J Intell Robot Syst (2018) 92:265–277
https://doi.org/10.1007/s10846-017-0712-5

Camera Recognition and Laser Detection based on EKF-SLAM in the Autonomous Navigation of Humanoid Robot

Shuhuan Wen[1] · Miao Sheng[1] · Chunli Ma[1] · Zhen Li[1] · H. K. Lam[2] · Yongsheng Zhao[3] · Jingrong Ma[1]

Received: 24 June 2015 / Accepted: 22 August 2017/ Published online: 3 November 2017

Abstract The ability of autonomous navigation of the humanoid robot under unknown environment is very important to real-life applications. EKF-SLAM based on the camera recognition and laser detection for humanoid robot NAO is presented in this paper. Camera recognition is used to recognize if the object is a landmark. Because the computational resources needed for the feature-based position estimation are quite expensive, the laser instead of the camera provides the position of the landmark. A fractional order proportional-integral (PI) controller is designed to reduce the derivation of the NAO robot from the desired path during autonomous navigation. Experiments show that the proposed method is valid and reliable for autonomous navigation of the NAO robot under unknown environment.

Keywords EKF-SLAM algorithm · Object recognition · Fractional order PI controller · Autonomous navigation · Laser detection

✉ H. K. Lam
hak-keung.lam@kcl.ac.uk

Shuhuan Wen
swen@ysu.edu.cn

Miao Sheng
1511194744@qq.com

Chunli Ma
mcl_neuq@sina.com

Zhen Li
zhenli_ysu@126.com

Yongsheng Zhao
yszhao@ysu.edu.cn

Jingrong Ma
748866261@qq.com

[1] Key Lab of Industrial Computer Control Engineering of Hebei Province, Yanshan University, Qinhuangdao, 066004, China

[2] Department of Informatics, King's College London, Strand, London, WC2R 2LS, UK

[3] Parallel Robot and Mechatronic System Laboratory of Hebei Province and Key Laboratory of Advanced Forging & Stamping Technology and Science of Ministry of National Education, Yanshan University, QinHuangDao, 066004, China

1 introduction

With the progress of science and technology, service robots at home have been widely used for people. Widespread attention has been paid on the humanoid robot as one of the most active research direction of the robot research areas. The study focused on the fields of a complex environment to complete some human tasks, just to name a few, such as human assistance, home care and transportation [1]. So it is necessary to study the robot to walk from one point to another independently, which is an essential component for basic action of robots.

Object recognition is a basic research in the field of machine vision, and it plays an important role in the realization of the target image understanding. In the process of moving object recognition and localization of mobile robot, the target position in the image, scale and viewpoint and the background images will change with the vision sensor of the robot. It increases the difficulty for the robot to identify and locate moving target [2]. Target recognition based on color features, invariance of rotation, translation, scaling is suitable for the target with bright color. However, when the color of an object is similar to the target in the environment, it will lead to the failure of target recognition [3]. So target

recognition based on color features in this paper is used to judge the contour of the object. Thus the object recognition is more accurate and reliable [4]. In this paper, the NAO robot will identify the target object with every step.

SLAM can estimate the structure of the surrounding world (the map), by using moving exteroceptive sensors, while it can simultaneously locate [5]. SLAM is derived from the Bayesian framework, and the consistency issue is a fundamental problem. EKF-SLAM can produce inconsistency in the linearization process [6]. EKF-SLAM uses the extended Kalman filter (EKF), which is developed from Kalman filter (KF). As the robot moves, this map keeps updating by the EKF. There are two critical steps, prediction (robot motion model) and correction (the robot sensor detects the landmarks that had been mapped before) [7].

In this paper, the robot using camera recognizes if the object which the robot sees is a landmark. Because the computational resources needed for the feature-based position estimation provided by the camera are quite expensive and the method cannot be run during the entire journey of the robot, a laser scanner is used to detect the position of the landmarks [8, 9]. The information of landmark does not need to be stored in the NAO robot before experiment. The method is introduced in section III(A). The parameters will be modified if the landmark changes. If the object is a landmark, the laser module is called to give the position of the robot. The proposed method in this paper is applied on a humanoid NAO robot [10]. Camera is used to capture landmark for robot navigation based on EKF-SLAM. The laser sensor is used to improve the accuracy and simplify the computation of EKF-SLAM algorithm. The fractional order controller is also added to reduce the deviation when the NAO robot walks [11]. In this paper, the NAO robot recognizes the predefined graphical markers through its camera, and obtains the landmark information.

In our early work in [12], the object recognition was completed by hand drawing contour through the third party API in choregraph. Then the NAO robot calls Vision Reco. module to complete the recognition. In this paper, the camera of NAO robot does not need to use the third party API to complete the object recognition. NAO robot can automatically obtain the image and recognize the object. Aldebaran Robotics company develops a graphical programming platform-Choregraphe for NAO robot, and also develops many functional models for this platform, such as audio, communication , data edit, motions, system, trackers, vision etc. In ref. [12], the authors used the vision model to realize the identification function of NAO robot. Because of strict requirement of object shape, low rate of identification, drawing the outline of the object in advance etc, this method is extremely limited. In this paper, the method we proposed could make the robot identify the object automatically, we just need to pre-establish the parameters of object in Python code. In our early work in [12], only laser was used to provide the position of the landmark. As long as the object appears in the detection range of the laser, the laser will give the position information of object no matter the object is landmark or not. This will result in over computation and wasting time. In this paper, the NAO robot provides the position information only when it recognizes the landmark. Thus the proposed method can reduce the computational time and make the robot walk autonomously.

The organization of this paper is as follows. Section 2 introduces the basic knowledge used in this paper, including the NAO robot, object recognition in the visual database of OpenCV, laser based EKF-SLAM algorithm and the fractional PI controller design. Section 3 describes the process of research. The results of simulation and experiment are showed in Section 4. Section 5 draws the conclusion.

2 basics Knowledge Introduction

2.1 NAO Humanoid Robot

NAO humanoid robot, as shown in Fig. 1, has a total of 25 degrees of freedom, flexible movement. It is also equipped with an inertial navigation unit (a gyroscope and an accelerometer). The use of some action engines and gait planning algorithms can make the NAO walk steadily, even realize some highly difficult movements [13].

NAO has two identical cameras, the angle between the top and bottom center of the camera is 40 degrees, the maximum available 30 frames / s, with a resolution of 640 × 480 images. The camera can be used to capture the far and near, and near (about 50cm) can only be used under the camera. Some parameters of the camera are shown in Table 1 [14]. Although NAO has two cameras, only one camera is active at any instance where the settings of the parameters can be changed through the camera API. There is one laser scanner on the robot NAO head. The type of the laser scanner used in our experiment is URG-04LX. The function of the laser is provided in Table 2 [15].

2.2 EKF-SLAM

EKF-SLAM uses mobile probability distribution model and observation model to obtain the state vector and covariance matrix. The input and output of EKF-SLAM are shown in Fig. 2 [17].

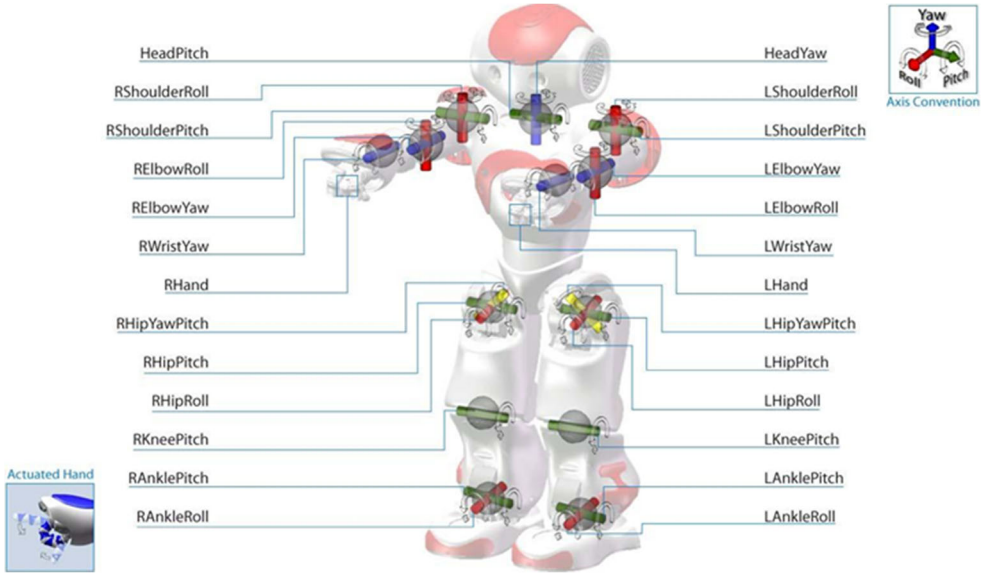


Fig. 1 NAO humanoid robot [16]

The state matrix **X** predicts the position of the robot and the state vector of the landmark. Its dimension is $3 + 2N$ where N denotes the number of signs [18–20].

$$\mathbf{X} = \begin{bmatrix} \mathbf{R} \\ \mathbf{M} \end{bmatrix}_{(3+2N)\times 1} \quad \mathbf{R} = \begin{bmatrix} x_r \\ y_r \\ \theta \end{bmatrix} \quad \mathbf{M} = \begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ L_N \end{bmatrix}_{2N\times 1} \tag{1}$$

The covariance **P** is a square matrix representing the uncertainty of system or deviation. Equation 2 is the covariance matrix, including many sub-variance matrices.

Where $\mathbf{P}_{RR}$ is the robot variance, $\mathbf{P}_{MM}$ is the signs variance, $\mathbf{P}_{RM}$ is the associated variance of the robot and road signs, $\mathbf{P}_{MR}$ is the transpose of $\mathbf{P}_{RM}$.

$$\begin{aligned} \mathbf{P} &= \begin{bmatrix} [\mathbf{P}_{RR}]_{3\times 3} & [\mathbf{P}_{RM}]_{3\times 2N} \\ [\mathbf{P}_{MR}]_{2N\times 3} & [\mathbf{P}_{MM}]_{2N\times 2N} \end{bmatrix} \\ &= \begin{bmatrix} [\mathbf{P_{RR}}]_{3\times 3} & \left[\mathbf{P_{Rl_1}}\right]_{3\times 2} \cdots \left[\mathbf{P_{Rl_N}}\right]_{3\times 2} \\ \left[\mathbf{P_{l_1R}}\right]_{2\times 3} & \left[\mathbf{P_{l_1l_1}}\right]_{2\times 2} \cdots \left[\mathbf{P_{l_1l_N}}\right]_{2\times 2} \\ \vdots & \vdots \quad \ddots \quad \vdots \\ \left[\mathbf{P_{l_NR}}\right]_{2\times 3} & \left[\mathbf{P_{l_Nl_1}}\right]_{2\times 2} \cdots \left[\mathbf{P_{l_Nl_N}}\right]_{2\times 2} \end{bmatrix} \end{aligned} \tag{2}$$

Table 1 camera parameters

Original image format	YUV422
Horizontal view angle	58 degree
Vertical view angle	34.8 degree
Focus range	30cm to infinity

EKF is a kind of minimum variance estimation method for nonlinear system equations. It can express system equations of linear system and estimate the state vector of nonlinear system. With a given initial state values, the covariance matrix and the observed value at time t, recursive the computation of state vector of t moments.

2.3 Fractional PI Controller

When the NAO moves, it may be interfered by sensor noise, environmental noise and other factors. It can lead to deviation of the target trajectory. In order to minimize the deviation, a fractional order PI controller is designed to compensate the deviation [12]. In ref. [12], simulation results show that the fractional controller based on EKF-SLAM can reduce the position deviation during the NAO walking under the unknown indoor environment. In this paper, the controller design is cited in ref. [12]. The key information is as follows.

This paper uses the fractional order PI (PI^{α}) as controller, whose transfer function is:

$$G(s) = \frac{U(s)}{E(s)} = K_p + \frac{K_i}{s^{\alpha}} = \frac{K_p s^{\alpha} + K_i}{s^{\alpha}}. \tag{3}$$

Table 2 Main characteristics of the NAO URG-04LX laser

Function	Data
Detection range	0.02–4m
Scan angle	240 degree
Scan time	100msec/scan(10.0Hz)
Resolution	1mm
Interface	USB 2.0, RS232

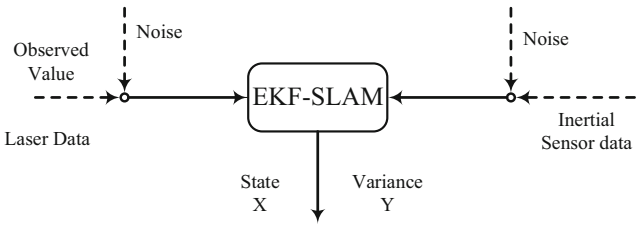


Fig. 2 The observed model and mobile model in EKF-SLAM connection

Al-Alaoui operator discretized, Eq. 3 can be written as:

$$G(z) = \frac{U(z)}{E(z)} = \frac{K_p\left(\frac{8}{7T}\right)^{\alpha}\frac{P_p}{Q_q} + K_i}{\left(\frac{8}{7T}\right)^{\alpha}\frac{P_p}{Q_q}} = \frac{K_p\left(\frac{8}{7T}\right)^{\alpha}P_p + K_i Q_q}{\left(\frac{8}{7T}\right)^{\alpha}P_p} \tag{4}$$

where, $P_p = (1 - z^{-1})^{\alpha}$, $Q_q = (1 + z^{-1}/7)^{\alpha}$. Equation 4 is changed through the three order PSE [21] expansion:

$$G(z) = \frac{U(z)}{E(z)} = \frac{K_p K^{\alpha}(p_3 z^{-3} + p_2 z^{-2} + p_1 z^{-1} + 1) + K_i(q_3 z^{-3} + q_2 z^{-2} + q_1 z^{-1} + 1)}{K^{\alpha}(p_3 z^{-3} + p_2 z^{-2} + p_1 z^{-1} + 1)} \tag{5}$$

where, $p_1 = -\alpha$, $p_2 = \frac{\alpha^2-\alpha}{2}$, $p_3 = \frac{-\alpha^3+3\alpha^2-2\alpha}{6}$, $q_1 = \frac{\alpha}{7}$, $q_2 = \frac{\alpha^2-\alpha}{98}$, $q_3 = \frac{\alpha^3-3\alpha^2+2\alpha}{2058}$, $K^{\alpha} = \left(\frac{8}{7T}\right)^{\alpha}$
So Eq. 5 can be expressed as

$$G(z) = \frac{(K_p K^{\alpha} p_3 + K_i q_3)z^{-3} + (K_p K^{\alpha} p_2 + K_i q_2)z^{-2} + (K_p K^{\alpha} p_1 + K_i q_1)z^{-1} + (K_i + K_p K^{\alpha})}{K^{\alpha}(p_3 z^{-3} + p_2 z^{-2} + p_1 z^{-1} + 1)}. \tag{6}$$

Reshuffling the terms in Eq. 7, we have

$$\begin{aligned} &U(z)K^{\alpha}(p_3 z^{-3} + p_2 z^{-2} + p_1 z^{-1} + 1) \\ &\quad = E(z)[(K_p K^{\alpha} p_3 + K_i q_3)z^{-3} + (K_p K^{\alpha} p_2 + K_i q_2)z^{-2} + (K_p K^{\alpha} p_1 + K_i q_1)z^{-1} + (K_i + K_p K^{\alpha})]. \end{aligned} \tag{7}$$

We define the control quantity in the time domain as $u(t)$ based on Eq. 7, we can get u(t) as follows:

$$\begin{aligned} u(t) \;=\; & (K_p p_3 + K_i q_3 K^{\alpha})e(t-3) + (K_p p_2 + K_i q_2 K^{\alpha})e(t-2) + (K_p p_1 + K_i q_1 K^{\alpha})e(t-1) \\ & + (K_p + K_i K^{\alpha})e(t) - p_1 u(t-1) - p_2 u(t-2) - p_3 u(t-3) \end{aligned} \tag{8}$$

The difference between the real position and estimation position is denoted as $e(t)$ which is employed as the input of controller. $u(t)$ is the output of controller which is employed to correct the walking process. Based on the results from a large number of experiments, when $K_p = 0.9$, $K_i = 0.2$, $\alpha = 0.1$, $T = 0.1$, the best result is achieved, i.e., $e(t)$ is minimized.

3 object Recognition and Landmarks Location

3.1 Object Recognition

To identify landmarks or obstacles, visual detection based on the camera of NAO robot is used. Firstly, semi-autonomous learning is used in order to access to the target threshold in the detection process. It can improve the light adaptability of visual processing, and perform object segmentation by using global threshold method [22] in BGR color space. Secondly, the pixel is traversed to extract target feature points. Finally, the outline of the object is extracted, and the contour perimeter, area, and other parameters are calculated. Thus the measured object is determined if it is a target object. It is a valid method to determine the target object in our study.

We use color histogram and color moment to extract the color feature of the target object [23, 24]. Color histogram describes the global distribution of the color of the image. The color histogram calculation divides the color space into several short intervals. Each interval is a bin histogram. This process is called color quantization. Then, the color histogram is obtained by counting the number of pixels in each

short interval. Color quantization uses vector quantization method. The calculation formula is as follows:

$$H(k) = \frac{n_k}{N}(k = 0, 1, \ldots, L-1) \tag{9}$$

where, k represents the characteristic value of image, L is the number of possible characteristic values, n_k is the number of image pixels with k characteristic value, N is the total number of pixels.

Different color space will generate different color histogram. The RGB space is the most common method. The color histogram could be obtained by statistics from R, G and B channels. It could be expressed by three one dimensional color histograms or one three dimensional color histogram.

Any distribution of color could be expressed by its moments. Because the distribution information of color converges on the low order moments, mean (μ_i), variance (σ_i) and skewness (s_i) of color are enough to express the distribution of image color. They are expressed as the following equations.

$$\begin{aligned}\mu_i &= \frac{1}{N}\sum_{j=1}^{N} p_{ij}, \sigma_i = \left(\frac{1}{N}\sum_{j=1}^{N}(p_{ij}-\mu_i)^2\right)^{\frac{1}{2}}, s_i \\ &= \left(\frac{1}{N}\sum_{j=1}^{N}(p_{ij}-\mu_i)^3\right)^{\frac{1}{3}}\end{aligned} \tag{10}$$

where p_{ij} is the i^{th} color component of j^{th} pixel in the image.

The color moments of the image need three color components and three low order moments of each color component. Compared with other expression of color distribution, the color moment is more succinct. In fact, the color moments is always combined with other image characters. This can avoid the disadvantages that the ability of judgment of low order moments is weak. In general, the color moments is firstly used to narrow down before using other image characters.

Images in Fig. 3 are obtained by using the NAO robot's camera and converted to numpy format.

Table 3 Function of OpenCV library

Library	Function
cv2.pyrMeanShiftFiltering	Image segmentation
cv2.cvtColor	RGB to HSV format of the image
cv2.split	Separate color channels
cv2.findContours	Find the contour of object in the image
cv2.moments	Calculate the moment of the contour

Though computer vision library based on image processing – OpenCV, three channel images are B, G, R instead of the usual R, G, B. So it is necessary to make conversion. The split function in the OpenCV library can separate image channels. The channel values are stored in R, G, B format, then they use merge function to merge channels. Thus the converted image is obtained. In the first image preprocessing, image segmentation using cv2.pyrMeanShiftFiltering (The function as show in Table 3) is extracted conveniently. The result of image segmentation is shown in Fig. 4.

Then image in Fig. 4 is converted into HSV format by using cv2.cvtColor. HSV (Hue, Saturation, Value) based on a color space creates the visual characteristics of color, which is known as hexagonal pyramid model. The model parameters are the color: hue (H), saturation (S), value (V). Here RGB format will be converted into HSV format, which is designed to reduce the optical effect of light on target recognition. The separation channel through the cv2.split method, the hue, saturation and brightness of the image information are stored in the H, S, V, but the standard H range should be 0-360. H is in the range of standard half 0-180 in the OpenCV HSV format. We set the threshold of saturation, hue, brightness three properties respectively, then binary image. The color of the recognized object in this paper is green. The green color values are in the range of 35-70, which satisfy the original pixels hue threshold value. The threshold value is set to 255, while the other position is set to 0. The above work is the pretreatment of the main

Fig. 3 The picture taken by the camera of NAO robot

Fig. 4 The results of image segmentation

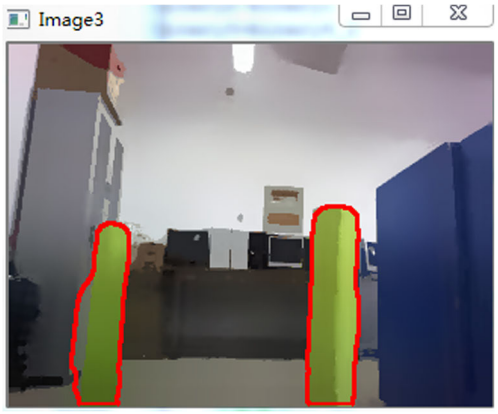


Fig. 5 The results of recognition of objects

working objects recognition. According to the original saturation and brightness control to distinguish the background and the object, other colors background interference are excluded, and finally the image median filtering can reduce noise in the image point.

Due to environments may also contain other similar colored objects to the target object, target object shape information is regarded as the primary way to be recognized. Firstly, the preprocessed binarized image can find all of the contour in the figure by using the cv2.findContours function. According to the contour of the area and perimeter, the ratio of area and perimeter square can exclude part of the background profile. Secondly, the contour of the moment is calculated by cv2.moments function. 'nu20' and 'nu02' which are the normalized values in the x and y contour direction of second-order central moments are used to determine the contour of target object. Because the contour of the object is approximately rectangular and set upright, 'nu02' is greater than 'nu20'. Finally, the center position of the object in the image is calculated, which is shown in eq.(11) and eq.(12).

$$x = c[' m10']/c[' m00'] \tag{11}$$

$$y = c[' m01']/c[' m00'] \tag{12}$$

where c ['m10'] is the first moment for the contour in x direction, c ['m01'] is the first moment for the contour in y direction, c ['m00'] is the area of contour. By using this method to recognize the object, the conditions are more stringent, then judgment is more accurate. Figure 5 shows the result of recognition of objects.

3.2 Location of Landmarks

NAO humanoid robot can get the location of objects by its own sensors, such as laser, sonar or camera. Sonar can only measure the distance between the robot and the object. Positioning error provided by camera is large. However, laser can provide accurate position of target object. The detection area of laser and its coordinate axis are shown in Fig. 6. There is error when NAO robot walks. The error will increase with accumulation. The robot walks five steps which will produce negligible error. So the robot NAO in the experiment stops in every five steps to retrieve the position of the target object, and the new location information will be send back to the robot. When the object is recognized as target object by the camera, the position information of the target object is given by laser. The laser information includes x-y coordinates, the distance and angle values of the target object. K-means clustering method [25] is used to extract useful readings from laser and exclude invalid value, such as zero.

The data is stored in a two dimensional 684 × 4 array. The data includes distance between laser and objects, angle, x and y in Cartesian space. The steps are shown as follows [26].

Step 1: Assign each observation to the cluster whose mean is closest to it.

$$S_i = \{x_p : \| x_p - m_i \| \leq \| x_p - m_j \|, \forall 1 \leq j \leq k \tag{13}$$

where x_p is a set of observations $\{x_1 \cdots x_n\}$ and assigned to exactly one S_i, m_i is an initial set of k means $\{m_1 \cdots m_k\}$.

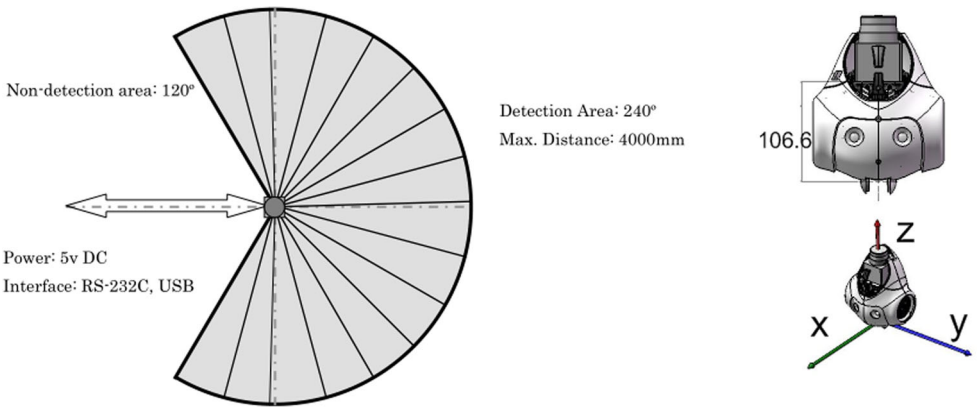


Fig. 6 The characteristics of laser

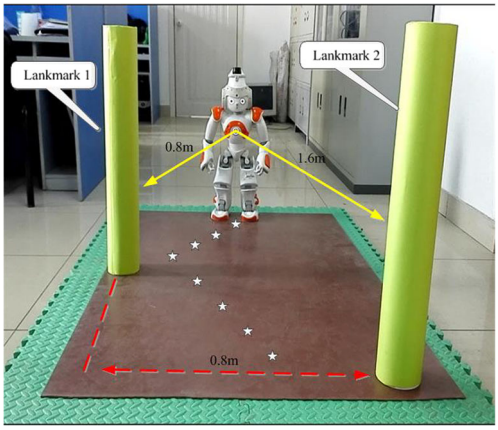


Fig. 7 The position information of NAO robot and landmarks

Step 2: Update the new means as the centroids of the observations in the new clusters.

$$m_i = \frac{1}{|S_i|} \sum_{x_j \in S_i} x_j \tag{14}$$

The algorithm stops when all m_i do not change any more.

4 experiment Results

The work in [12, 27] used laser to detect the location information of landmarks. But laser cannot judge whether the object is the target object (that is landmark) in its detection range. The laser data is from all the objects in its detection range. This is the limitation in the work [12]. In this paper, camera and laser are combined to be used in SLAM. Camera is used to recognize if the object is a landmark. If the object is a landmark, then laser data is called. If it is not a landmark, NAO robot will not call laser to detect the location information of landmark. This method saves the computational resource of NAO robot and time. Fractional PI controller is used in this paper to decrease the error in the walking of NAO robot. The work in [27] used PI controller to decrease the error. However, PI controller is not better than Fractional PI controller [12].

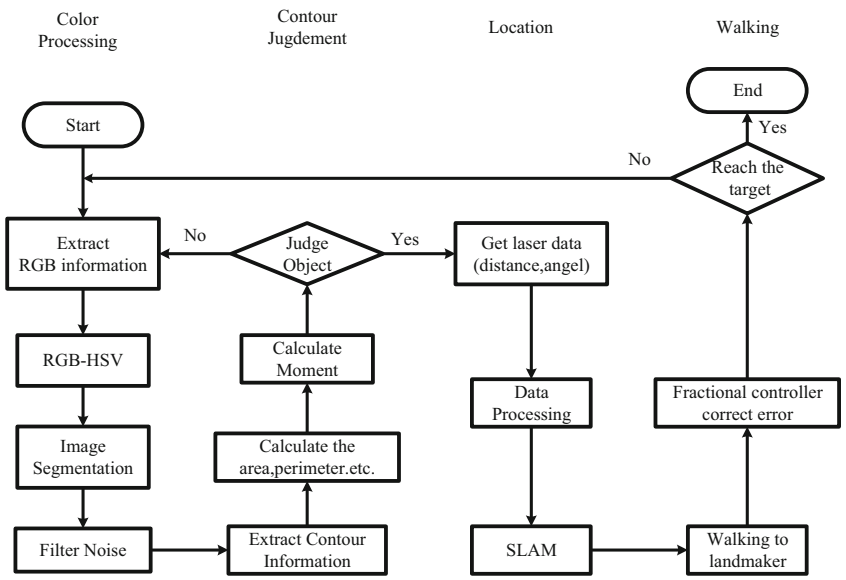
4.1 The Scene of Two Landmarks

We consider two kind of experiment scenes. The first scene is that two landmarks which need to be recognized. The second scene is about one landmark and one non-landmark. The position information of NAO robot and landmarks is shown in Fig. 7. The distance between the NAO robot and the landmark on its left side is about 0.8m. The distance between another landmark from the NAO robot is about 1.6m. We set the safety distance between NAO robot and landmark to be 0.3m. The experiment step is shown in Fig. 8. The NAO robot recognizes the target object by using camera before walking. When it finds two landmarks, then the NAO robot calls the laser data of both landmarks. If the transversal distance between the two landmarks is greater than 0.6m, the robot can walk straight in the middle of both landmarks. NAO robot will stop walking if the distance between NAO robot and landmark is less than 0.3m. The experimental scene is shown in Fig. 9.

The length range of laser sensor is (200, 1000)mm, angle range is $(-\pi/3, \pi/3)$. The initial position of NAO robot is $[x \approx 0, y \approx 0, \theta \approx 0^r]$. The recognition process of the experiment is shown in Fig. 10.

The NAO robot recognizes the landmarks and plans the path by the recognized landmarks. Figure 10 shows the recognition process when the NAO robot walks. Figure 10a is the recognition result before the NAO robot begins to walk.

Fig. 8 The experiment step of two landmarks



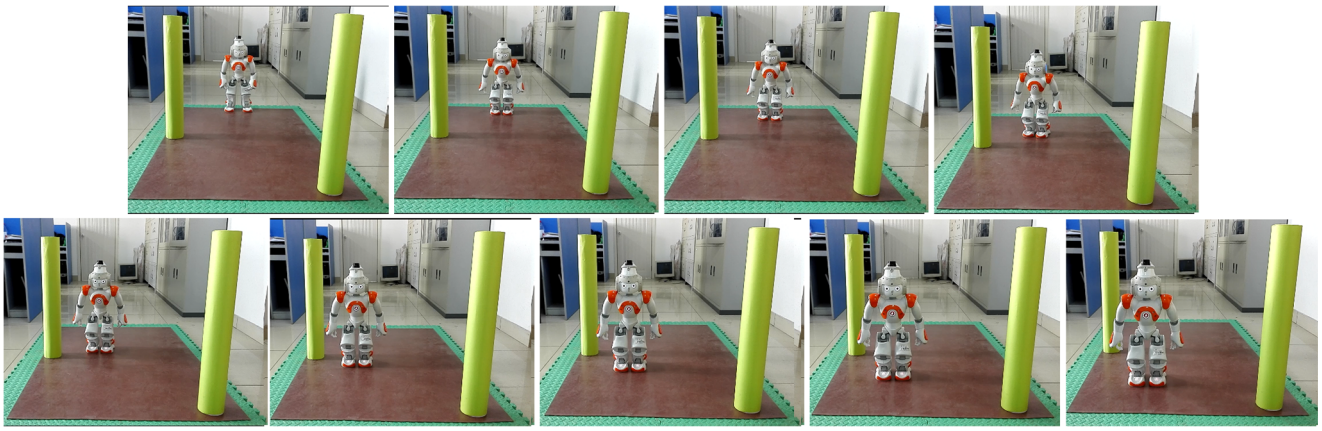
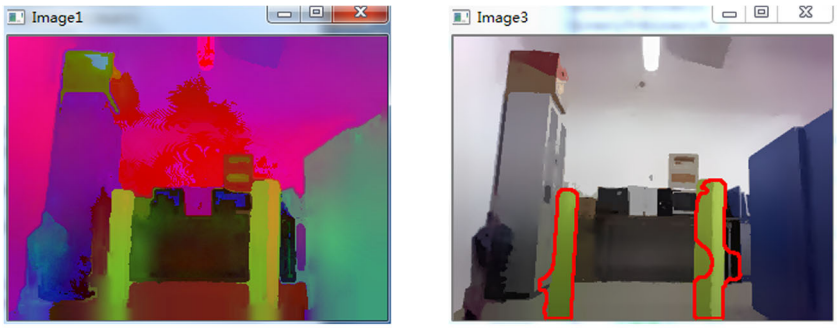
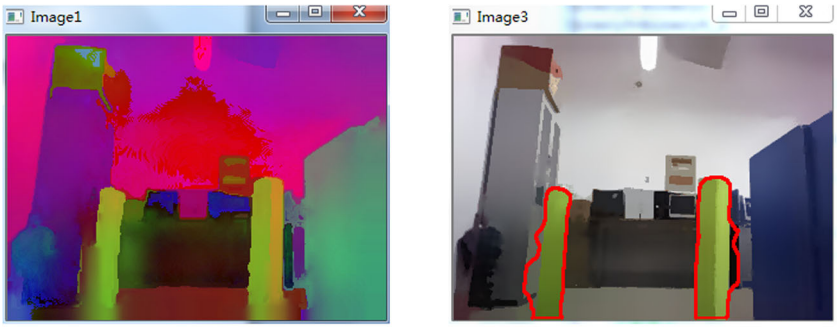


Fig. 9 The experiment scene of two landmarks

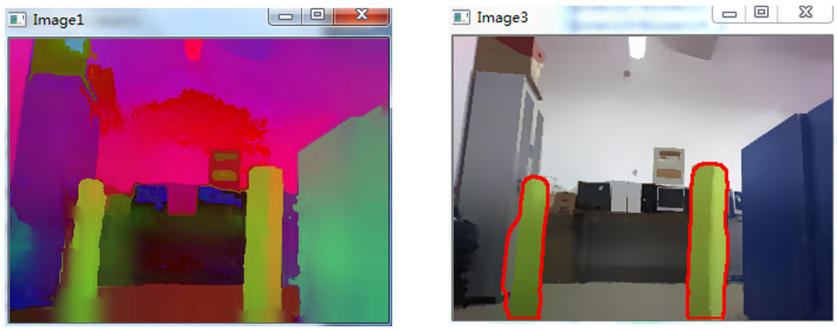
Fig. 10 The recognition process of NAO robot in walking (two landmarks)



(a) The recognition results for the first time



(b) The recognition results for the second time



(c) The recognition results for the third time

Table 4 Partial data of two landmarks detected by laser

Landmark 1				Landmark 2			
Distance(m)	Angle(rad)	X value(m)	Y value(m)	Distance(m)	Angle(rad)	X value(m)	Y value(m)
1.053	1.457	0.119	1.046	0.695	-0.344	0.654	-0.234
1.040	1.458	0.116	1.033	0.696	-0.339	0.656	-0.232
1.024	1.454	0.119	1.017	0.694	-0.327	0.657	-0.223
1.010	1.453	0.118	1.003	0.693	-0.316	0.659	-0.215
0.995	1.450	0.119	0.988	0.690	-0.307	0.658	-0.209
0.978	1.444	0.123	0.970	0.689	-0.297	0.659	-0.202
0.965	1.445	0.120	0.957	0.700	-0.289	0.671	-0.200
0.952	1.445	0.119	0.945	0.721	-0.285	0.692	-0.203
0.937	1.440	0.122	0.929	0.728	-0.280	0.700	-0.201
.....	.....	.....	.....	.....	.....	.....	.....

The left figure in Fig. 10a is the processing results of the hue (H), saturation (S), brightness (V) treatment of objects which the NAO robot views. The right figure in Fig. 10a is the contour of extracted landmarks. Figure 10b and c are the second and third recognition results respectively. These results show that the proposed recognition method in this paper is feasible and valid when NAO robot walks.

The unit of data obtained by laser is millimeter. In order to make the calculation more easily, we transform the unit into meter. Laser will create values and store them in the array by scanning the object. Firstly, NAO robot transposes this array and filters the invalid values. Then the obtained data is divided into several parts by clustering method according to the amount of objects. Lastly, the mean value of distance and angel data detected by laser are calculated and transformed into coordinate values. Table 4 shows a part of data obtained by laser.

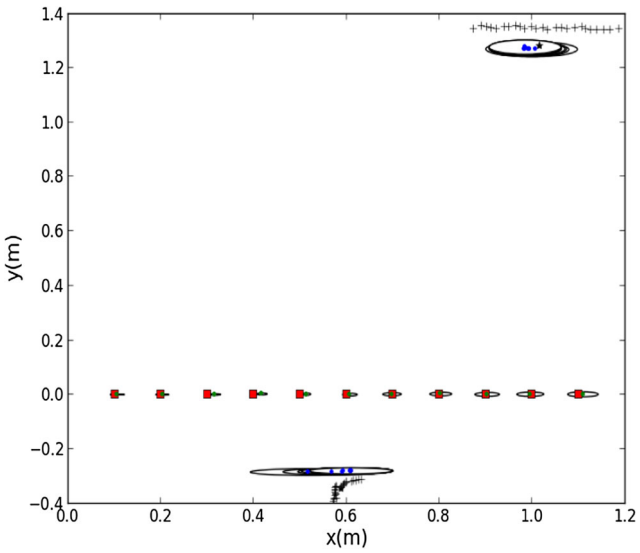


Fig. 11 The NAO robot walking process under two landmarks

Figure 11 is the result of the walking trajectory of NAO robot. The hollow elliptical dots represent the walking trajectory of the NAO. The hollow ovals represent the covariance matrix. The range constituted by the black '+' is the signs observed by laser sensor. The star represents the actual position of landmarks. The dots around landmarks mean the estimated location by NAO robot.

4.2 The Scene of One Landmark and One non-Landmark

In order to verify that the recognition process is only valid for the specified landmark, we do an experiment that the position information of NAO robot and landmark and non-landmark are shown in Fig. 12. The experiment step is shown in Fig. 13. The green object is the recognized landmark. The

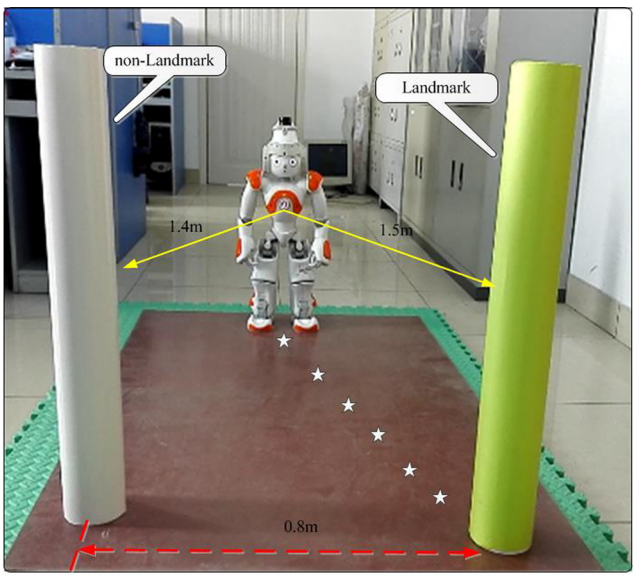


Fig. 12 The position information of NAO robot and landmark and non-landmark

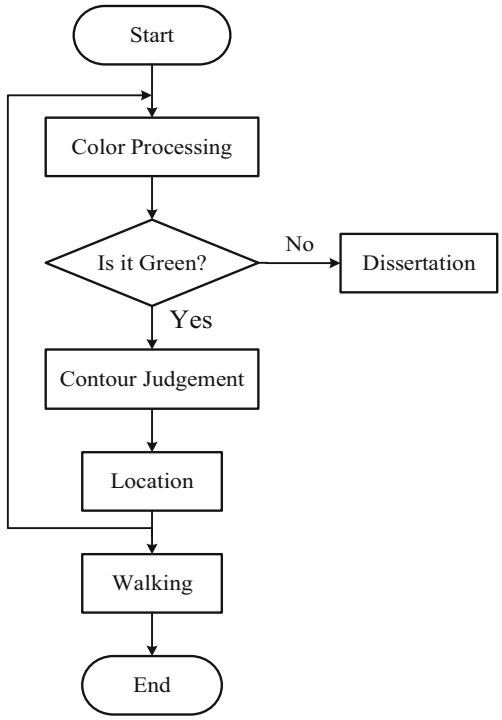


Fig. 13 The experiment step of landmark and non-landmark

white object is not the recognized landmark. In this paper, we do not study the obstacle avoidance. So the NAO robot can keep a safe distance from the white object and stop at a certain distance during the walking process. The experiment scene is shown in Fig. 14.

Figure 15 shows the recognition processing of the NAO robot under one landmark and one non-landmark in the experiment. We can see that the NAO robot can identify the landmark accurately. Figure 15 shows the process of recognition. Figure 15abc are the first, second and third recognition results.

When one landmark and one non-landmark coexist in the same environment, the NAO robot can call position information of the landmark by laser when it meets and recognizes the landmark. That means the NAO robot cannot call position information of the landmark by laser when it sees other non-landmark. Because laser can detect all the position information of the object, we need only extract the position information of landmark by clustering method. Then the NAO robot can walk toward the landmarks. The part of data detected by laser in our experiment is shown in Table 5.

Figure 16 shows the walking trajectory of NAO robot under the experiment environment of one landmark and one non-landmark. We can see that the trajectory of the NAO robot is good. It shows the proposed method in this paper is valid.

In [12, 26], laser is used to detect the landmarks location based on the EKF-SLAM algorithm. But the NAO robot is unable to judge if the object is a landmark in the range of laser detection. So the NAO robot needs to regard all objects in the range of laser detection as landmarks. The limitation is to reduce the range of laser detection in order to prevent non-landmarks from emerging in the range of laser detection. This is not appropriate for the real environment. In this paper, camera object recognition is used to judge if the object is a landmark before laser detects the location of the landmark. If the object is a landmark, then the NAO robot calls laser to provide the location information of landmark. Otherwise, the NAO robot will not call laser. This method improves the real-time of the EKF-SLAM algorithm.

Figure 11 shows the NAO robot walking trajectory under two landmarks in the experiment. In Figs. 11 and 16, the hollow elliptical dots represent the walking trajectory of NAO robot. The red rectangle represents the reference walking trajectory of NAO robot. The hollow ovals represent the covariance matrix. The black '+' means the range observed by laser sensor. The star means the actual position of landmarks. The dots around landmarks mean the estimated location by NAO robot. The actual trajectory and the reference trajectory are almost coincident, which demonstrates

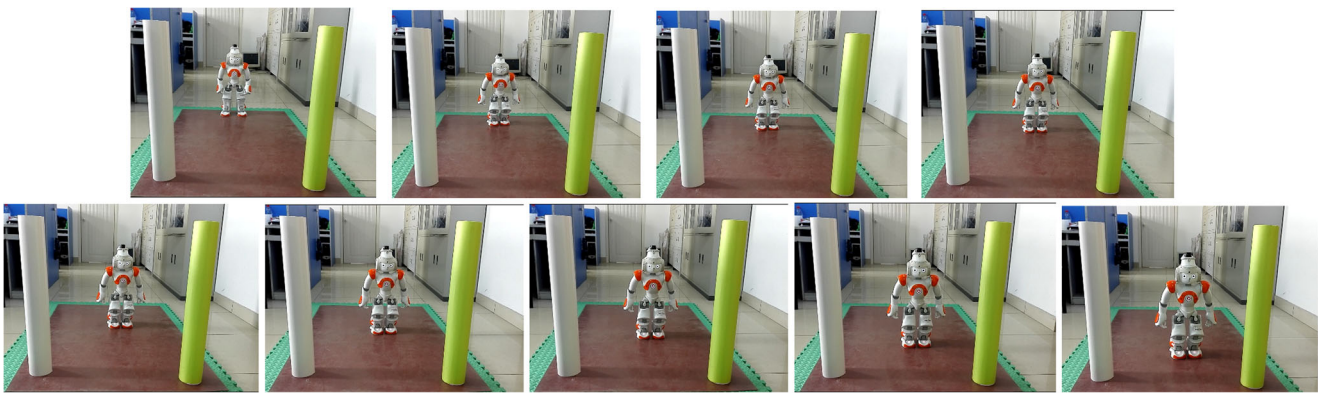
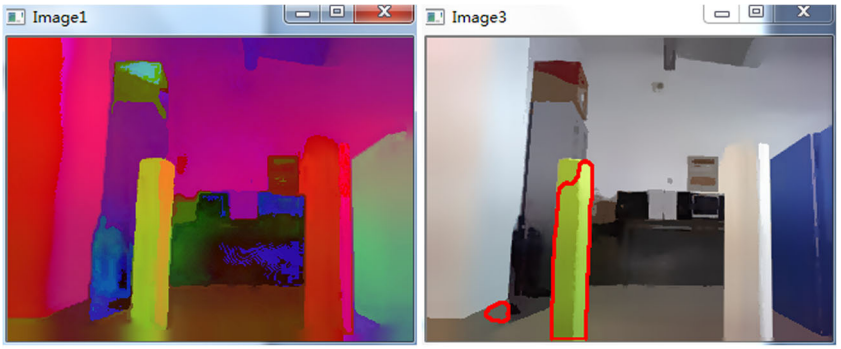
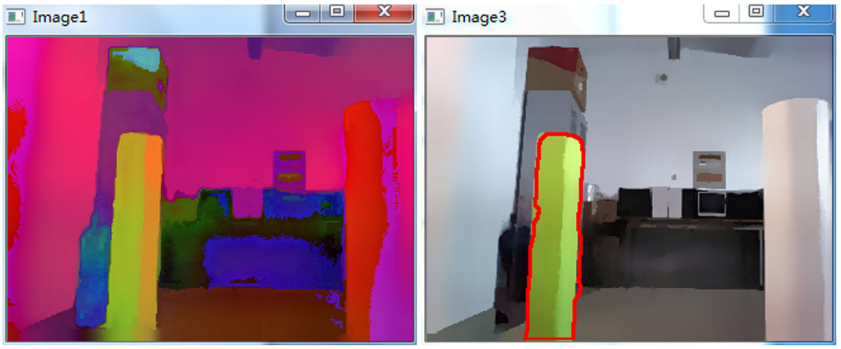


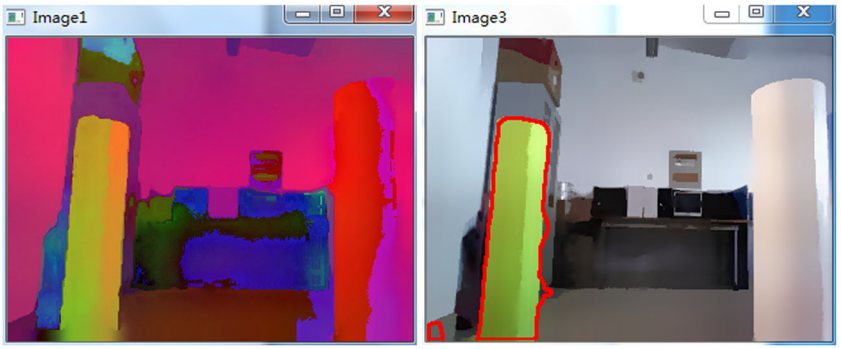
Fig. 14 The experiment scene of landmark and non-landmark



(a) The recognition results for the first time



(b) The recognition results for the second time



(c) The recognition results for the third time

Fig. 15 The recognition process of NAO robot in walking (landmark and non-landmark)

Table 5 Partial data of landmark and non-landmark detected by laser

Landmark				Non-Landmark			
Distance(m)	Angle(rad)	X value(m)	Y value(m)	Distance(m)	Angle(rad)	X value(m)	Y value(m)
0.905	-0.428	0.823	-0.376	1.077	0.312	1.025	0.331
0.908	-0.422	0.828	-0.372	1.067	0.317	1.014	0.333
0.910	-0.417	0.832	-0.369	1.057	0.321	1.003	0.334
0.913	-0.411	0.837	-0.365	1.044	0.321	0.991	0.330
0.916	-0.406	0.841	-0.362	0.959	0.327	0.908	0.308
0.882	-0.384	0.818	-0.331	0.944	0.324	0.895	0.301
0.884	-0.379	0.821	-0.327	0.934	0.326	0.885	0.299
0.886	-0.373	0.825	-0.323	0.920	0.324	0.872	0.293
0.892	-0.363	0.834	-0.317	0.910	0.327	0.862	0.292
.....	.....	.....	.....	.....	.....	.....	.....

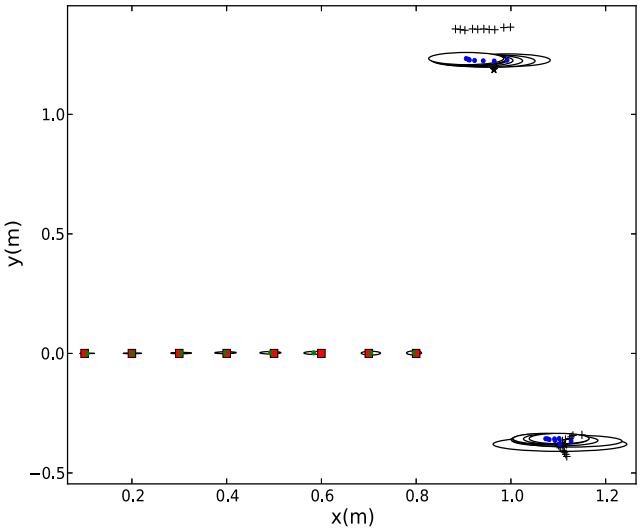


Fig. 16 The NAO robot walking process under one landmark and one non-landmark

that the walking error of NAO robot reduced greatly by using fractional PI controller. Figure 15 demonstrates that the NAO robot can judge the landmark and non-landmark correctly. Figure 16 shows that laser can also detect the location information of non-landmark.

The NAO robot running time is listed in Table 6. In this paper, the running time means the interval time including from sending instructions by computer to receiving instructions, executing instructions and returning the results to the computer by the NAO robot.

5 Conclusion and Future Work

In this paper, EKF-SLAM based on the camera recognition and laser detection was developed. EKF-SLAM algorithm is used under indoor environment simultaneous localization and map building. Fractional controller is used to reduce the error during the walking process of the NAO robot. Camera and laser are combined to recognize the landmarks and provide the position information of the landmarks. The experiment results show that the proposed method is feasible, and can finish walking of the NAO robot autonomously under indoor environment.

This paper only used two landmarks. We could study more complex situation, such as more landmarks. Better recognition method will be used to improve the result of object recognition. In the real environment, the limitation of this paper is that we do not improve the EKF-SLAM algorithm. Future work is to improve the EKF-SLAM algorithm.

Table 6 Two scene Running time

	The scene of one landmark and one non-landmark	The scene of two landmarks
The time of every 5 steps	3s	3s
The time of recognizing object	3s	3s
The time of the whole procedure	45s	First landmarks 26s Second landmarks 56s Total time: 82

Acknowledgments The work was partly supported by the National Natural Science Foundation of China (Project No. 61773333), and Scholars Studying Abroad Science and Technology Activities of Hebei Province of China (Project No. C201400355), the major project of Science and technology in Hebei Universities (Project No. ZD2016150).

References

1. Hirose, N., Tajima, R., Sukigara, K., Minoru, T.: Personal robot assisting transportation to support active human life-reference generation based on model predictive control for robust quick turning. In: 2014 IEEE International Conference on Robotics and Automation (ICRA), pp. 2223–2230 (2014)
2. Cinque, L., Olsen, K.A., Levialdi, S., et al.: Color-based image retrieval using spatial chromatic histograms. Image Vision Comput **19**(13), 979–986 (2001)
3. Chen, T.-C., Chung, K.-L.: An efficient randomized algorithm for detecting circles. Comput Vis Image Understand **83**(2), 172–191 (2001)
4. Chen, H.H., Su, J.S.: A syntactic approach to shape recognition. Proc Intl Computer Symp, 103–122 (1986)
5. Zhang, X., Rad, A.B., Wong, Y.-K.: Sensor fusion of monocular cameras and laser rangefinders for line-based simultaneous localization and mapping (slam) tasks in autonomous mobile robots. Sensors **12**(1), 429–452 (2012)
6. Huang, S., Dissanayake, G.: Convergence and consistency analysis for extended kalman filter based slam. IEEE Trans. Robot. **23**, 1036–1049 (2007)
7. Sola, J.: Simulataneous localization and mapping with the extended kalman filter, unpublished. Available: http://www. joansola.eu/JoanSola/eng/JoanSola.html (2013)
8. Mohareri, O., Rad, A.B.: A visioin-based location positioning system via augmented reality: an application in humanoid robot navigation. Intern J Humanoid Robot **3**, 1–26 (2013)
9. Yamashita, A., Kitaoka, M., Kaneko, T.: Motion planning of biped robot equipped with stereo camera using grid map. Interna-Tional J Autom Technol **5**(5), 639–647 (2011)

10. Moutarlier, P., Chatila, R.: An experimental system for incremental environment modelling by an autonomous mobile robot, Presented at Experimental Robotics I (1990)
11. Chen, Z., Samarabandu, J., Rodrigo, R.: Recent advances in simultaneous localization and map-building using computer vision. Adv. Rob. **21**(3-4), 233–265 (2007)
12. Wen, S., Chen, X., Zhao, Y., Rad, A.B., Othman, K.M., Zhang, E.: The study of fractional order controller with SLAM in the humanoid robot. Advances in Mathematical Physics (2014)
13. http://www.aldebaran-robotics.com/documentation/index.html. [NAO Software 1.14 Documentation]
14. http://www.ros.org/wiki/nao. [Robot Operating System]
15. Maier, D., Bennewitz, M., Stachniss, C.: Self-supervised obstacle detection for humanoid navigation using monocular vision and sparse laser data in robotics and automation (ICRA), IEEE International Conference, pp. 1263–1269 (2011)
16. NAO Datasheet, 2013. Aldebaran Robotics
17. Bailey, T., Nieto, J., Guivant, J., Stevens, M., Nebot, E.: Consistency of the EKF-SLAM algorithm. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (2006)
18. Bosse, M., Newman, P., Leonard, J., Soika, M., Feiten, W., Teller, S.: An atlas framework for scalable mapping. In: IEEE International Conference on Robotics and Automation, pp. 1899–1906 (2003)
19. Chong, S., Kleeman, L.: Feature-based mapping in real,large scale environments using an ultrasonic array. Int J Robot Res **18**(1), 3–19 (1999)
20. Sola, J.: Simulataneous Localization and Mapping with the Extended Kalman Filter, Avery quick guide with MATLAB code (2013)
21. Wing, J.K., Cooper, J.E., Sartorius, N.: Measurement and classification of psychiatric symptoms: an instruction manual for the PSE and Catego Program. Cambridge University Press, Cambridge (2012)
22. Shao-ling, D., Shu-guang, D.: Research on the color detection of LCD Instrument panel based on Machine Vision. In: 2010 International Symposium on Information Science and Engineering (ISISE), pp. 75–77. IEEE (2010)
23. Swain, M.J., Ballard, D.H.: Color indexing. Int J Comput Vis **7**(1), 11–32 (1991)
24. Stricker, M., Orengo, M.: Similarity of color images. SPIE Storage Retr Image Video Databases III **2185**, 381–392 (1995)
25. Kanungo, T., Mount, D.M., Netanyahu, N.S., Piatko, C.D., Silverman, R., Wu, A.Y.: An efficient k-means clustering algorithm: Analysis and implementation. IEEE Trans Pattern Anal Mach Intell **24**(7), 881–892 (2002)
26. Wen, S., Othman, K.M., Rad, A.B., Zhang, Y., Zhao, Y.: Indoor SLAM using laser and camera with closed-loop controller for NAO humanoid robot. In: Abstract and Applied Analysis. Hindawi Publishing Corporation, Hindawi (2014)
27. Othman, K.M.: Implementation of EKF-SLAM on NAO humanoid robot. PhD thesis, Applied Sciences: School of Mechatronic Systems Engineering (2013)

Shuhuan Wen was born in Heilongjiang, China, on July 16, 1972. She received the PhD degree in control theory and control engineering from the Yanshan University, Qinhuangdao, China in 2005. She is currently a Professor of automatic control in the Department of Electric Engineering, Yanshan University. She has coauthored one book, about 40 papers. Her research interests include humanoid robot control, force/motion control of parallel robot, Fuzzy control, 3-D object recognization and reconstruction. Dr. Wen was a Visiting Scholar of the Ottawa University, Carleton University and Simon Fraser University in Canada from 2011 to 2013.

Miao Sheng was born in Hubei, China, on February 02, 1994. He received the Bachelor degree in the Department of Physics and Electric Engineering, Hubei University of Arts and Sciences in 2016. His research interests are focused on humanoid robot vision system.

Chunli Ma was born in ShanXi, China, in May, 1989. She received the Bachelor degree in the Department of Control Engineering, Northeastern University at Qinhuangdao in 2013. She has coauthored about one journal paper. Her research interests are focused on humanoid robot control, the simultaneous locatization and mapping (SLAM), 3-D object recognization and reconstruction.

Zhen Li was born in HeNan, China, in December, 1991. He received the Bachelor degree in the Department of Electric Engineering, PingDingShan University in 2014. He has coauthored about one journal papers. His research interests are focused on humanoid robot control and SLAM algorithm.

H. K. Lam received the B.Eng. (Hons.) and Ph.D. degrees from the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong, in 1995 and 2000, respectively. During the period of 2000 and 2005, he worked with the Department of Electronic and Information Engineering at The Hong Kong Polytechnic University as Post-Doctoral Fellow and Research Fellow respectively. He joined as a Lecturer at King's College London in 2005 and currently is a Reader. His current research interests include intelligent control and computational intelligence. He has served as a program committee member and international advisory board member for various international conferences and a reviewer for various books, international journals and international conferences. He is an associate editor for IEEE Transactions on Fuzzy Systems, IET Control Theory and Applications, International Journal of Fuzzy Systems and Neorocomputing; and guest editor for a number of international journals. He is in the editorial board of Journal of Intelligent Learning Systems and Applications, Journal of Applied Mathematics, Mathematical Problems in Engineering, Modelling and Simulation in Engineering, Annual Review of Chaos Theory, Bifurcations and Dynamical Systems and The Open Cybernetics and Systemics Journal. He is an IEEE senior member. He is the coeditor for two edited volumes: Control of Chaotic Nonlinear Circuits (World Scientific, 2009) and Computational Intelligence and Its Applications (World Scientific, 2012), and the coauthor of the monograph: Stability Analysis of Fuzzy-Model-Based Control Systems (Springer, 2011).

Yongsheng Zhao received his bachelor's degree and the master's degree in mechanical engineering from Northeast Heavy Machinery Institute, Qiqihaer, Heilongjiang Province, China in 1983 and 1987, respectively. He received the PhD degree in mechanical engineering from Yanshan University, Qinhuangdao, Hebei Province, China in 1999. He is currently a professor in Robotics Research Center at Yanshan University. He is currently the vice-chancellor of Yanshan University. His research interests include parallel robot, force sensor, numerical control technique, Fuzzy control, etc. He has coauthored one book, more than 80 papers.

Jingrong Ma was born in Hebei, China, in December, 1991. She received the Bachelor degree in the Department of Control Engineering, Hebei United University in 2014. She has coauthored about one journal paper. Her research interests are focused on humanoid robot control and 3-D object recognization.