with an external signal synchronization and trigger I/O. The FPGA chips simultaneously performed the image capture, compression, preprocessing and transmission. The DSP accepted the input signals of the sample images obtained by the FPGA and then obtained the sample features using a feature extraction algorithm. The chips then outputted the extracted features to the robot PC to establish the image feature library. During the target recognition process, the DSP received the signals of the captured images and the external sync signals, which were feature library vectors, using a feature-matching

https://doi.org/10.1017/S0263574714001970 Published online by Cambridge University Press

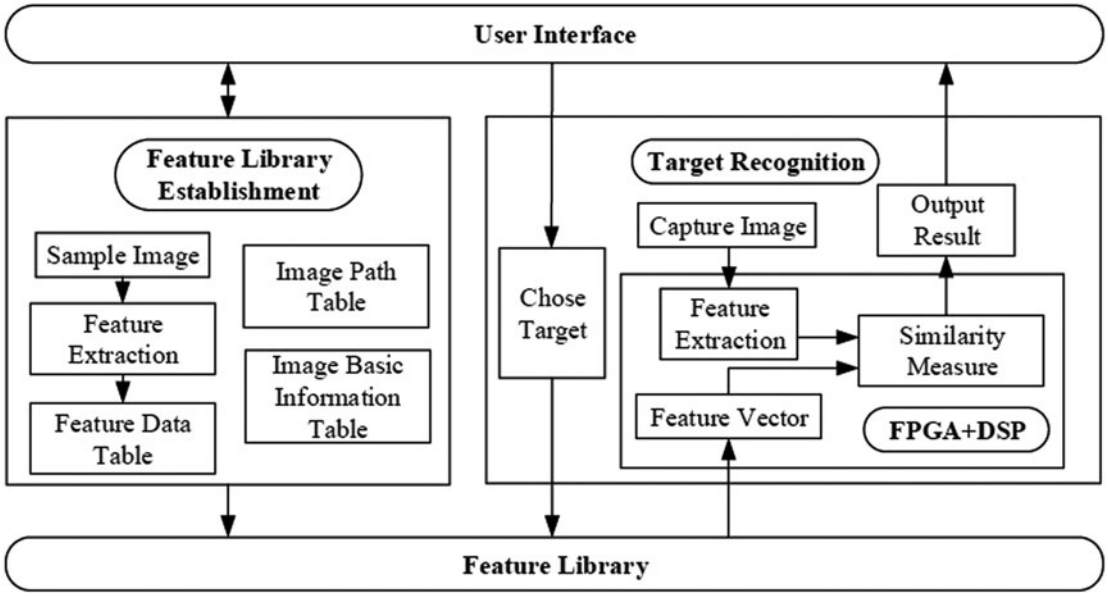


Fig. 3. Flowchart for object recognition.

algorithm. Then the target recognition results were obtained. Finally, the chips fed back the results to the robot PC.

2.2. Target recognition system scheme

The complete image feature library based on the target recognition was divided into two parts: feature library establishment and target recognition. Fig. 3 shows the system scheme of the target recognition.

The target feature library included the image path tables, the basic information tables and the image feature data tables. First, the target image for storage in the feature library was selected. Next, the sample images of the selected target were captured using a digital camera. Finally, the sample images were preprocessed and the features were extracted and the data stored. Establishing the image path tables and the basic information tables allowed easier extraction of the information in the feature data tables.

Fig. 3 shows the target recognition procedures. In the first step, the desired target was selected via the user interface. Next, the target was transmitted to the feature library. The target feature data in the database was obtained through the database retrieval system. This step was prepared for the subsequent feature matching. In the second step, the digital camera captured the image and then input it into the image processor, which contained DSP and FPGA. In the third step, the DSP corrected the illumination brightness of the image based on the average gray level and then extracted the feature vectors of each foreground target. In the final step, the similarity measure of each target feature vector and the vectors in the feature library were calculated. The target that corresponded to the minimum Euclidean distance was the recognized object. Then output the result.

3. Feature Library Establishment Method

The establishment of the feature library is important for the target recognition process. The selection of the library storage data and the number of data have a direct influence on the speed and accuracy rate of the feature matching. This paper presents a new method with illumination stability and affine invariant to extract the feature vector. First, multi-images from nine viewpoints were captured. The new images were obtained at normal illumination conditions using gray level transformation of each image. Next, three affine moment invariants were extracted from each new image. In order to eliminate the micro changes of the feature vector caused by the different viewpoints, average moment invariants were calculated. Finally, three average affine moment invariants were obtained as the vector for storage in the library. This method effectively enhanced the target recognition speed and accuracy rate.

https://doi.org/10.1017/S0263574714001970 Published online by Cambridge University Press

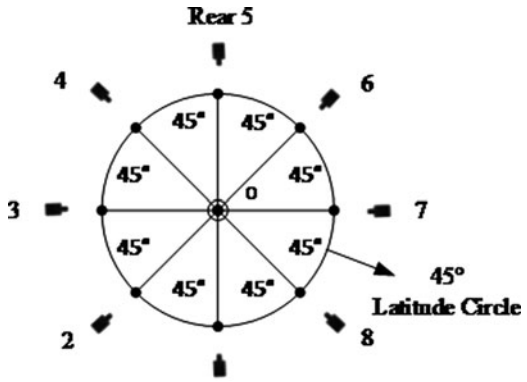


Fig. 4. Camera positions for systematic comparison.

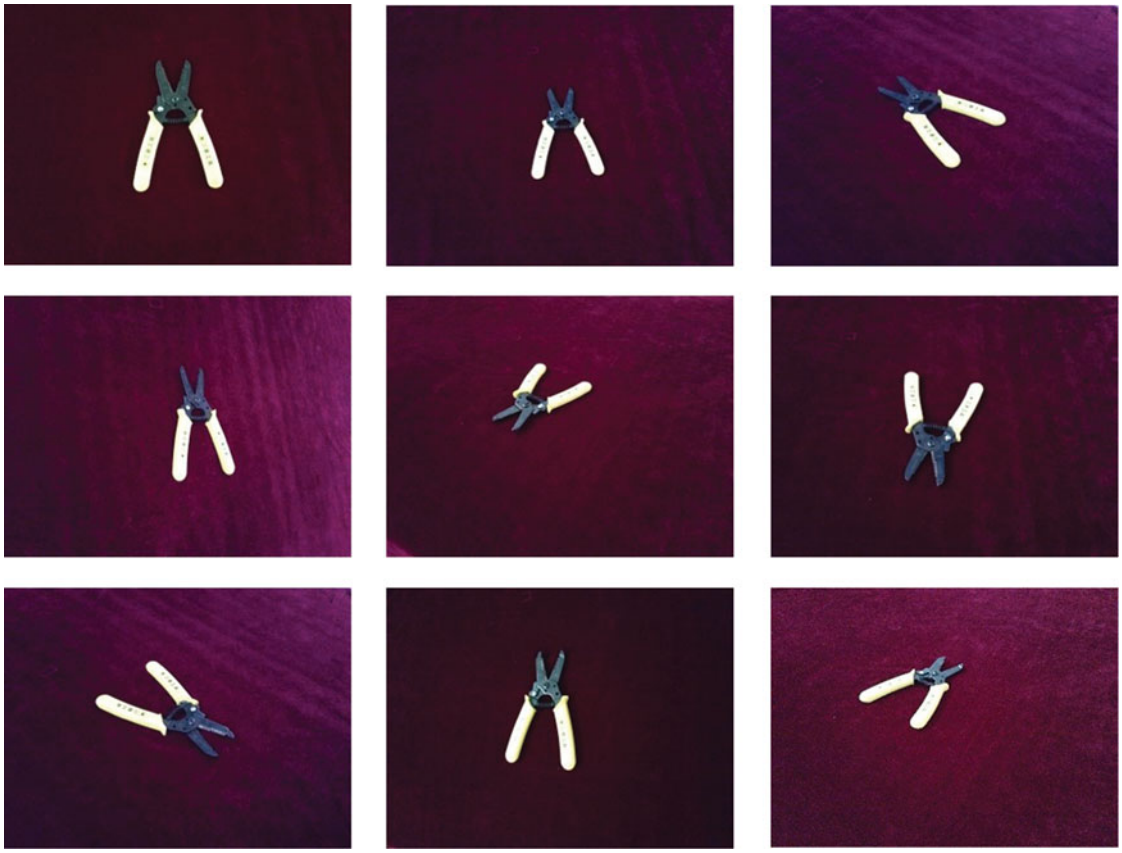


Fig. 5. Images captured from different viewpoints.

3.1. Image capture

During target recognition process, the robot captured images from various viewpoints. The image feature was extracted and matched with the feature data stored in the feature library. In order to address the micro changes that occur from the different viewpoints, a method was used for capturing multi-images from nine viewpoints. The captured target images where input into the storage library, which decreased the feature change and increased the accuracy rate. The illumination was normal and there was no interference during image capture process.

Nine images were captured from nine viewpoints. Each image capture was the same distance from the target and camera. Based on the size of the target, a distance of 500 mm was selected. Fig. 4 shows the camera positions for a systematic comparison of the different viewpoints at the top view. A common pair of pliers was used as the sample target. Nine sample images were captured (Fig. 5).

https://doi.org/10.1017/S0263574714001970 Published online by Cambridge University Press

Table I. Three illumination conditions.

Illuminance/(lux)	Illumination condition
250	Strong
150	Medium
50	Weak

A sample image was captured from the first viewpoint, which was directly above the target object. Fig. 5 labels this viewpoint as 'Viewpoint 0'. Next, the camera was moved to a latitudinal circular pathway and rotated in 45-degree increments along this pathway at a distance of 500 mm away from the target. This procedure was repeated for viewpoints 1 to 8 (Fig. 4). In total, there were 9 viewpoints.

3.2. Feature extraction algorithm

After capturing the images, the sample images were preprocessed to obtain the new gray images with the illumination correction. Next, the affine moment invariants were extracted; this eliminated the influence of both illumination and the affine invariants. This paper presents a new method of combining the gray level transformation and the affine moment invariants feature extraction algorithm.

3.2.1. Gray level transformation.

A gray level transform function stretched the contrast ratio in order to obtain a sharp contrast ratio image for feature extraction. The whole process contained three parts: first, determining the illumination condition based on illumination intensity; second, calculating the average gray value of the image and obtaining the gray level transform function; third, transforming the sample image using the obtained function. The new gray image was then used for feature extraction in the next step.

① Determine the illumination condition based on illuminance.

By controlling the number of light sources in the experimental environment and measuring the corresponding illumination intensity, we can quantitative analysis the illumination conditions. There are three illumination conditions divided based on opening different light sources.(Table I)

② Calculate the average gray value and obtain the gray level transform function.

Assume the size of image f is $M \times N$ and $f(x, y)$ is the gray value of each pixel; therefore, the average gray value of the image is:

$$\bar{f} = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} f(x, y) \tag{1}$$

Normalize the gray value of image $f(x, y)$ to [0,1] to provide the normalized image $F(x, y)$. The Sigmoid function was used as the gray level transform function:

$$F_c(x, y) = \frac{1}{1 + e^{-\gamma[F(x,y)-c]}} \tag{2}$$

Where γ represents the level of incline and c is the center value of the contrast function. Guntruk's paper[11] presented the selection of different parameters γ and c provided the different transform images with various contrast ratios.

Calculate the average gray values of the captured images in three illumination conditions and obtain the thresholds for the target which were 80 and 120 through repeated experiment. After referencing the paper[11] and combining the illumination condition the following values were chosen: $\gamma = 50$ and $c = 0.2$ or 0.3 or 0.5. Table II shows the relation between illumination conditions and parameters, γ and c.

https://doi.org/10.1017/S0263574714001970 Published online by Cambridge University Press

Table II. Relation between illumination and parameters.

Illumination condition	Average gray value	γ	c
Strong	$\bar{f} > 120$	50	0.5
Medium	$80 < \bar{f} < 120$		0.3
Weak	$\bar{f} < 80$		0.2

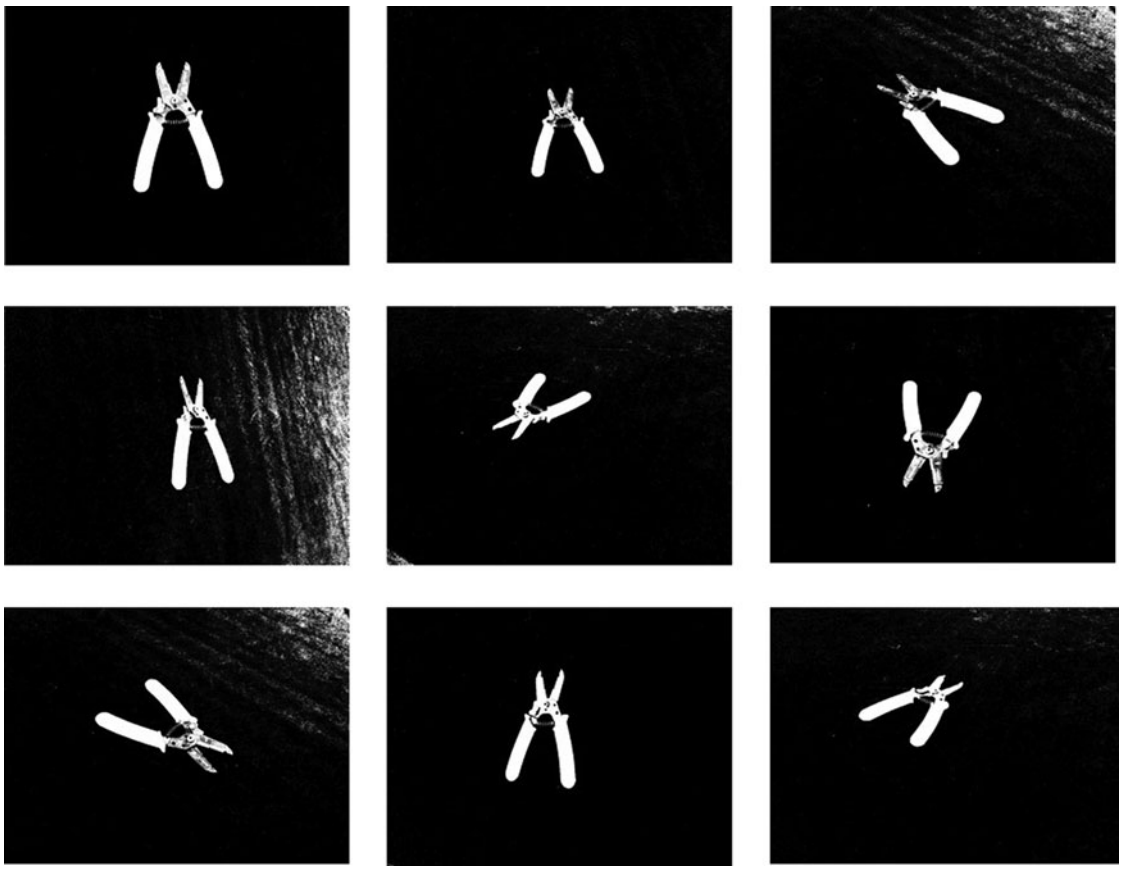


Fig. 6. Gray level transformation images.

The experimental results showed that the assignment of these parameters obtained images with optimal contrast ratio.

③ Transform the sample images using the obtained function.

The gray level transform function was used to process the sample images and obtain the new gray images, which were then used for feature extraction in the next step. Fig. 6 shows the new gray images.

3.2.2. Affine moment invariants.
The target position change can be described by the affine transformation model:[14]

$$\begin{cases} u = a_0 + a_1 x + a_2 y \\ v = b_0 + b_1 x + b_2 y \end{cases} \tag{3}$$

The theory of algebraic invariants was used to derive the method of affine moment invariant extraction. The affine transformation (3) can be deconstructed into the following six transformations with

https://doi.org/10.1017/S0263574714001970 Published online by Cambridge University Press

one-parameter:

$$\mathbf{1.}\begin{cases} u = x + \alpha \\ v = y \end{cases} \quad \mathbf{2.}\begin{cases} u = x \\ v = y + \beta \end{cases} \quad \mathbf{3.}\begin{cases} u = \omega x \\ v = \omega y \end{cases},$$

$$\mathbf{4.}\begin{cases} u = \delta x \\ v = y \end{cases} \quad \mathbf{5.}\begin{cases} u = x + ty \\ v = y \end{cases} \quad \mathbf{6.}\begin{cases} u = x \\ v = t'x + y \end{cases}$$

If the affine moment invariants satisfied the above six formulas, then they were invariant under general affine transformation.

Formula (3) is described as :

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} x_a \\ y_a \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} = \mathbf{A} \begin{bmatrix} x \\ y \end{bmatrix} + \mathbf{B} \tag{4}$$

Where **A** is a second order singular matrix, including rotation, scaling and oblique transformation parameters; and **B** is the translation vector. The two-dimensional (p+q)th order moments m_{pq} and the central moments $\mu_{pq}(p \geq 0, q \geq 0)$ are defined as:

$$m_{pq} = \sum_{(x,y)\in R} x^p y^q f(x, y) \tag{5}$$

and

$$\mu_{pq} = \sum_{(x,y)\in R} (x - \bar{x})^p (y - \bar{y})^q f(x, y) \tag{6}$$

Where $(\bar{x}, \bar{y})$ is the coordinate of the center of gravity of a given object, which are $\bar{x} = \frac{m_{10}}{m_{00}}$, $\bar{y} = \frac{m_{01}}{m_{00}}$.

The normalized moments are defined as:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^r}, \quad r = \frac{p+q}{2} + 1, \quad p + 1 = 2, 3, \ldots \tag{7}$$

The affine moment invariants were built using center moments. Special theorems for invariants provided the three required invariants. Proofs can be found in ref. [23].

Theorem 1: Apolar

$$\frac{1}{2} \sum_{i=0}^{k} (-1)^i \binom{k}{i} \mu_{i,k-i} \mu_{k-i,i} \tag{8}$$

If k is even, it is invariant after normalization.

Theorem 2: Hankel determinant

$$\begin{vmatrix} \mu_{0,k} & \mu_{1,k-1} & \cdots & \mu_{q-1,k-q+1} \\ \mu_{1,k-1} & \mu_{2,k-2} & \cdots & \mu_{q,k-q} \\ \vdots & & & \\ \mu_{u-1,k-u+1} & \mu_{u,k-u} & \cdots & \mu_{u+q-2,k-u-q+2} \\ \mu_{0,k'} & \mu_{1,k'-1} & \cdots & \mu_{q-1,k'-q+1} \\ \mu_{1,k'-1} & \mu_{2,k'-2} & \cdots & \mu_{q,k'-q} \\ \vdots & & \cdots & \\ \mu_{v-1,k'-v+1} & \mu_{v,k'-v} & \cdots & \mu_{v+q-2,k'-v-q+2} \\ \vdots & & & \end{vmatrix} \tag{9}$$

The Hankel determinant is invariant after normalization, where $k = q + u - 2, k' = q + v - 2$.

https://doi.org/10.1017/S0263574714001970 Published online by Cambridge University Press

Table III. Affine moment invariants at different viewpoints.

Viewpoint	$I^1/10^{-2}$	$I^2/10^{-6}$	$I^3/10^{-4}$
0	4.7361	−17.642	−12.648
1	4.9785	−23.672	−14.072
2	4.1443	−19.651	−9.6748
3	3.8880	−15.552	−8.5092
4	3.3524	−7.4247	−5.5980
5	4.4995	−14.391	−10.026
6	3.6942	−9.3469	−6.8261
7	4.1840	−15.673	−9.6699
8	3.8365	−11.585	−7.8046
Average	4.1459	−14.993	−9.4254

Theorem 3: Discriminant of polynomial

$$\sum_{i=0}^{k} \binom{k}{i} \mu_{i,k-i} x^{k-i} \tag{10}$$

After dividing by $\mu_{0,k}$ and normalization is invariant.

The above theorems provide the invariants from second-order and third-order moments:

$$I_1 = \left(\mu_{20}\mu_{02} - \mu_{11}^2\right)/\mu_{00}^4 \tag{11}$$

$$I_2 = \left(\mu_{30}^2\mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + 4\mu_{30}\mu_{12}^3 + 4\mu_{21}^3\mu_{03} - 3\mu_{21}^2\mu_{12}^2\right)/\mu_{00}^{10} \tag{12}$$

$$I_3 = \left[\mu_{20}\left(\mu_{21}\mu_{03} - \mu_{12}^2\right) - \mu_{11}\left(\mu_{30}\mu_{03} - \mu_{21}\mu_{12}\right) + \mu_{02}\left(\mu_{30}\mu_{12} - \mu_{21}^2\right)\right]/\mu_{00}^7 \tag{13}$$

Three features, $I_1 I_2 I_3$, were used for object recognition in the following feature library construction and experimentation.

3.2.3. Feature analysis.

Under the same conditions, nine images were captured from different viewpoints (Fig. 4). The gray level transformation function was used to transform each image (Fig. 5); the affine moment invariants of the new images (Fig. 6) were then extracted. Table III shows the 27 obtained features. The parameter was $c = 0.3$ because the capture occurred with general illumination.

The numbers in the first row represent the nine camera positions. The corresponding features are the affine moment invariants at each viewpoint. To decrease the calculations and increase the accuracy ratio, the average of the 27 invariants was calculated to obtain three affine moment invariants. Table I shows the values of invariants, I^1,I^2 and I^3. The feature vector was stored in the library as: $(I^1, I^2, I^3) = (4.1459, -14.993, -9.4254)$.

3.3. Feature library design method

The feature database was designed after obtaining the average values of the target affine moment invariants. Also, the path table and index were constructed to obtain the data quickly during the object recognition process. The image feature library included the feature data tables, the image path tables and the image basic information tables (Fig. 7).

The image path tables were designed to allow easy access to the library data. The image basic information tables were used to store basic information of the library images. Each image record contained N extendable attributes (Table IV).

The affine moment invariants of the object were stored in the feature data tables. They were used to match features for object recognition. Table V shows parts of the information of the feature data tables. Each image record included five attributes, three of which were the invariants.

The image path tables were used to find the image basic information tables and the database. The basic information and the feature vector in the feature library were also obtained. Table VI shows the eight procedures used to realize object recognition.

https://doi.org/10.1017/S0263574714001970 Published online by Cambridge University Press

Table IV. Attribute description for basic image information.

Function name	Data type	Function
ID	Int	Record the NO.
Img_Name	Varchar	Image name
Date	Varchar	Date
Time	Varchar	Time
Resolution	Int	Resolution ratio
Sto_form	Varchar	Memory format

Table V. Attribute description for image feature database.

Function name	Data type	Function
ID	Int	Record the NO.
Img_Name	Varchar	Image name
Img_feat_1	Varchar	NO.1 feature vector
Img_feat_2	Varchar	NO.2 feature vector
Img_feat_3	Varchar	NO.3 feature vector

Table VI. Description of program classes.

Items	Function
Con_data	Connect the feature library
W_Name	Select the recognizing target
Inva_Date	Get the feature vector in the library and save it
RIBF	Read image
O_S_Info	Open and save the information of the image
Inva_Desc	Extract the affine moment invariants of the image
Feat_Matc	Feature matching
Result	Output the result

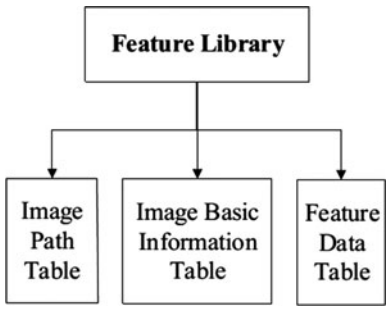


Fig. 7. Target feature library.

4. Target Recognition Method

The target recognition process was completed using the image feature library. This process contained image acquisition, preprocessing, feature extraction and feature matching.

4.1. Target image input

Changes to the viewpoint of the robot camera allowed the robot visual system to capture images under different illumination, affine transformation, multi-target interference and partial occlusion. The average gray level was used to determine the condition of the image and then transform the image gray value. This process allowed the gray image to be obtained for feature extraction.

https://doi.org/10.1017/S0263574714001970 Published online by Cambridge University Press

Table VII. Affine moment invariants and distances in different illumination.

Items	S^1	S^2	S^3	d	$d/\|\mathbf{I}\|$
1	4.3003	−12.172	−10.286	2.9534	0.1624
2	4.2922	−12.198	−10.439	2.9767	0.1637
3	4.4498	−16.931	−11.094	2.5753	0.1416

4.2. Feature extraction and matching method

The feature extraction was different from the processing of the feature library establishment because it has to segment the objects separately in addition to segment the objects and the background. First, the objects and the background were segmented in HSV color space which has the best divisibility performance compared with other color spaces. Then, a dynamic partial threshold segmentation method was used to segment the target and interference objects based on histogram, threshold and shape, etc., and the local region was determined by the size of smooth filter. This kind of segmentation allowed the targets to be separated from the background and ensured that each target was independent, and the invariants of each target were then extracted. The target invariants were recorded as S_i^1, S_i^2, S_i^3, where i was the serial number of the target.

The feature vector stored in the feature library was retrieved through the image path tables after the object recognition target was selected. Then, the vector was recorded as $\mathbf{I} = (I^1, I^2, I^3)$. Finally, the feature vector of the image was extracted and recorded as $\mathbf{S}_i = (S_i^1, S_i^2, S_i^3)$. The minimum distance in the 3-D Euclidean feature space provided the classification:

$$d = |\mathbf{S}_i - \mathbf{I}| = \sqrt{\sum_{j=1}^{j=3} \left(S_i^j - I^j\right)^2} \tag{14}$$

In order to eliminate the impact of the vector module on the Euclidean distance, d was divided by $|\mathbf{I}|$. The threshold selected was 0.25; if $d/|\mathbf{I}| < 0.25$, then the corresponding target is the optimal match.

5. Experimental Results

In order to show the performance of the affine moment invariants after gray level transformation for object recognition, an image feature library was established and used for experimentation. The experiments demonstrated that even with different illumination, affine transformation, complex conditions and comprehensive factors, effective and accurate experimental results were still achieved for object recognition.

5.1. Target recognition results under different illumination

To test the sensitivity of our method with varying illumination, three images were captured under three different illumination at a same viewpoint which were shown in Fig. 8 (a). The viewpoint can be arbitrary, as the feature library was established by the average value of nine features at nine viewpoints. The purpose of this experiment is to obtain the target recognition accuracy when the illumination condition changed. The test images were 640*480 pixels in size.

The illumination conditions for each image were determined by the average gray values and each illumination condition corresponded to a parameter c**.** As the average gray value of captured image under illumination 1 is more than 120, based on the above algorithm parameter $c = 0.5$. While the average gray value of illumination 2 is just between 120 and 80, parameter $c = 0.3$. The image under illumination 3 has the lowest average gray value and it is less than 80 for the weakest illumination, so parameter $c = 0.2$. Using parameters c and γ, gray level transform function is obtained and the transformed images used to be extracted affine moment invariants are shown in Fig. 8 (b) with the values of c.

Table VII shows the values of affine moment invariants and the distances for each of the three different illuminations. The values of $d/|\mathbf{I}|$ showed that each match was accurate for all of them were less than the threshold 0.25. In total, 45 images were obtained by capturing images of the pliers from various viewpoints under different illumination conditions. An accuracy rate of 95.6% was obtained

https://doi.org/10.1017/S0263574714001970 Published online by Cambridge University Press

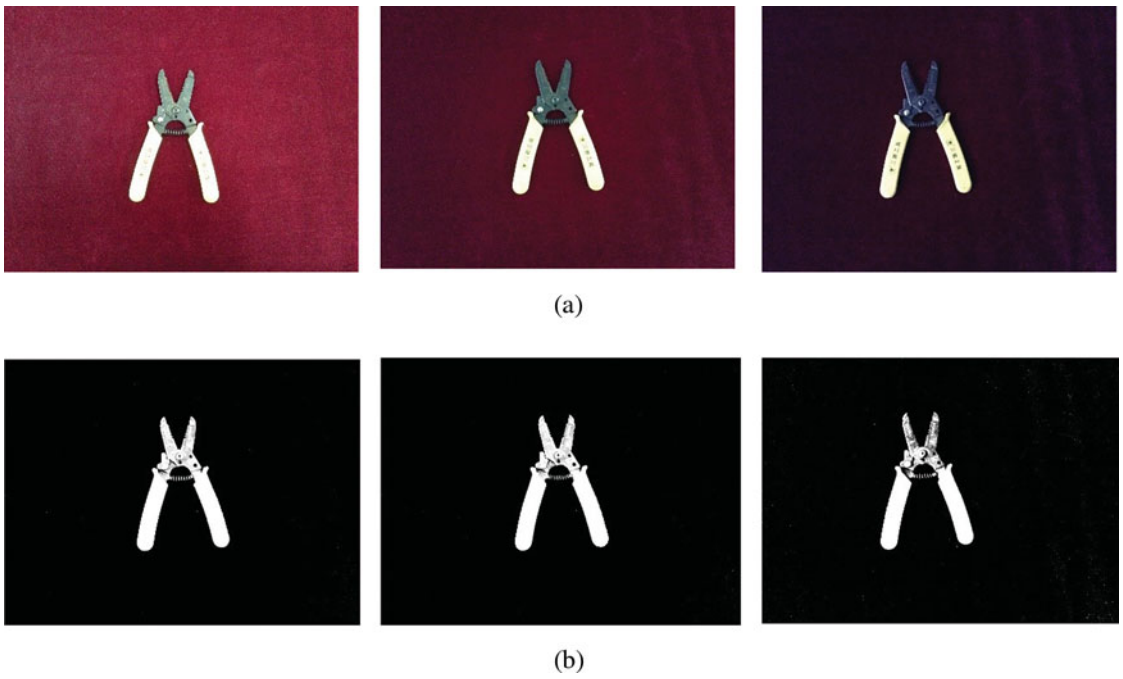


Fig. 8. Images under different illumination.

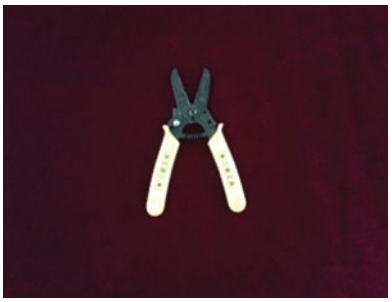


Fig. 9. Sample image.

by repeating the capture process. Furthermore, the frame speed was obtained too, which reached 28 frames per second.

5.2. Target recognition results under affine transformation

In this section, to test the effect of a change in target position, a series of transformations were set up. The position of the target, including horizontal translation, rotation, and scaling, was changed under uniform, general illumination condition. Fig. 9 shows the sample image.

Fig. 10 shows the captured images under affine transformation. Fig. 10 (a) and (A) shows the translation images, with a 100 mm and 200 mm translation, respectively. Fig. 10 (b) and (B) shows the images of rotation. Fig. 10 (c) and (C) shows the images at different scales.

Table VIII shows the values of affine moment invariants and distances. For this experiment, all transformed images were correctly matched with the models from the database, using affine moment invariants except the enlarged 1.2 times one as the value of $d/|\mathbf{I}|$ was approximately equal 0.25. The analysis showed that it might occur for the effect of uncertain noise. All other targets were matched correctly. In total, 54 images were captured from the various viewpoints and the six different affine transformations. An accuracy rate of 96.3% was obtained during the object recognition process. At the same time, we obtained the frame speed with 30 frames per second.

Due to the limitation of the vision range, the target with a horizontal translation 200 mm has reached the edge of the image, i.e., the target can be accurately recognized within the scope

https://doi.org/10.1017/S0263574714001970 Published online by Cambridge University Press

Table VIII. Affine moment invariants and distances under different affine transformation.

Items	S^1	S^2	S^3	d	$d/\|\mathbf{I}\|$
Sample	4.2922	−12.198	−10.439	2.9767	0.1637
a	4.2426	−12.182	−9.9912	2.8690	0.1577
b	3.9913	−12.011	−8.7654	3.0581	0.1681
c	4.3273	−12.169	−11.541	2.4503	0.1347
A	4.4599	−14.827	−11.081	1.6933	0.0931
B	4.4860	−15.504	−11.303	1.9754	0.1086
C	4.8839	−16.289	−13.575	4.7684	0.2425

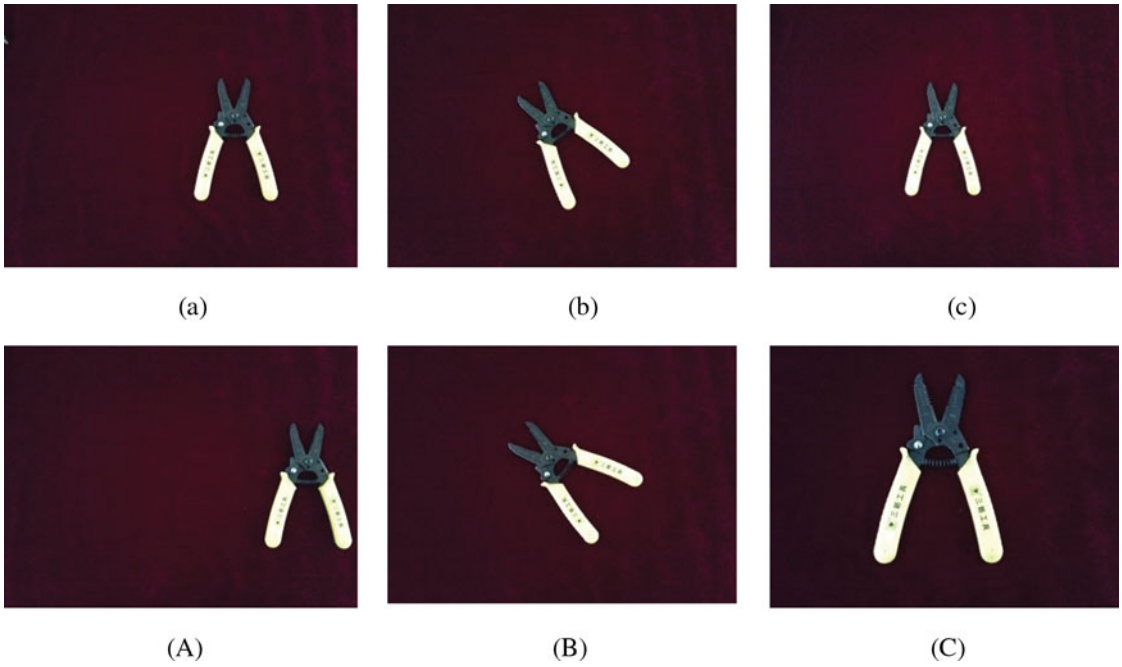


Fig. 10. Images under affine transformation.

of 0 200 mm translation. To obtain the rotation and scale change on the influence of target recognition, we captured images at different rotation angles and scales, and calculated the values of $d/|\mathbf{I}|$.

Fig. 11 shows the result curves. Fig. 11 (a) shows that target can be accurately recognized within the scope of 360-degree rotation. Fig. 11 (b) shows that target can be accurately recognized within the scope of 0.6 1.2 times scaling.

5.3. Target recognition results under complex conditions

The third experiment addressed recognition in complex conditions under general illumination. The camera captured occlusion images and non-occlusion images (Fig. 12). Fig. 12 (a) and (b) show the four targets in each image; the yellow pair of pliers was the designated object for recognition. The images were segmented into four separate targets based on the histogram and threshold which introduced in section IV-B specifically, and then the affine moment invariants were extracted.

Table IX shows the values of the affine moment invariants of each target and also the distances between the targets and the targeted object in the feature space. The values of $d/|\mathbf{I}|$ show that there were only two correctly matched targets: A-1 and B-1. In total, 160 occlusion and non-occlusion images of the pliers were captured under different complex conditions and viewpoints. The accuracy rate for non-occlusion images was 93.1% and for occlusion images was 86.9%. With the increase of occlusion factors, the frame speed declined slightly with the values of 25 fps and 22fps under non-occlusion condition and occlusion condition respectively.

https://doi.org/10.1017/S0263574714001970 Published online by Cambridge University Press

Table IX. Affine moment invariants and distances under complex conditions.

NO.	S^1	S^2	S^3	d	$d/\|\mathbf{I}\|$
A-1	4.5848	−16.669	−11.71	2.8672	0.1576
A-2	4.6884	12.042	−17.939	28.349	1.5586
A-3	6.4214	−12.654	−29.025	19.869	1.0924
A-4	4.8141	20.886	−7.8669	35.919	1.9748
B-1	4.4005	−13.9086	−12.008	2.8126	0.1546
B-2	5.7872	11.782	−19.226	28.560	1.5702
B-3	5.3073	−13.9896	−25.066	15.716	0.8641
B-4	4.6204	10.437	0.11756	27.166	1.4936

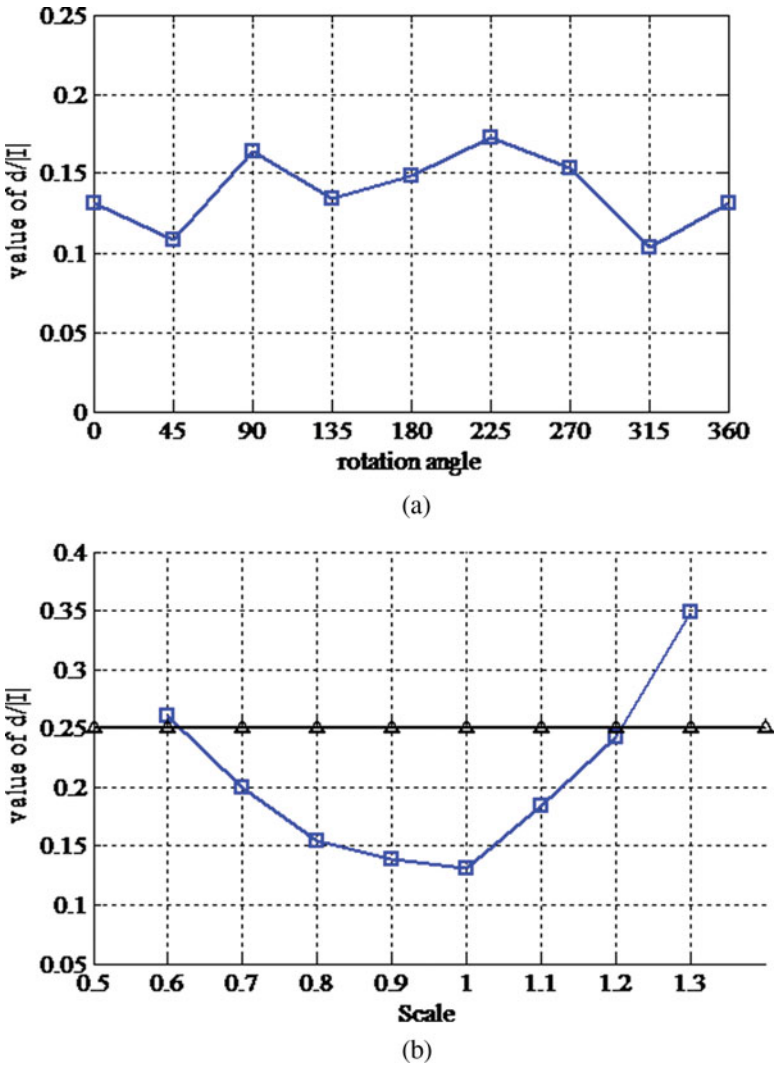


Fig. 11. Result curves under affine transformation (a) Value of d/|I| curve under different rotation angles (b) Value of d/|I| curve under different scales.

5.4. Target recognition results under comprehensive factors

To test our method for images taken from targets in complex scenes and varying illumination, we captured a set of targets images under different illumination, affine transformation and occlusion shown in Fig. 13 (a) and (b). Fig. 13 (A) and (B) show transformations of the original images using the contrast function, taking into account the illumination conditions. By calculating the average gray value of each test image, we got the value of parameter c which was shown in Fig. 13.

https://doi.org/10.1017/S0263574714001970 Published online by Cambridge University Press

Table X. Affine moment invariants and distances under comprehensive factors.

NO.	S^1	S^2	S^3	d	$d/\|\mathbf{I}\|$
A-1	4.7195	−13.384	−10.463	1.9986	0.1099
A-2	6.04856	16.759	−12.519	31.959	1.7571
A-3	5.5657	−12.748	−16.59	7.6412	0.4201
A-4	3.9437	−0.1236	−0.7367	17.223	0.9469
B-1	7.2524	10.394	−14.519	26.079	1.4338
B-2	3.0545	−17.911	−12.208	4.1772	0.2297
B-3	5.4331	2.4352	−15.217	18.410	1.0122
B-4	6.4990	6.6705	0.88712	24.108	1.3255

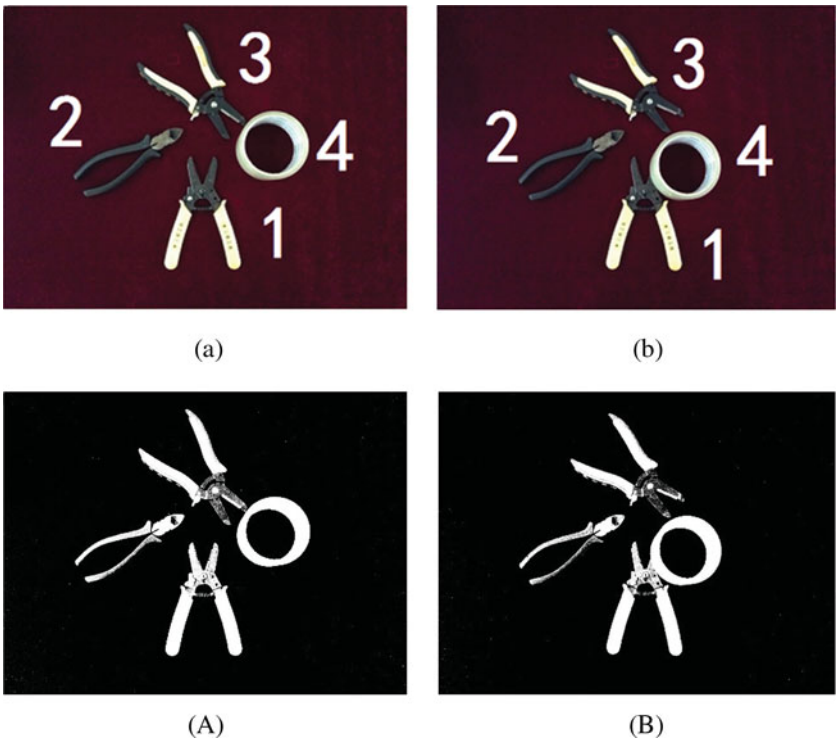


Fig. 12. Images under complex conditions.

Table X shows the values of affine moment invariants for each target and also the distances. The values of $d/|\mathbf{I}|$ show that only two targets matched correctly: A-1 and B-2. The results showed that the recognition method was quiet accurate.

An additional 80 images were captured under comprehensive factors. An accuracy rate of 88.8% was obtained by repeating the capture process. Moreover, we got the average frame speed was 20 frames per second.

Table XI shows the accuracy rate and frame speed under different conditions based on the above research and experiments and the results of studies only based on color feature or color & simple shape features. The results of these experiments demonstrated that object recognition by affine moment invariants based on gray level transformation was sufficiently accurate. Furthermore, the establishment of an image feature library based on affine moment invariants decreased the calculation time for the vector stored in the library only contained three features and the similarity measurement algorithm was calculated quickly. This method can solve the interference of different illumination, affine transformation, and occlusion under complex conditions. Although the speed of operation declined slightly, it was still within acceptable parameters.

https://doi.org/10.1017/S0263574714001970 Published online by Cambridge University Press

Table XI. Target recognition results and processing time comparison of different factors.

Items	Accuracy Rate (%)	Frame Speed (frame·s^{-1})
Color feature	50.0	30
Color & simple shape features	80.0	25
Illumination	95.6	28
Affine Transformation	96.3	30
Complex Conditions with Non-Occlusion	93.1	25
Complex Conditions with Occlusion	86.9	22
Comprehensive Factors	88.8	20

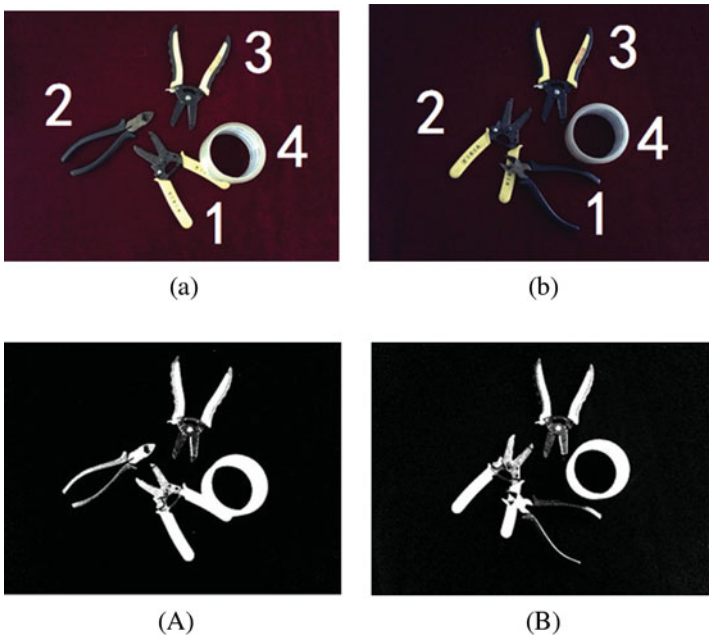


Fig. 13. Images under comprehensive factors.

6. Conclusion

This paper proposed a novel method to significantly increase the accuracy rate for robot object recognition in the complex environment of a space station. This method made the following contributions.

(1) A new robot visual target recognition method based on gray level transformation, affine moment invariant extraction and feature matching. This method can solve the problem of target recognition in complex space conditions.
(2) A new image-capture method that used nine viewpoints and calculated the average moment invariant of each viewpoint image. Also, the establishment of a feature library that included image path tables, basic information tables and image feature data tables. This enabled the robot to quickly and accurately obtain target recognition from any viewpoint.
(3) The construction of an experimental robot visual target recognition system. This system was used to conduct a series of experiments under the special conditions of a space station, including different illumination, affine transformation, occlusion and non-occlusion, and comprehensive factors. The results showed that this method is effective and feasible.

The robotic visual system will be studied further in order to address the actual space station environment in which recognition would occur. A viable system would also need to demonstrate effectiveness in coping with uncertainties different to those presented in this paper, such as the recognition of moving targets.

https://doi.org/10.1017/S0263574714001970 Published online by Cambridge University Press

The algorithm used in this study only focused on robotic visual recognition because the aim was to propose a method for enhancing the speed and accuracy rate of a space robot visual system. Future study should focus on the visual navigation of a space robot, including system optimization, enhanced human and robot safety, and energy consumption.

Acknowledgements
The authors wish to express their gratitude to the National High Technology Research of China (Grant 2011AA040202), the Beijing Science Foundation (Grant 4122065) and the National Natural Science Foundation of China (Grant 60925014 and 61273348) for their support of this work.

References

1. W. De-han and H. Wei-fen, "Astronaut task analysis," *Aerospace China* **7**, 37–40 (1997).
2. W. Zheng, W. Li and XU Run-tao, "Study on operational reliability of astronaut in spacecraft cabin," *Space Med. Med. Eng.* **19**(5), 358–362 (2006).
3. C. Jin-dun, "Safety analysis for astronaut and the personal protective equipment," *Space Med. Med. Eng.* **12**(6), 418–422 (1999).
4. Z. Qian-xiang, "The research of human reliability in manned spaceflight," *Aerospace Shanghai* **4**, 26–29 (2001).
5. Lin Xiao, http://www.spacechina.com/n25/n148/n272/n4793/c104843/content.html
6. Z. Qian-xiang, "Lighting design ergonomics research in manned space mission," *Chinese J. Illuminating Eng.* **17**(2), 1–5 (2006).
7. B. V. Funt and G. D. Finlayson, "Color constant color indexing," *IEEE Trans. Pattern Anal. Mach. Intell.* **17**(5), 522–529 (1995).
8. M. J. Swain and D H. Ballad, "Color indexing," *Int. J. Comput. Vis.* **7**(1), 11–32 (1991).
9. S. K. Nayar and R. M. Bolle, "Reflectance based object recognition," *Int. J. Comput. Vis.* **17**(3), 219–240 (1996).
10. T. Gevers and A. W. M. Smeulders, "Color based object recognition," *Pattern Recognition* **32**(3), 453–464 (1999).
11. G. Murat and K. Bahadir , "Illumination robust interest point detection," *Comput. Vis. Image Understanding* **113**, 565–571 (2009).
12. M. K. Hu, "Visual pattern recognition by moment invariants," *IER Trans. Inform. Theory* **8**(2), 179–187 (1962).
13. T. H. Reiss, "The revised fundamental theorem of moment invariants," *IEEE Trans. Pattern Anal. Mach. Intell.* **13**(8), 830–834 (1991).
14. J. Flusser and T. Suk, "Pattern recognition by sffine moment invariants," *Pattern Recognition* **26**(1), 167–174 (1993).
15. J. Flusser and T. Suk, "Pattern recognition by means of affine moment invariants," Res. Rep. 1726, Institute of Information Theory and Automation, 1991.
16. Z. Yang and F. S. Cohen, "Cross- weighted moments and affine invariants for image registration and matching," *IEEE Trans. Pattern Anal. Mach. Intell.* **21**(8), 804–814 (1999).
17. T. Suk and J. Flusser, "Combined blur and affine moments invariant and their use in pattern recognition," *Pattern Recognition* **36**, 2895–2907 (2003).
18. J. Flusser, T. Suk and B. Zitova, *Moments and moment invariants in pattern recognition* (Wiley, Chichester, 2009).
19. T. Suk and J. Flusser, "Affine moment invariants generated by graph method," *Pattern Recognition* **44**(9), 2047–2056 (2011).
20. D. G. Lowe, "Object recognition from local scale invariant features," **In:** *Proceedings of the International Conference on Computer Vision* Corfu Greece (1999) pp. 1150–1157.
21. D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.* **60**(2), 91–110 (2004).
22. J. M. Morel and G. Yu, "ASIFT: A new framework for fully affine invariant image comparison," *SIAM J. Imaging Sci.* **2**(2), 1–31 (2009).
23. I. Schur, *Vorlesungen Uber Invariantentheorie* (Springer, Berlin 1968).

https://doi.org/10.1017/S0263574714001970 Published online by Cambridge University Press