Smooth Vision-Based Autonomous Land Vehicle Navigation in Indoor Environments by Person Following Using Sequential Pattern Recognition*

Ching-Heng Ku, Wen-Hsiang Tsai†
Department of Computer and Information Science
National Chiao Tung University
Hsinchu, Taiwan 300
Republic of China

Received November 7, 1997; revised October 17, 1998;
accepted December 5, 1998

A new approach to autonomous land vehicle (ALV) navigation by the person following is proposed. This approach is based on sequential pattern recognition and computer vision techniques, and maintenance of smoothness for indoor navigation is the main goal. The ALV is guided automatically to follow a person who walks in front of the vehicle. The vehicle can be used as an autonomous handcart, go-cart, buffet car, golf cart, weeder, etc. in various applications. Sequential pattern recognition is used to design a classifier for making decisions about whether the person in front of the vehicle is walking straight or is too right or too left of the vehicle. Multiple images in a sequence are used as input to the system. Computer vision techniques are used to detect and locate the person in front of the vehicle. By sequential pattern recognition, the relation between the location of the person and that of the vehicle is categorized into three classes. Corresponding adjustments of the direction of the vehicle are computed to achieve smooth navigation. The approach is implemented on a real ALV, and successful and smooth navigation sessions confirm the feasibility of the approach.

*This work was supported by the National Science Council under Grant NSC86-2213-E009-114.
†To whom all correspondence should be addressed.

CCC 0741-2223/99/050249-14

1. INTRODUCTION

Many methods[1–8] have been proposed for ALV navigation in indoor environments. These methods depend on special features in environments, such as lines,[1,2] corners,[3,4] models,[5,6] or special landmarks.[7,8] When an ALV navigates in different indoor environments, these features may not always be available. In such cases, the ALV has to change its navigation strategy to suit new environments. In this study, a navigation method by a person following is proposed, by which a person can lead the ALV to any place the person wants to reach, and this creates many possible applications of the ALV system. The system can be used as an autonomous handcart, go-cart, buffet cart, shopping cart, dust cart, golf cart, weeder, etc. Besides, the system can also be used as a learning system to collect information in certain environments, including open paths or locations of obstacles. On the other hand, because the indoor environment is usually complicated and small, the navigation methods of the road following,[9,10] vehicle following,[11,12] or car following[13–15] proposed for outdoor environments are usually unsuitable to indoor navigation. Navigation by the person following instead is more convenient in such cases.

In the area of human body tracking, some methods have been proposed, such as stereo and motion measurements,[16] multiclass statistical modeling of color and shape,[17] and the so-called Pfinder (''person finder'') and Spfinder (''stereo person finder'').[18] These methods are not useful for autonomous navigation because the detection of an entire human body shape spends a lot of time and is not necessary for real-time navigation. In this study, a special shape attached on the back of the person is used for fast detection and tracking of the person to shorten the processing time.

Besides, some systems[19–21] for following a person have been proposed. Matsuda et al.[19] proposed a tracking control system for an autonomous mobile robot with multiple sensors in which supersonic distance sensors and infrared direction sensors are configured. Hanebeck and Schmidt[20] proposed a high performance multisonar system which can be used for tracking a walking person. These two systems do not use the information of the visual sensor. Mori and Sano[21] proposed a guide dog robot which follows a moving pedestrian by tracking the jacket or trousers image of the person. This system does not use the information of the relation between consecutive images.

In this study, a new approach to the person following for autonomous land vehicle navigation using image sequences is proposed. The major goal is to achieve smooth navigation by sequential pattern recognition techniques. To guide the ALV to follow the person who walks in front of the vehicle, the location of the person is detected from input images first. The relation between the location of the person and that of the vehicle is then categorized into three classes. The first class is that the location of the person is too right or too left to the vehicle. The second is that both the person and the vehicle move in a straight trajectory. In addition, the last class is that the location of the person does not belong to the above two classes. In the sequential pattern recognition process, two features are proposed for use in a quadratic classifier. According to the classification result, corresponding adjustments of the direction of the vehicle are proposed. The proposed method can solve the problem of the drift location due to small differences between the vehicle heading and the heading of the person. The direction of the vehicle is not adjusted in the case of drift location when the location of the person is close to the trajectory of the vehicle. The vehicle keeps its direction until the location of the person is too right or too left of the vehicle. The adjustment of the ALV speed is implemented using the fuzzy technique.[22] It is hoped that the ALV could navigate smoothly when the person walks in different directions, so that the turning angle of the front wheels of the ALV need not be adjusted immediately in every cycle. Furthermore, due to the use of a rectangular shape on the back of the person, the computing time for the detection of the location of the person is shortened and the walking direction of the person is easily detected. In short, the main contribution of this study is the use of sequential pattern recognition and the detection of the simple shape to achieve fast navigation with smooth trajectories.

A flowchart of the person following system is shown in Figure 1. The navigation system is composed of a vision system, a decision system, and a control system. In the vision system, input images are captured by a color CCD camera that is mounted on the vehicle. Additionally, image processing and computer vision techniques are employed to detect feature points from the image of the person who walks in front of the vehicle and calculate the three-dimensional coordinates of the feature points. The location and speed of the person are computed accordingly and used in the control system. In the decision system, the sequential pattern recognition

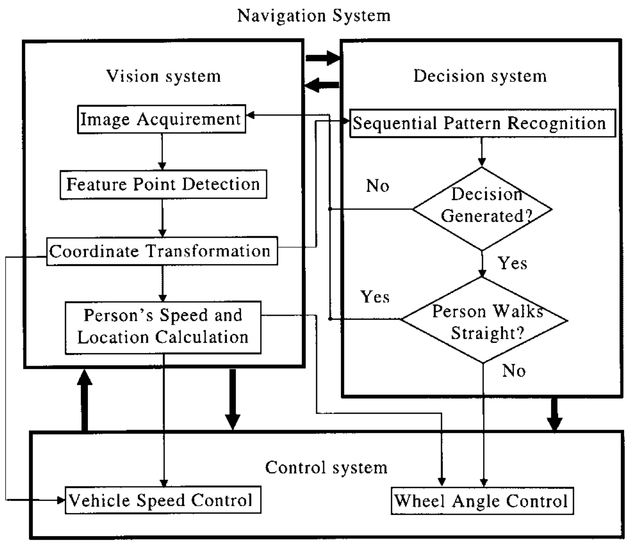

Figure 1. Flowchart of the proposed system.

technique is used to decide whether the person walks straight or is too right or too left of the vehicle. The control system is triggered by the decision system. The three-dimensional coordinates of the location of the person are used in the control system to control the turning angle of the front wheels of the ALV. Besides, the speed of the person and the distance between the person and the vehicle are used to adjust the speed of the vehicle by a fuzzy control technique.[22] A more detailed description of the navigation process is described as follows.

Step 1. Image acquirement: Acquire an image of the front view of the vehicle, which in normal cases, includes the person who walks in front of the vehicle.

Step 2. Feature point detection: Detect some feature points of the rectangular shape attached on the back of the person by a region growing technique using the color information.

Step 3. Coordinate transformation: Transform the image coordinates of the feature points obtained in Step 2 into the 3D space coordinates to calculate the speed and location of the person.

Step 4. Sequential pattern recognition: Apply the sequential pattern recognition technique to make a decision about whether the person walks straight or is turning, based on the use of a quadratic classifier.

Step 5. Speed calculation of the person: Calculate the speed of the person from the locations of the person in consecutive cycles as a reference for adjusting the speed of the vehicle.

Step 6. Wheel angle control: Adjust the turning angle of the front wheels of the vehicle according to the location of the person when the system decides that the person is turning.

Step 7. Vehicle speed control: Apply the fuzzy control technique to adjust the speed of the vehicle according to the speed of the person and the distance between the person and the vehicle.

In the remainder of this article, the proposed person following technique is described in section 2, including a review of the sequential pattern recognition technique, the description of the proposed classifier, and the features used for the classifier. Wheel angle adjustment strategy is described in section 3. Employed image processing techniques for the detection of feature points are described in section 4. Experimental results are found in section 5. Conclusions are given in section 6.

2. PROPOSED DECISION MAKING BY SEQUENTIAL PATTERN RECOGNITION

In the process of ALV navigation by the person following, the turning angle of the front wheels of the vehicle is decided not only by the location of the person but also by the turning status of the person. When a person means to walk straight, he/she may not be able to march in a real straight trajectory. That is, the location of the person may not be always right in front of the vehicle and might sometimes be a little bit on the right side or on the left of the moving direction of the vehicle. In such cases, the turning angle of the front wheels of the vehicle should keep at 0°. It is undesirable to turn the ALV when the person is not too left or too right of the direction of the vehicle. On the contrary, if the location of the person is too left or too right, the result of the classifier should mean to adjust the direction of the ALV even when the heading difference is small. In short, the main concept here is that the direction of the vehicle should not be adjusted even when there is a small difference between the vehicle heading and the heading of the person.

To implement the above idea, we develop in this study a method based on the sequential pattern recognition technique to decide whether the person walks straight or not. This method does not take as input just a single location of the person; instead, a sequence of the locations of the person extracted from an image sequence is used to provide sufficient information for decision making. The method results in more precise decisions and guides the vehicle to navigate more smoothly.

2.1. Principle of Using Sequential Pattern Recognition for the Person Following Navigation

The principle of sequential pattern recognition is to use a sequence of input data and postpone decision making until sufficient confidence is accumulated. The decidable and undecided areas in the feature space are separated by two threshold values. By this method, the turning angle of the vehicle is not adjusted in every cycle; turning is activated only when a decision is made.

Traditionally, a two-class classifier used for making decisions takes the form,

$$\begin{aligned} h(X) > 0 \qquad &\rightarrow X \in \omega_1 \\ h(X) < 0 \qquad &\rightarrow X \in \omega_2 \end{aligned} \tag{1}$$

where X is an input feature vector, $h(X)$ is a discriminant function, ω_1 and ω_2 denote the two classes, and "$\rightarrow$" means "decide." In sequential pattern recognition,[23] the value of $h(X)$ is not compared with zero but with two threshold values, a and b. We assume that a is smaller than b. When the value of the discriminant function is smaller than a or larger than b, the input feature vector is classified and a decision is made. Otherwise, the value of $h(X)$ is judged to be located on an undecided area and decision making is postponed. In this situation, the value of $h(X)$ is computed for the next input feature vector and added to the prior value of $h(X)$. The discriminant function of the classifier is

$$S_m = \sum_{i=1}^{m} h(X_i) \tag{2}$$

where m denotes the number of all the input feature vectors observed so far. In addition, the classifier is of the form,

$$\begin{aligned} &S_m \leq a \rightarrow X's \in \omega_1 \\ &a < S_m < b \rightarrow \text{take the } (m+1)\text{th sample} \\ &b \leq S_m \rightarrow X's \in \omega_2 \end{aligned} \tag{3}$$

For our case here, ω_1 means the first class that the location of the person is too right or too left of the vehicle, and ω_2 means the second class that both the person and the vehicle move in a straight trajectory. Whenever a decision is made, S_m is reset to 0 and m to 1. This means that S_m is reinitialized and a new decision making process is started. If the vector X does not belong to ω_1 or ω_2, the vector X belongs to an undecided area, which is the last class, and the next sample is taken.

The two threshold values, a and b, in (3) control the error of the sequential pattern recognition result.[24] As the absolute values of a and b become larger, the error reduces, while the number m of observations required to reach the decision increases. If the values of a and b are equal to zero, the three classes are combined into two classes and the undecided area does exist. Also, the vector X will belong either to ω_1 or to ω_2 according to the value of classifier $h(X)$ in Eq. (1), and a larger error will occur. For example, when a person walks straight but his location is a little right of the vehicle, the motion of the person does not belong to ω_1 or ω_2, and the next sample should be considered. If the discriminant function makes a decision at this time, the decision result may be wrong and errors occur. Hence, if b is larger than zero and a is smaller than zero, then the larger the range between a and b, the larger the range of the last class. This will force the discriminant function to use more samples to make a decision and the error caused in making decision will be reduced.

The values of a and b are assigned according to the experiment. If the value of a is taken to be too small, the person will be out of the camera view because it will take too much time to make a decision that the person belongs to ω_1. If the value of a is, on the contrary, taken to be too large, the discriminant function judges most of the time that the person belongs to ω_1 so that the vehicle adjusts its direction frequently and smooth navigation is hard to achieve. According to the value of the discriminant function S_m and the two threshold values a and b, the relation between the location of the person and that of the vehicle is categorized.

2.2. Features for Sequential Pattern Recognition

By (3), a decision is made according to the value of S_m which is calculated from the sequential values of the function $h(X_i)$. In this section, the feature vector X selected for use in this study is described.

Conceptually, if the X coordinate value l of the location of the person in the vehicle coordinate system (VCS) is larger than zero, the person is regarded to stand on the right side of the vehicle;

otherwise, he is regarded to be on the left side of the vehicle. When a person stands on the right side of the vehicle, he/she may turn left, right, or go straight. This means that l is not sufficient for decision making about the movement of the person. We thus choose another feature which is the length w of the projection of the body of the person onto the X axis of the VCS. Because the value of w is calculated from the real width w_b of the body of the person, it is not affected by locations of the person. This means that these two features are independent. When the person walks straight, the value of w is stable and keeps on almost the same value. On the contrary, the value of w becomes smaller when the direction of the person is not parallel to the vehicle. So, the value of w is a good feature for predicting the direction of the person. An illustration of the values of l and w when the person turns right is shown in Figure 2, in which P_l and P_r, to be described in detail in section 4, denote two detected feature points of the back of the person. w_b is the length between P_l and P_r. Assume that their coordinates are $(x_{p_l}, y_{p_l}, z_{p_l})$ and $(x_{p_r}, y_{p_r}, z_{p_r})$ in the VCS, respectively. The valuer of l is $(x_{p_l} + x_{p_r})/2$ and that of w is $(x_{p_r} - x_{p_l})$. The selected feature vector X is

$$X = \begin{bmatrix} l \\ w \end{bmatrix} \tag{4}$$

2.3. Classifier for Sequential Pattern Recognition

Since the feature vector X is two-dimensional, the decision boundary of the classifier is a two-dimensional curve in the L-W space. When the person stands on the right side of the vehicle, l is larger than zero; otherwise, l is smaller than zero. Besides, when the person walks straight, w is equal to the width w_b of the body of the person; otherwise, w is smaller than w_b.

In this study, it is found that the use of the two values of l and w is enough to identify three classes and derive the classifier $h(X)$. The first class is that the location of the person is too right or too left of the vehicle. We use the width w_v of the vehicle as a boundary. If the absolute value of the X coordinate of the location of the person is larger than w_v, the location of the person is regarded to belong to the first class; otherwise, the direction of the person has to be considered. The second class is that both the person and the vehicle move in a straight trajectory. In the optimal situation, this means that w is equal to w_b. However, the direction of the person may drift. We use a half value of w_b as a boundary. If the value of w is smaller than a half value of w_b, the location of the person is not regarded to belong to the second class. It means that the person is turning right or left. Based on the above observation, we design a classifier $h(X)$ to be a parabola curve which passes through $(0, w_b/2)$ and (w_v, w_b), as illustrated in Figure 3. The equation for $h(X)$ is derived accordingly as follows (the details are omitted),

$$h(X) = W - \left(\frac{w_b}{2w_v^2}\right)L^2 - \frac{w_b}{2} \tag{5}$$

According to the classifier $h(X)$, the L-W plane is separated into three areas, S, T_r, and T_l, as illus-

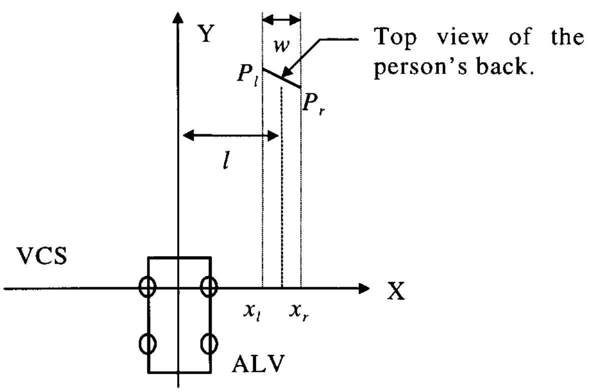

Figure 2. Top view of the vehicle coordinate system (VCS) and illustration of the X coordinate value l of the location of the person and the length w of the projection of the body of the person when he turns right.

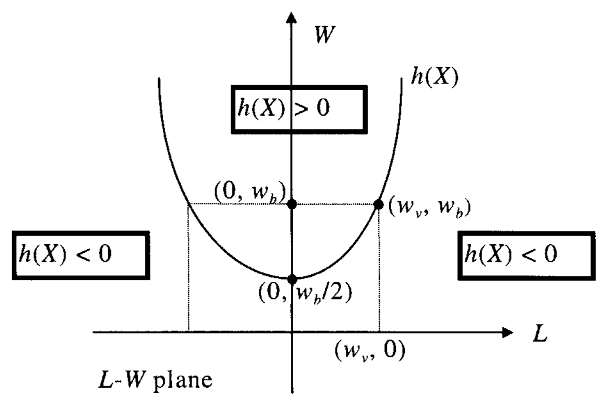

Figure 3. Illustration of the classifier, $h(X)$. w_b and w_v are the widths of the person and the vehicle, respectively.

trated in Figure 4. The area S means that the person walks straight and belongs to ω_2. The areas T_r and T_l mean that the location of the person is too right and too left of the vehicle, respectively, and belongs to ω_1. If the location of the person is in the area T_r or T_l, the direction of the vehicle has to be changed. It does not mean that the person is turning right or left. Besides, if the person is not too left ($l < 0$) and the person turns right (w small), the feature vector X is located near the boundary of the classifier $h(X)$ and the absolute value of $h(X)$ is a little larger than zero. In this case, according to Eq. (3), the vector X belongs to an undecided area, i.e., to class 3, and the vehicle keeps the current direction and the person is still in the camera view. If the person is too left ($l < 0$) and the person turns right (w small), then the location of the person belongs to ω_1, and the vehicle adjusts the direction of the vehicle according to the location of the person. Hence, the location of the person will not be out of the camera view in the next cycle. If the person walks very fast so that the person is out of the camera view, the vehicle will stop.

With the three classes, corresponding adjustments of the direction of the vehicle are described in the next section.

3. WHEEL ANGLE CONTROL

The wheel angle control is triggered after the decision that the person belongs to the class ω_1 or ω_2 is made. If the person belongs to the second class ω_2, the wheel angle of the vehicle is adjusted to be 0°. If the person belongs to class 3, the wheel angle of the vehicle keeps on the current one. If the person belongs to the first class ω_1, the direction of the vehicle is adjusted to follow the person. The turning angle δ of the front wheels of the vehicle is computed according to the coordinates (x_{p_b}, y_{p_b}) of the location of the person in the VCS.

When the location of the person belongs to ω_1, it does not mean that the person always turns 90°. The person might walk straight but with his location too right of the vehicle. In such a case, the direction of the vehicle is adjusted. Otherwise, the person will be out of the camera view. The angle δ is calculated according to the current location of the person in the vehicle coordinate system (VCS) and the adjustment of the direction of the vehicle toward the person is made accordingly. The derivation of the turning angle δ for the front wheels of the vehicle is based on the kinematic trajectory of the vehicle.

As shown in Figure 5 which is a top view of the VCS, assume that the vehicle is located at A and the person is located at P_b with coordinates (x_{p_b}, y_{p_b}) in the VCS. Let the vehicle move a distance S from location A to location P_b by turning an angle δ. Let S be the kinematic trajectory of the vehicle and δ be the angle that we want to obtain. By assuming that x_{p_b} and y_{p_b} have been calculated (see section 4), the values of P and u in Figure 5 can be computed as

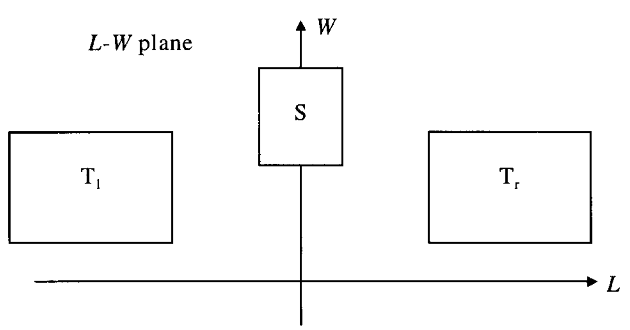

Figure 4. The area of S means that the person walks straight (class 2). The areas of T_r and T_l mean that the location of the person is too right and too left of the vehicle, respectively, (class 1).

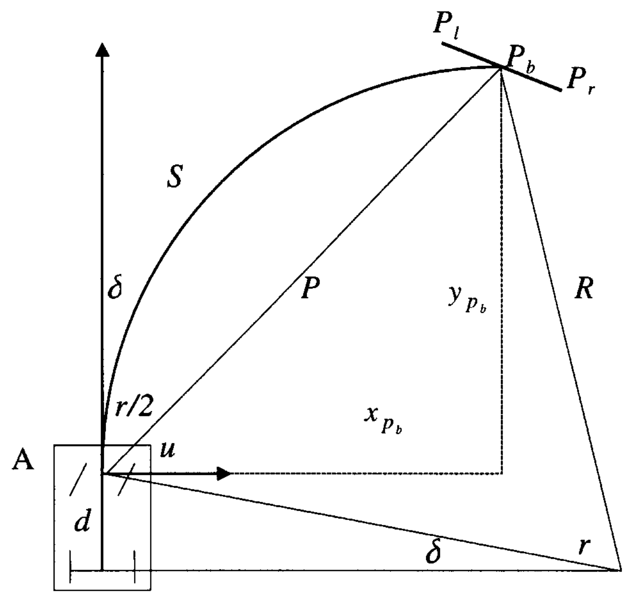

Figure 5. Analysis of the path S on which the vehicle navigates from location A to location P_b by turning an angle δ. The coordinates of location P_b are (x_{p_b}, y_{p_b}) in the VCS.

follows,

$$P = \sqrt{x_{p_b}^2 + y_{p_b}^2}$$
$$u = \tan^{-1}\left(\frac{y_{p_b}}{x_{p_b}}\right) \quad (6)$$

where P, x_{p_b}, and y_{p_b} compose a right triangle and u is the corresponding angle of y_{p_b}. On the other hand, P, R, and u can be acquired by the following equations[25]:

$$P = R\sqrt{2(1-\cos r)} \quad (7)$$

$$R = \frac{d}{\sin \delta} \quad (8)$$

$$u = \frac{\pi}{2} - \delta - \frac{r}{2} \quad (9)$$

where R is the rotation radius, d is the distance between the front wheels and the rear wheels, r is the corresponding angle of S, and δ is the turning angle of the front wheels. From Eqs. (7)–(9), we can obtain

$$G = \frac{P}{2d}$$
$$D = \frac{\pi}{2} - u \quad (10)$$
$$\delta = \tan^{-1}\left(\frac{\sin D}{\cos D + G}\right)$$

where d is a constant, and P and u are calculated according to (6). After substituting the values of P, u, D, and G into the last equation in (10), we get a more detailed equation for δ as follows,

$$\delta = \tan^{-1}\left(\frac{\sin\left(\frac{\pi}{2} - \tan^{-1}\left(\frac{y_{p_b}}{x_{p_b}}\right)\right)}{\cos\left(\frac{\pi}{2} - \tan^{-1}\left(\frac{y_{p_b}}{x_{p_b}}\right)\right) + \frac{\sqrt{x_{p_b}^2 + y_{p_b}^2}}{2d}}\right) \quad (11)$$

4. VISION SYSTEM

In this study, it is desired to control the vehicle to navigate smoothly by following a person who walks in front of the vehicle. An important source of information used in this study is the location and the direction of the person. Only by the use of correct coordinates of the location of the person can the proposed system drive the vehicle to navigate correctly. Because the detection of the whole human body wastes a lot of time, we attach a rectangular shape on the shirt of the person for fast detection and tracking of the person to shorten the processing time. Images with the shirt of the person are grabbed by a color CCD camera. The two upper corners of the rectangular shape are detected by image processing techniques. We have two reasons for detecting just the two upper corners of the rectangular shape instead of the four. One is that the two upper corners are sufficient to provide the width of the shape. The other is that the two lower corners are out of the camera view when the person is close to the vehicle. The second reason is related to the position of the camera. In this study, the camera is mounted on the front of the vehicle. Besides, the position of the camera is set at an average height of human beings and with a little tilt downward to the ground. Hence, the whole shape can be detected when the distance between the person and the vehicle is larger than 1 or 2 m. In the experiment, the position of the camera is 35 cms higher than the feature points to be detected. When the two upper corners are detected in the image, the three-dimensional coordinates of the detected corners in the vehicle coordinate system are calculated to obtain the location of the person. In this study, the height of the detected corners is assumed to be known by measuring it directly in advance. The camera calibration is implemented using the information of vanishing lines.[26] The coordinate system[27] and the transformations between them are adopted from refs. 28 and 29 for use in this study.

4.1. Feature Point Detection by Image Processing Techniques

The feature point we use for locating the person are two points on the back of the person which represent proportionally the width of the person. As shown in Figure 6a, a rectangular shape is attached on the shirt of the person. The color of the rectangular shape is yellow. We segment the shape out of the back of the person by thresholding. Next, a region growing method is used to detect the border of the rectangle. Finally, the intersection point of

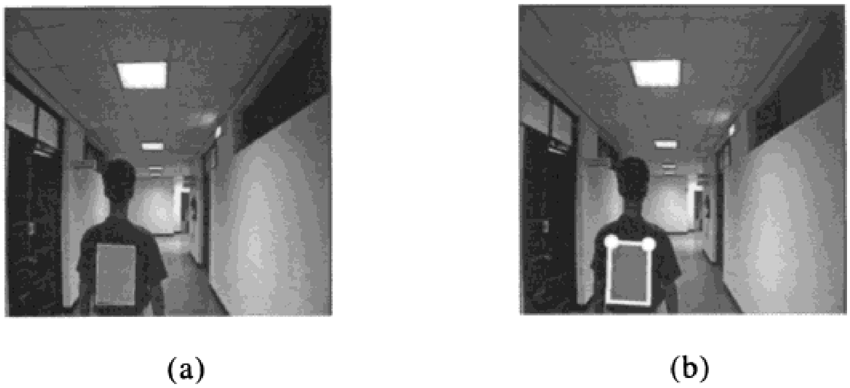

Figure 6. The experimental result of feature point detection. (a) is a captured image with the segmented rectangular shape. (b) Detected feature points.

the left border and the upper border, and that of the right border and the upper border, are detected as the two desired feature points, as shown in Figure 6b.

4.2. Three-Dimensional Coordinate Calculation by Computer Vision Techniques

In this study, the ALV navigation environment is described by three coordinate systems, the image coordinate system (ICS), the camera coordinate system (CCS), and the vehicle coordinate system (VCS),[27] as shown in Figure 7.

The transformations between the ICS, the CCS, and the VCS are described as follows. Assume that any point P in the image plane has the CCS coordinates $(u_P, -f, w_P)$, where (u_P, w_P) specify the coordinates in the ICS and f is the focus length. We get the VCS coordinates (x_P, y_P, z_P) of point P in the image[4] as

$$\begin{aligned} x_P &= u_P(\cos\theta\cos\psi + \sin\theta\sin\phi\sin\psi) \\ &\quad + f(\sin\theta\cos\phi) + w_P(\sin\theta\sin\phi\cos\psi \\ &\quad + \cos\theta\cos\psi) + x_d \\ y_P &= u_P(\sin\theta\cos\psi + \cos\theta\sin\phi\sin\psi) \\ &\quad - f(\cos\theta\cos\phi) - w_P(\cos\theta\sin\phi\cos\psi \\ &\quad + \sin\theta\sin\psi) + y_d \\ z_P &= u_P(\cos\theta\sin\psi) - f\sin\phi + w_P(\cos\theta\cos\psi) + z_d \end{aligned} \tag{12}$$

where (x_d, y_d, z_d) are the VCS coordinates specifying the translation vector from the origin of the VCS to the origin of the CCS, θ, ϕ, and ψ are the pan,

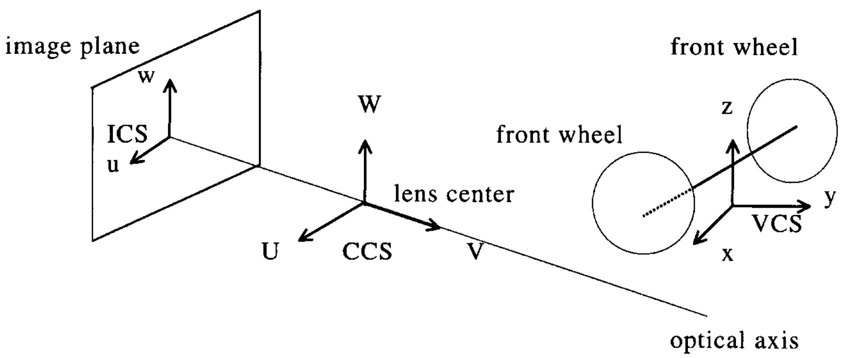

Figure 7. The three coordinate systems, including the image coordinate system (ICS), the camera coordinate system (CCS), and the vehicle coordinate system (VCS).

tilt, and swing angles, respectively, of the camera with respect to the VCS. These camera parameters are calibrated[26] in advance. In addition, the equation of a line LP, which passes the lens center and P, is

$$\frac{x - x_d}{x_P - x_d} = \frac{y - y_d}{y_P - y_d} = \frac{z - z_d}{z_P - z_d} = k \tag{13}$$

As shown in Figure 8, let point P' be the intersection point of the plane $z = h$ and the line LP. By substituting $z = h$ into Eq. (13), the desired VCS coordinates $(x_{P'}, y_{P'}, z_{P'})$ of point P' can be solved to be

$$\begin{aligned} x_{P'} &= x_d + \frac{h - z_d}{z_P - z_d}(x_P - x_d) \\ y_{P'} &= y_d + \frac{h - z_d}{z_P - z_d}(y_P - y_d) \\ z_{P'} &= h \end{aligned} \tag{14}$$

In this study, assume the image coordinates of the feature points P_l and P_r are (u_{p_l}, w_{p_l}) and (u_{p_r}, w_{p_r}), respectively. In addition, the height z_{p_b} of the detected feature points is assumed to be known. The coordinates $(x_{p_l}, y_{p_l}, z_{p_l})$ and $(x_{p_r}, y_{p_r}, z_{p_r})$ of the features points P_l and P_r are obtained by Eqs. (12) and (14) with $h = z_{p_b}$ in the vehicle coordinate system. Hence, the coordinates $(x_{p_b}, y_{p_b}, z_{p_b})$ of the location of the person P_b can be calculated as follows,

$$\begin{aligned} x_{p_b} &= \frac{x_{p_l} + x_{p_r}}{2} \\ y_{p_b} &= \frac{y_{p_l} + y_{p_r}}{2} \\ z_{p_b} &= z_{p_l} = z_{p_r} \end{aligned} \tag{15}$$

5. EXPERIMENTAL RESULTS

The proposed method has been implemented and used to steer a real ALV autonomously to follow a person, as shown in Figure 9. Smooth navigation was demonstrated while the person walks in different directions. In our experiments, the width w_v and

Figure 9. The vehicle is guided automatically to follow a person who walks in front of the vehicle.

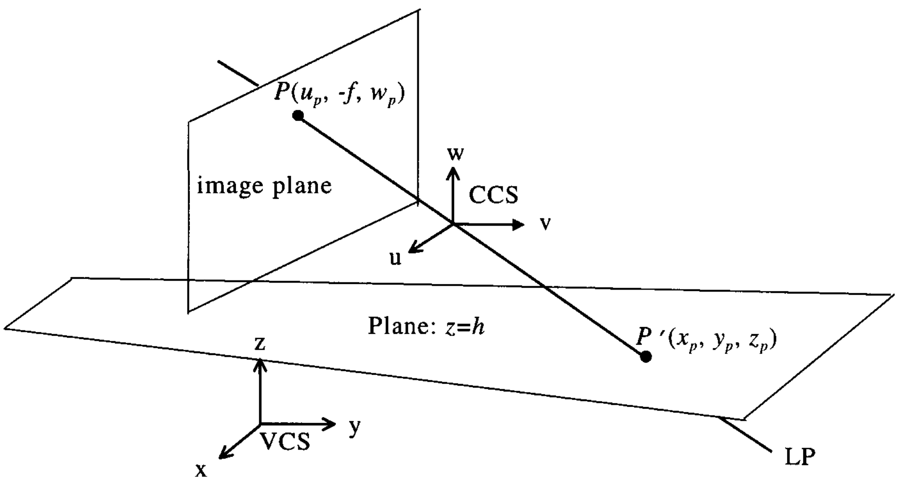

Figure 8. Illustration of the coordinate transformations between the camera coordinate system (CCS) and the vehicle coordinate system (VCS) when the point P' is located in the plane $z = h$.

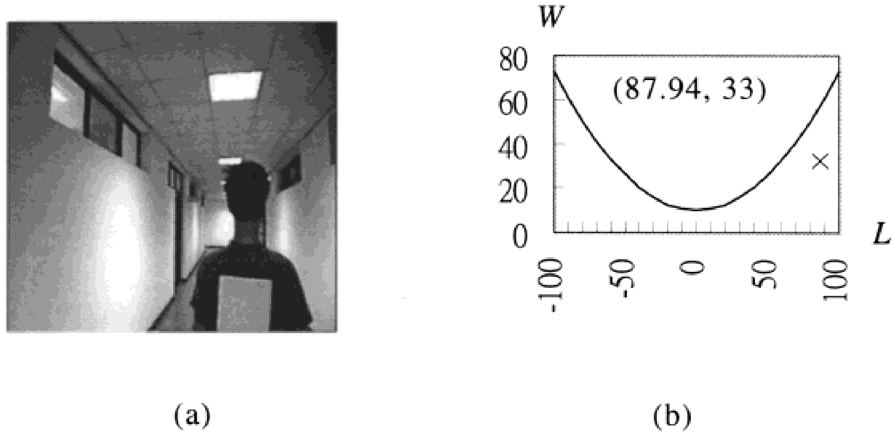

Figure 10. A situation where the location of the person is too right of the vehicle. (a) The result of the image processing. (b) The location × of the value of the discriminant function in the *L*-*W* plane. The coordinates, x_b and y_b, of the location of the person in the VCS are 87.94 and 201.07, respectively. The value w is 33. The value of the discriminant function $h(X)$ is -166.23 and the turning angle, δ, of the front wheels is 10.22°.

the length of the ALV are 40 and 120 cm, respectively. The length between the front wheels and the rear wheels of the ALV is 82 cm. The width w_b of the rectangular shape is 30 cm. The X-Y-Z coordinates of the camera in the VCS are (17.7, 3.7, 167.9). The pan, tilt, and swing angles of the camera are -0.007, 0.005, and 0.03, respectively. In the experiment, the threshold values a and b used in Eq. (3) are -15 and 15, respectively.

The proposed relations between the location of the vehicle and that of the person and the corresponding actions taken by the ALV are verified in the experiments. Figures 10 and 11 illustrate that the location of the person is too right and too left, respectively, and both cases belong to the first class ω_1. Figure 10 shows that the person walks in a straight trajectory and Figure 11 shows that the person turns left. The values of the discriminant

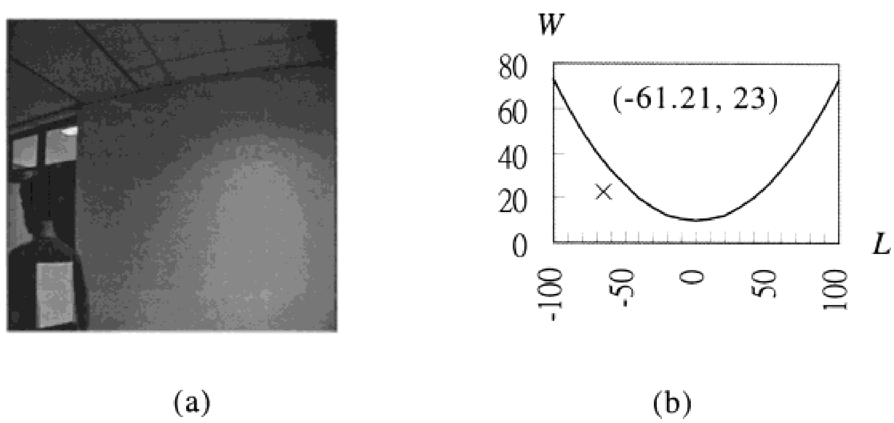

Figure 11. A situation where the location of the person is too left of the vehicle. (a) The result of the image processing. (b) The location × of the value of the discriminant function in the *L*-*W* plane. The coordinates, x_b and y_b, of the location of the person in the VCS are -61.21 and 202.15, respectively. The value w is 23. The value of the discriminant function $h(X)$ is -80.03 and the turning angle, δ, of the front wheels is $-7.46°$.

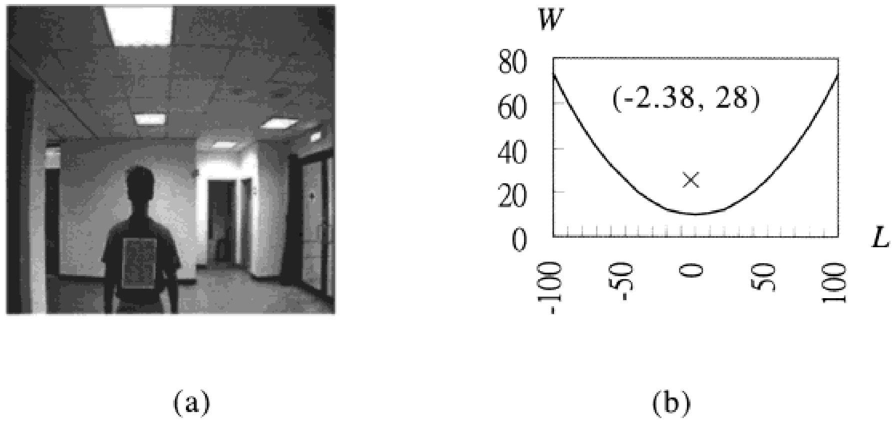

Figure 12. The situation that the person walks straight. (a) The result of the image processing. (b) The location × of the value of the discriminant function in the L-W plane. The coordinates, x_b and y_b, of the location of the person in the VCS are -2.38 and 232.39, respectively. The value w is 28. The value of the discriminant function $h(X)$ is 17.90 and the turning angle, δ, of the front wheels is 0°.

function are both smaller than the threshold value a. The feature vectors X are located in the areas of T_r and T_l, respectively, in the L-W plane (denoted with "×"). In Figure 10, the two lower corners of the shape are out of the camera view and the two upper corners are detected stably. Figure 12 shows that the person walks straight and the case belongs to the second class ω_2. The value of the discriminant function is larger than the threshold value b and the feature vector is located in the area S in the L-W plane. Figure 13 shows that the feature vector is located in the undecided area. The value of the discriminant function is in the range from the threshold value a to the threshold value b. The direction of the vehicle keeps on the current one. Figure 14 shows a sequence of experimental images in a real indoor environment. A top view of a sequence of the trajectory of the vehicle and that of the person is shown in Figure 15. In the experimental results, if the value of δ is smaller than zero,

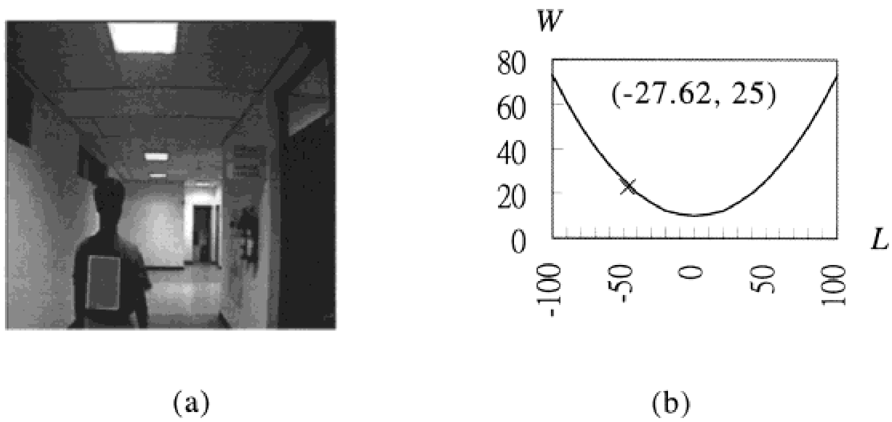

Figure 13. The situation that the feature vector is located in the undecided area. (a) The result of the image processing. (b) The location × of the value of the discriminant function in the L-W plane. The coordinates, x_b and y_b, of the location of the person in the VCS are -27.62 and 224.54, respectively. The value w is 25. The value of the discriminant function $h(X)$ is -3.23.

Figure 14. A sequence of experimental images in the corridor of a real indoor environment.

then it means that front wheels of the vehicle turn left; otherwise, the vehicle turns right.

The accuracy in the computation of the location of the person is shown in Table I, which includes the location of the person measured by hand and that computed by the calibrated vision system. The distance error rate in the table is defined as the ratio of the Euclidean distance between the measured position of the person and the computed position of the person to that between the measured position of the person and the origin of the VCS, which is expressed as

$$\varepsilon = \frac{\sqrt{(x_e - x_c)^2 + (y_e - y_c)^2}}{\sqrt{x_e^2 + y_e^2}} \tag{15}$$

where (x_e, y_e) are the VCS coordinates of the measured location of the person and (x_c, y_c) are the VCS coordinates of the computed location of the person. The average distance error rate computed from all of the error rates in the table is 5.6%, which is acceptable for safe navigation.

6. CONCLUSIONS

In this study, a new approach to ALV navigation by the person following in indoor environments has been proposed. This approach uses sequential pattern recognition and computer vision techniques to achieve smooth navigation. A rectangular shape attached on the back of the person is detected by the computer vision technique and two effective features are used in calculating the location of the person. Then sequential pattern recognition is used in making decisions about whether the person walks straight or is too left or too right of the vehicle. This

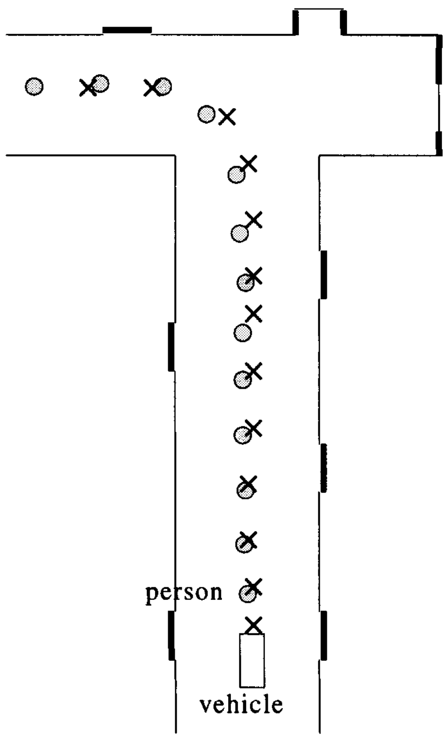

Figure 15. A top view of the experimental indoor environment and a continuous navigation session, in which each black circle indicates the location of the person in a navigation cycle and × means the location of the vehicle.

Table I. Computed location of the person vs. measured location of the person.

	Measured location of the person (x_e, y_e)	Computed location of the person (x_c, y_c)	Distance error rate (%)
1	(10, 100)	(11.06, 93.88)	6.2
2	(15, 200)	(16.40, 186.81)	6.6
3	(10, 300)	(14.31, 311.92)	4.2
4	(60, 200)	(66.80, 187.70)	6.7
5	(5, 250)	(3.96, 233)	6.8
6	(−10, 150)	(−7.97, 145.22)	3.5
7	(−20, 220)	(−24.75, 229.25)	4.7
8	(40, 180)	(49.73, 174.33)	6.1

approach has been implemented on a real ALV, and successful and smooth navigation sessions confirm the feasibility of the approach.

The authors thank Kuang-Hsiung Chen in the Computer Vision Laboratory of the Department of Computer and Information Science at National Chiao Tung University for helpful discussions and assistance in the experiments of this research.

REFERENCES

1. M.S. Chang, P.Y. Ku, L.L. Wang, and W.H. Tsai, Indoor autonomous land vehicle guidance by line following using computer vision techniques, Proc 1989 Workshop Comput Vision, Graph and Image Processing, Taipei, Taiwan, Republic of China, 1989, pp. 135–144.
2. B. Matthews, M. Ruthemeyer, D. Perdue, and E.L. Hall, Line following for a mobile robot, Proc SPIE—The Int Soc Opt Eng, Vol. 2588, 1995, pp. 610–617.
3. W.J. Ke and W.H. Tsai, Indoor autonomous land vehicle guidance by corner tracking using computer vision, Proc 1991 Workshop on Comput Vision, Graph and Image Processing, Tainan, Taiwan, Republic of China, 1991, pp. 133–139.
4. Y.M. Su and W.H. Tsai, Autonomous land vehicle guidance for navigation in buildings by computer vision, radio, and photoelectric sensing techniques, J Chinese Institute Eng, 17 (1994), 63–73.
5. H.I. Christensen, N.O. Kirkeby, S. Kristensen, L. Knudsen, and E. Granum, Model-driven vision for indoor navigation, Robot Autonomous Syst, 12 (1994), 199–207.
6. L.L. Wang, P.Y. Ku, and W.H. Tsai, Model-based guidance on the longest common subsequence algorithm for indoor autonomous vehicle navigation using computer vision, Automat Construction, 2 (1993), 123–137.
7. T. Fukuda, S. Ito, N. Oota, F. Arai, Y. Abe, K. Tanaka, and Y. Tanaka, Navigation system based on ceiling landmark recognition for autonomous mobile robot, Proc IECON Int Conf Ind Electron, Control, and Instrumentation, 3, 1993, pp. 1466–1471.
8. A. Gilg and G. Schmidt, Landmark-oriented visual navigation of a mobile robot, IEEE Trans Ind Electron, 41 (1994), 392–397.
9. C. Thrope, M.H. Hebert, T. Kanade, and S.A. Shafer, Vision and navigation for Carnegie-Mellon NAVLAB, IEEE Trans Pattern Anal Machine Intell, 10 (1988), 362–373.
10. J.D. Crisman and C.E. Thorpe, SCARF: A color vision system that tracks roads and intersections, IEEE Trans Robot Automat, 9 (1993), 49–58.
11. C.C. Chien, M.C. Lai, and R. Mayr, Vehicle following controller design for autonomous intelligent vehicles, Conference on Intelligent Robotics in Field, Factory, Service, and Space, Houston, TX, Vol. 1, 1994, pp. 212–221.
12. D. Noll, M. Werner, and W. von Seelen, Real-time vehicle tracking and classification, Proc Intell Veh Symp, 1995, pp. 101–106.
13. K.I. Kim, S.Y. Oh, S.W. Kim, H. Jeong, J.H. Han, C.N. Lee, B.S. Kim, and C.S. Kim, An autonomous land vehicle PRV II: progresses and performance enhancement, Proc Intell Veh '95 Symp, Detroit, MI, 1995, pp. 264–269.
14. M. Schwarzinger, T. Zielke, D. Noll, M. Brauckmann, and W. von Seelen, Vision-based car-following: detection, tracking, and identification, IEEE Intelligent Vehicles Symposium, 1992, pp. 24–29.
15. H. Schneiderman, M. Nashman, A.J. Wavering, and R. Lumia, Vision-based robotic convoy driving, Mach Vis Appl, 8 (1995), 359–364.
16. H.K. Nishihara, H.J. Thomas, and E. Huber, Real-time tracking of people using stereo and motion, Proc SPIE—The Internat Soc Opt Eng, Vol. 2183, 1994, pp. 266–273.
17. C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, Pfinder: Real-time tracking of the human body, Proc SPIE—Integration Issues Large Commercial Media Delivery Syst, Vol. 2615, 1996, pp. 89–98.
18. A. Azarbayejani, C. Wren, and A. Pentland, Real-time 3-D tracking of the human body, Proc IMAGE'COM 96, France, May 1996.
19. K. Matsuda, M. Kamihira, M. Ohchi, and T. Furukawa, Design and implementation of tracking control system of autonomous mobile robot with

multiple sensors, Rep Fac Sci Eng Saga Univ, 24 (1995), 73–82.
20. U.D. Hanebeck and G. Schmidt, A new high performance multisonar system for fast mobile robots, IROS Proc IEEE/RSJ/GI Internat Conf Intell Robots and Syst Advanced Robotic Syst Real World, Vol. 3, 1994, pp. 1853–1860.
21. H. Mori and M. Sano, A guide dog robot Harunobu-5—following a person, Proc IROS, IEEE/RSJ Internat Workshop Intelli Robots and Syst Intell Mech Syst, Vol. 1, 1991, pp. 397–402.
22. C.H. Ku and W.H. Tsai, Smooth autonomous land vehicle navigation in indoor environments by person following using sequential pattern recognition, fuzzy speed control and computer vision techniques, Proc Joint Conf Fifth Int Conf Automation Technol 1998 Int Conf Prod Res (Asia Meeting), Grand Hotel, Taipei, Taiwan, Republic of China, July 1998, p. 115.
23. K. Fukunaga, Introduction to statistical pattern recognition, 2nd ed., Academic Press, San Diego, CA, 1990.
24. A. Wald, Sequential analysis, Wiley, New York, 1947.
25. S.D. Chang and W.H. Tsai, Model-based guidance of autonomous land vehicles in indoor environments by structured light using vertical line information, J Electr Eng, 34 (Dec. 1991).
26. L.L. Wang and W.H. Tsai, Camera calibration by vanishing lines for 3D computer vision, IEEE Trans Pattern Anal Machine Intell, 13 (1991), 370–376.
27. S.D. Cheng and W.H. Tsai, Model-based guidance of autonomous land vehicle in indoor environments by structured light using vertical line information, J Electr Eng, 34 (1991), 441–452.
28. R.M. Haralick and L.G. Shapiro, Computer and robot vision, Vol. 2, Addison-Wesley, Reading, MA, 1993.
29. A. Watt, 3D computer graphics, Addison-Wesley, Reading, MA, 1993.