

Intelligent Pose Control of Mobile Robots Using an Uncalibrated Eye-in-Hand Vision System

T. I. James Tsay and Y. F. Lai
*National Cheng Kung University
 Taiwan (R.O.C.)*

1. Introduction

The use of material transfer units constructed from AGVs and robot manipulators is becoming popular in production lines with small production volumes and periods. This new material transfer unit is also called a mobile robot, which can quickly adapt to the change of the factory layout. A mobile robot is sufficiently mobile that it can very flexibly perform tasks in the production line. A guidance control system drives a mobile base with a robot manipulator mounted on it to a predefined station so the robot manipulator can pick up a workpiece in the station.

During pick-and-place operations, the relative location between a predefined station and the robot manipulator cannot be anticipated due to the possibility of the non-planar ground and/or the positioning errors caused by the guidance control system of the mobile base. Uncertainties in either the location of the mobile base or the level of ground flatness are responsible for a position mismatch, which normally causes failure in the pick-and-place operation. A more direct way of dealing with this problem is to use a CCD camera to provide visual feedback. The camera is typically mounted on the end-effector of the robot manipulator to compensate for these uncertainties using a visual servoing scheme.

Hutchinson et al., 1996 classified the visual feedback control systems into four main categories. Position-based control aims to eliminate the errors determined by the pose of the target with respect to the camera; the features extracted from the image planes are used to estimate the pose in space. The control law then sends the command to the joint-level controllers to drive the servomechanism. This control architecture is the so-called hierarchical control and is referred to as the dynamic position-based look-and-move control structure. The control architecture is referred to as the position-based visual servo control structure if the servomechanism is directly controlled by the control law mentioned above rather than by the joint-level controller. The errors to the control law in image-based control, are directly governed by the extracted features of the image planes. The control architecture is hierarchical and referred to as the dynamic image-based look-and-move control structure if the errors are then sent to the joint-level controller to drive the servomechanism. The visual servo controller eliminates the joint-level controller and directly controls the servomechanism. Such control architecture is termed an image-based visual servo control structure.

Source: Mobile Robots, Moving Intelligence, ISBN: 3-86611-284-X, Edited by Jonas Buchli, pp. 576, ARS/pIV, Germany, December 2006

The mobile robots (Ting et al., 1997) made by Murata Machinery, Ltd. meet the necessary cleanliness standards required in current semiconductor production facilities. The mobile base of a mobile robot is first guided to a predefined station using CCD cameras mounted on the ceiling. Then, the eye-in-hand manipulator performs a pick-and-place operation by employing a static position-based look-and-move control structure. A pair of LEDs is installed in front of a specific location, where a cassette is picked up or put down, to provide a reference frame. The relative positions of the LEDs and the cassette is fixed and known. The trajectory along which the end-effector of the manipulator approaches the cassette can be divided into two parts. The first part of the trajectory is determined off-line using a teaching box to drive the end-effector to a location that allows the CCD camera mounted on the end-effector to obtain a pose to capture the image of both LEDs. After the end-effector follows the part of the trajectory planned off-line, the CCD camera guides the second part of the trajectory. The location of the cassette with respect to the manipulator base frame is first computed through the coordinate transformations and perspective transformation in response to the image of two LEDs. The end-effector is then commanded to move to the above computed location and pick up the cassette. The use of two LEDs makes success in the pick or place task highly dependent on the pose of the end-effector relative to the LEDs when the image of the two LEDs is captured. Consequently, the non-horizontality of the ground and the large positioning errors of the mobile base may cause the pick or place task to fail.

Beyond the conventional visual servo control methods referred to above, behavior-based methods for visual servo control have already been developed in the literature (Anglani et al., 1999; Kim et al., 2001; Wasik & Saffiotti, 2002, 2003). A behavior-based system has been proposed to perform grasping tasks in an unstructured environment, in cases in which the position of the targets is not already known (Wasik & Saffiotti, 2002, 2003). The controller maps input from the image space to the control values defined in the camera space. The control values are then transformed to joint controls by a Jacobain transformation, revealing that either a hand-eye calibration process has been implemented or the hand-eye relationship is known beforehand.

This study employs an uncalibrated eye-in-hand vision system to provide visual information for controlling the end-effector of the manipulator to approach and grasp the target workpiece. This work focuses on developing a novel behavior-based look-and-move control strategy to guide the manipulator to approach the object and accurately position its end-effector in the desired pose. Notably, the pose of the workpiece in relation to the end-effector of the manipulator is never numerically estimated. Also, the relationships between the deviations of the image features and the displacements of the end-effector are never determined. This strategy is based on six predefined image features. In the designed neural fuzzy controllers, each image feature is taken to generate intuitively one DOF motion command relative to the camera coordinate frame using fuzzy rules, which define a specific visual behavior. These behaviors are then combined and executed in turns to perform grasping tasks.

2. Image Processing

This section proposes the image processing method that we used to extract target object features in a visual servo system. In this study, the workpiece to be picked up by the

manipulator's end-effector is a rectangular parallelepiped. The images of the workpiece are captured from above by a CCD camera mounted on the end-effector, as the end-effector moves toward the workpiece. The image of the workpiece's top surface is a quadrangle. Only information about the quadrangle is of interest, so the captured image is firstly preprocessed to obtain a clear image of the quadrangle. Then, the selected image features can be calculated from the quadrangle.

Fig. 1 presents a complete image processing flowchart. The captured image is first preprocessed using the thresholding technique and opening and closing operations to yield a clear binary image of the workpiece's top surface, from which the area, the center of gravity, and the principal angle of the quadrangle in the image plane can be obtained. The Laplace operator is then applied to estimate quickly the approximate locations of the corners. Once the approximate corner locations are determined, four square windows can be obtained from the approximate locations of four corners, (x_i, y_i) , $i = 1, 2, 3, 4$, as shown in Fig. 2, to specify that all the edge points encircled by each window are on the same edge, to determine the precise location of the corners. Each square window can be established by the following procedure.

1. Select two adjacent corners, (x_1, y_1) and (x_2, y_2) .
2. Determine the width and length of the window, w and l .

$$w = |x_1 - x_2| \times 0.8 \quad \text{pixels}$$

$$l = |y_1 - y_2| \times 0.8 \quad \text{pixels}$$

Neither the width nor the length of the window are chosen as the full coordinate differences between two adjacent corners' locations to prevent incorrect encirclement due to the error in quickly estimating the approximate locations of the corners. However, if $|x_1 - x_2| < 5$ pixels, then the width is defined as $w = 10$ pixels, to ensure that at least a few edge points are encircled by the window, and if $|y_1 - y_2| < 5$ pixels, the length is defined as $l = 10$ pixels, to ensure that at least a few edge points are encircled by the window.

3. Calculate the coordinates of the window's upper left corner and the window's lower right corner, $w_{left}^{upper} = (x_{left}^{upper}, y_{left}^{upper})$ and $w_{right}^{lower} = (x_{right}^{lower}, y_{right}^{lower})$, from the following equations.

$$x_{left}^{upper} = (x_1 + x_2) / 2 - w / 2$$

$$y_{left}^{upper} = (y_1 + y_2) / 2 - l / 2$$

$$x_{right}^{lower} = (x_1 + x_2) / 2 + w / 2$$

$$y_{right}^{lower} = (y_1 + y_2) / 2 + l / 2$$

The coordinates of points on a single edge can then be fitted into a line equation using least-error-squares. The precise location of each corner can then be obtained from the point of intersection of two adjacent lines.

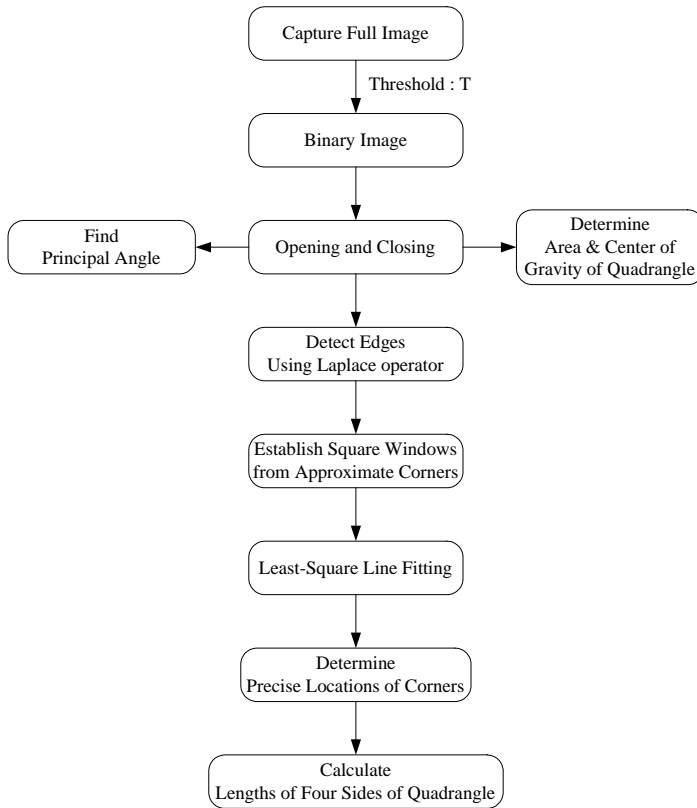


Fig. 1. Complete image processing flowchart.

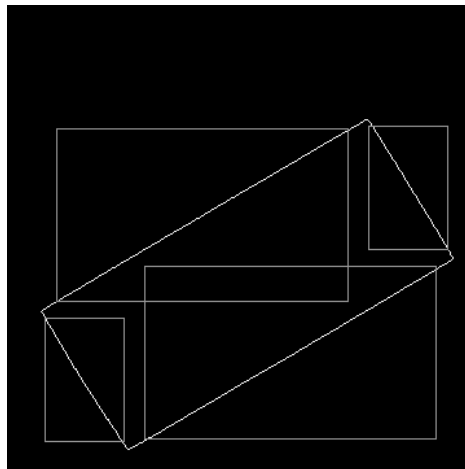


Fig. 2. Using corners' approximate locations to specify four windows to yield four line equations.

3. Control Strategy for Approaching the Target

Six designed neural fuzzy controllers map image features in the image space onto motion commands in the camera space. Also, each mapping from image space to camera's Cartesian space is mediated by fuzzy rules, and defines a particular vision-based behavior. Notably, these behaviors are defined from the perspective of the camera but not from that of the end-effector or external observer. This section defines a decomposition of the manipulation task into a set of basic behaviors, which are combined so that the eye-in-hand camera can gradually reach the desired position and orientation. The camera is fixed on the end-effector, so when the end-effector reaches the desired pose, the gripper can be commanded to grasp the workpiece.

3.1 Selection of Image Features

Six image features are adopted to determine the translational and orientational motion in 3D. Each image feature can uniquely direct one D.O.F. of motion relative to the camera frame. Before the image features are defined, the coordinate symbols are explicated. Specifically, as shown in Fig. 3, let $\delta^E X_1$, $\delta^E X_2$ and $\delta^E X_3$, represent differential changes in translation along the ${}^E X$, ${}^E Y$ and ${}^E Z$ axes of the end-effector frame. Let $\delta^E X_4$, $\delta^E X_5$ and $\delta^E X_6$ denote the differential changes in orientation about the ${}^E X$, ${}^E Y$ and ${}^E Z$ axes of the end-effector frame. Additionally, the camera frame must be considered. The origin of the camera frame is at the center of the camera, while its z axis is the optical axis of the camera. Similarly, $\delta^C X_i$ for $i=1, 2, \dots, 6$, represent the differential movements with respect to the camera frame.

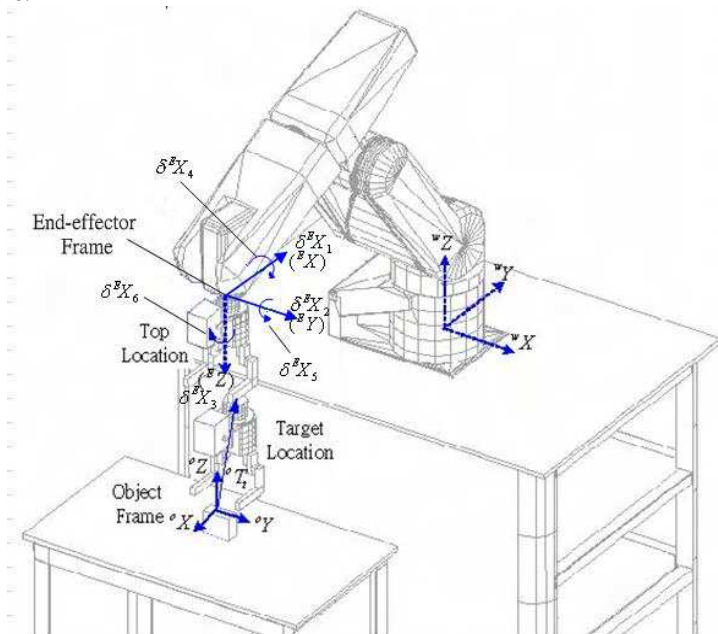


Fig. 3. End-effector and object frames.

In this study, the workpiece to be picked up by the manipulator's end-effector is a rectangular parallelepiped with a quadrangular image, as illustrated in Fig. 4. Six image features are extracted. Five image features, F_1 , F_2 , F_4 , F_5 and F_6 used in (Suh & Kim, 1994) are adopted. The other image feature called relative distance (RD) used in (Bien & Park, 1993) is taken as F_3 . Clearly, the image features are defined as follows.

$F_1 = V$, coordinate of the center of gravity of the quadrangle in the image plane. It directs camera motion ${}^c X_1$;

$F_2 = U$, coordinate of the center of gravity of the quadrangle in the image plane. It directs camera motion ${}^c X_2$;

$F_3 = RD = (A_1 - A_2)/A_1$, where A_1 is the area of the quadrangle in the desired pose, and A_2 represents the area of the quadrangle in the present pose. F_3 guides camera motion ${}^c X_3$;

$F_4 = L_1/L_2$, ratio of lengths of two opposite sides of the quadrangle about the axis, which passes through the center of gravity of the quadrangle and is parallel to the V axis in the image plane. F_4 directs camera motion ${}^c X_4$;

$F_5 = L_3/L_4$, ratio of lengths of two opposite sides of the quadrangle about the axis, which passes through the center of gravity of the quadrangle and is parallel to the U axis in the image plane. F_5 directs camera motion ${}^c X_5$, and

$F_6 = \text{angle between the } U \text{ axis and the axis of minimum moment of inertia of the quadrangle in the image plane. It guides camera motion } {}^c X_6$.

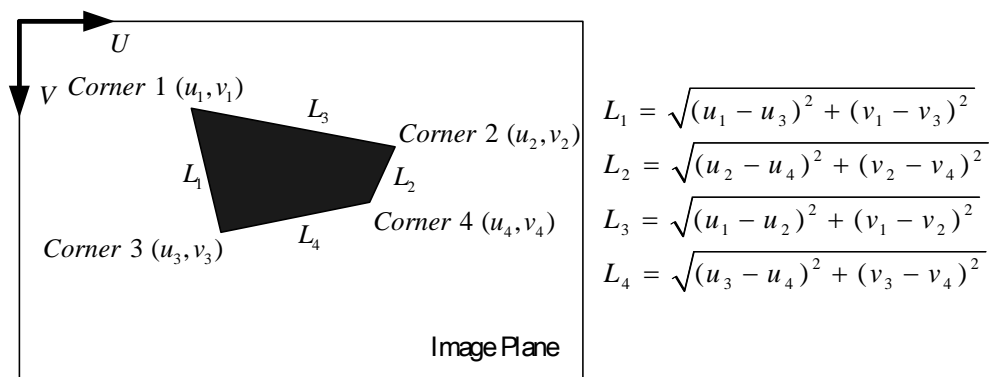


Fig. 4. Quadrangle in the workpiece image.

Each δF_i , for $i=1, 2, \dots, 6$, is defined as the difference between F_i and the feature value of the reference image at the target location, F_i' . F_i' is the value of F_i measured by the teach-by-showing method. This reference image is captured by the vision system when the end-effector is driven by a teaching box to the target location, ${}^o T$, relative to the object frame, as shown in Fig. 3. The target location is defined as follows. [The end-effector is initially driven by a teaching box to a location that allows the gripper to grasp the workpiece. Then, the

end-effector is driven to another location, which is a safe distance from the preceding location (in this case, 10 cm above it). This "another location" is called the "target location".] The reference features that correspond to the reference image at the target location are F_1^r , F_2^r , F_3^r , F_4^r , F_5^r and F_6^r .

3.2 Motion Planning Based on Behavior Design

In this work, fuzzy rules are used to enable the controllers to map image features in the image space onto motion commands in the camera space. These are then transformed to the commands in relation to the end-effector frame, which eventually control the manipulator. The final control values are relative motion commands sent to the position controller of the manipulator. The primary motivation for this process is that, after the notion of the camera frame has been introduced, the control rules for implementing basic vision-based motions such as "Center" or "Zoom" can be very easily written. No analytical model of the system is required.

1) Approach and Surround

The complete manipulation task involved in implementing a human-like visual servoing method is first divided into two complex behaviors - Approach and Surround - and one basic operation, Catch. The Approach behavior is the translational motion of the camera toward the workpiece, which is further divided into two basic behaviors - Center and Zoom. The Surround behavior is the orientational motion of the camera to keep the workpiece in the gripper, and is further divided into three basic behaviors - Yaw, Pitch and Roll. The Catch is a non vision-based operation. Only when the end-effector has reached the target location is the Catch activated and moves the end-effector 10 cm forward. The gripper then closes to grasp the workpiece. Fig. 5 displays the hierarchical composition of the behaviors, which are defined as follows.

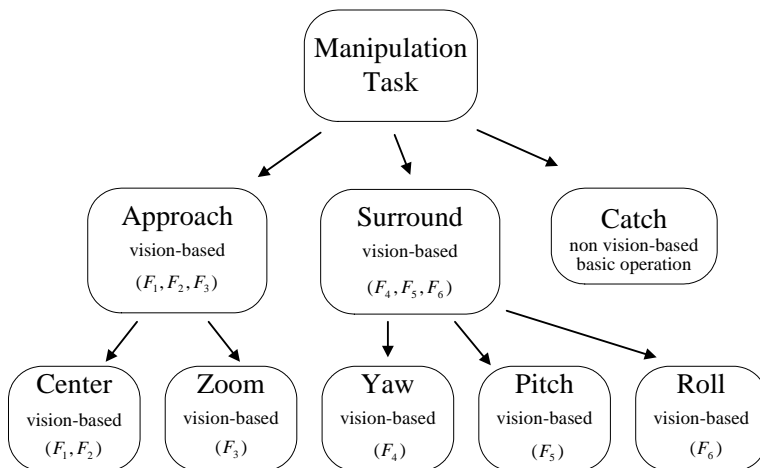


Fig. 5. Hierarchical composition of the behaviors.

Center is based on the first two image features, F_1 and F_2 . This behavior translates the camera along the cX and cY axes of the camera frame to keep the center of gravity of the quadrangular image at the desired pixel in the image plane.

Zoom is based on F_3 ; it moves the camera along the cZ axis of the camera frame to keep the size of the object as a predefined value.

Yaw is based on F_4 ; it rotates the camera about cX to keep the ratio of the lengths of the two short sides equal to F_4^r .

Pitch is based on F_5 ; it rotates the camera about cY to keep the ratio of the lengths of the two long sides equal to F_5^r .

Roll is based on F_6 ; it rotates the camera about cZ so that the principal angle equals that in the reference image, in which the gripper's two fingers are arranged parallel to the short sides of target.

Vision-based behaviors are defined from the perspective of an eye-in-hand camera, so movements are performed relative to the camera frame.

2) Neural Fuzzy Controller

The main shortcoming of traditional fuzzy controllers is that they are unable to learn. The best control law or membership functions can be determined by experience. However, the manipulation tasks are non-linear and coupled. None set of membership functions is good for the entire work environment. With respect to learning capacity of the artificial neural network, back-propagation architecture is the most popular and effective for solving complex and ill-defined problems. Therefore, six simple neural fuzzy controllers (NFCs), employing back-propagation (Kim et al., 1995; Nomura et al., 1992) are designed herein. One image feature is input to each controller, which changes one D.O.F. of the camera motion as the output. The back-propagation algorithm is used only to adjust the consequents of fuzzy rules for each neural fuzzy controller at each iteration during the manipulation. Restated, the camera is guided intuitively according to the image features on the image plane. For example, if the area of the quadrangular image is smaller than that in the reference image, the camera appears to be far from the workpiece, and so the camera is moved forward. Otherwise, the camera is moved backward. The other five D.O.F. of motion of the camera are controlled in the same manner.

Fuzzy singleton rules are adopted to simplify the neural fuzzy controllers; they are defined as follows;

$$\text{if } \delta F_i \text{ is } A_i^j, \text{ then } \delta^c X_i^* \text{ is } w_i^j \quad (1)$$

where input variable δF_i is an image feature error; output variable $\delta^c X_i^*$ denotes a relative motion command in the camera frame; A_i^j are linguistic terms of the precondition part with membership functions $\mu_{A_i^j}(\delta F_i)$, and w_i^j represent the real numbers of the consequent part, $i=1, 2, \dots, 6$ and $j=1, 2, \dots, 7$. That is, each $i=1, 2, \dots, 6$, can be regarded as a neural fuzzy controller with seven rules to control one D.O.F. of motion relative to the camera frame. Such a neural fuzzy system has a network structure as presented in Fig. 6. Herein, a simplified defuzzifier is used. The final output $\delta^c X_i$ of the neural fuzzy system is calculated by,

$$\delta^c X_i = \sum_{j=1}^7 \mu_{A_i^j} w_i^j \quad (2)$$

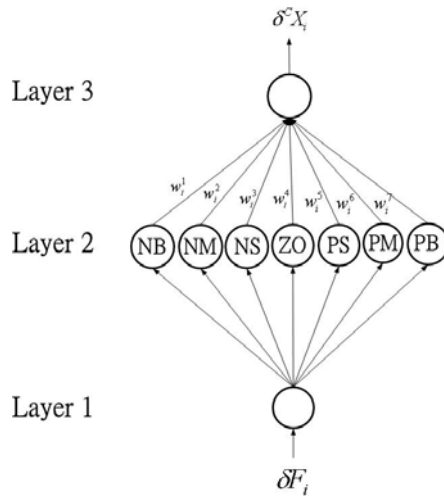


Fig. 6. Structure of the neural fuzzy controller with fuzzy singleton rules.

The parameter learning of NFC with fuzzy singleton rules can be the tuning of the real numbers w_i^j and the input Gaussian membership functions $\mu_{A_j}(\delta F_i)$, including the mean-points and the standard deviations (Lin & Lee, 1996). In this investigation, the mean-points and the standard deviations of the input membership functions are fixed to simplify parameter tuning. Only real numbers w_i^j are tuned on-line. Accordingly, the error function to be minimized is defined by

$$E_i = \frac{1}{2}(F_i^r - F_i)^2 = \frac{1}{2}(\delta F_i)^2 \tag{3}$$

The derivative of an error function E_i with respect to the j -th consequent can be obtained as

$$\frac{\partial E_i}{\partial w_i^j} = \frac{\partial E_i}{\partial(\delta^c X_i)} \frac{\partial(\delta^c X_i)}{\partial w_i^j} \tag{4}$$

However, unfortunately, determining $\partial E_i / \partial(\delta^c X_i)$ is difficult since the dynamics should be taken into account to see how the change in control input influences E_i . Trickily, however, $\partial E_i / \partial(\delta^c X_i)$ can be approximately computed from the difference ratio.

$$\frac{\partial E_i}{\partial(\delta^c X_i)} \approx \frac{dE_i}{d(\delta^c X_i)} \approx \frac{E_i(t) - E_i(t-1)}{\delta^c X_i(t) - \delta^c X_i(t-1)} \tag{5}$$

To decrease E_i with respect to w_i^j , the consequent changes at time stage t , and $\Delta w_i^j(t)$ can be chosen to be proportional to $-\partial E_i / \partial w_i^j$.

$$\Delta w_i^j(t) = -\eta \frac{\partial E_i}{\partial w_i^j} \tag{6}$$

where η is the learning-rate parameter. Hence, a learning rule for adapting the consequent at time stage t can be given as

$$w_i^j(t+1) = w_i^j(t) + \Delta w_i^j(t) \quad (7)$$

3.3 Rough Motion Transformation

In a manufacturing environment, hand-eye calibration process is generally time consuming. In this study, the pose of the camera in relation to the end-effector is invariant, as shown in Fig. 7, so the camera and the end-effector can be treated as a rigid body. After the motion of the rigid body has been analyzed, the transformation from the output values in relation to the camera frame to the motion commands with respect to the end-effector frame is obtained.

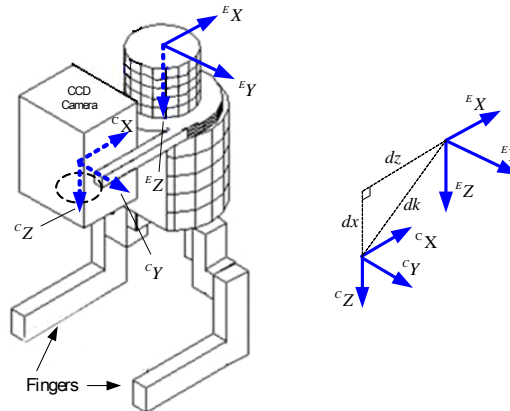


Fig. 7. End-effector and camera frames.

Fig. 8 reveals that if the controller output is only a rotation $\delta^c X_6$ about $^c Z$ but the command sent to the manipulator is a rotation about $^e Z$, then the unexpected camera displacements in the $^c X$ and $^c Y$ directions are $\delta^c X_1^+$ and $\delta^c X_2^+$, respectively. Similar situations occur in $\delta^c X_4$ and $\delta^c X_5$. Therefore, the motion command sent to the manipulator should be transformed as,

$$\begin{aligned} \delta^e X_1 &= \delta^c X_1 - dz \cdot \sin(\delta^c X_6) \cdot \tan(\delta^c X_6/2) - \\ &\quad 2dk \cdot \sin(\delta^c X_5/2) \cdot \cos(90^\circ - \delta^c X_5/2 - \tan^{-1}(dx/dz)) \\ \delta^e X_2 &= \delta^c X_2 + dz \cdot \sin(\delta^c X_6) + dx \cdot \sin(\delta^c X_4) \\ \delta^e X_3 &= \delta^c X_3 + dx \cdot \sin(\delta^c X_4) \cdot \tan(\delta^c X_4/2) - \\ &\quad 2dk \cdot \sin(\delta^c X_5/2) \cdot \sin(90^\circ - \delta^c X_5/2 - \tan^{-1}(dx/dz)) \\ \delta^e X_4 &= \delta^c X_4 \\ \delta^e X_5 &= \delta^c X_5 \\ \delta^e X_6 &= \delta^c X_6 \end{aligned} \quad (8)$$

where dx represents the distance between axes $^c X$ and $^e X$, and dz is the distance between axes $^c Z$ and $^e Z$. The two lines associated with dx and dz are assumed to be mutually perpendicular. They are measured roughly using a ruler and the naked eye. Accordingly,

$dk = \sqrt{(dx)^2 + (dz)^2}$ is assumed. Apparently, this hand-eye configuration is inaccurate. However, the designed neural fuzzy controllers handle the inaccuracy by tuning the consequents of the fuzzy rules according to the back-propagation algorithm. This process saves considerable time without the intensive computation associated with hand-eye calibration.

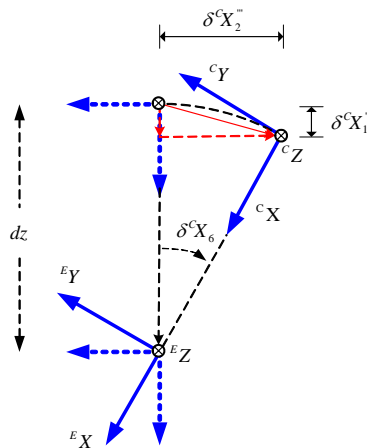


Fig. 8. Unexpected camera displacement caused by a rotation about ${}^E Z$.

3.4 Control Strategy

When the camera is far from the workpiece, the image features F_4 and F_5 are unstable because they are sensitive to the corner detection in image processing. Consequently, the Yaw and Pitch behaviors are activated only when the camera is close to the workpiece. On the contrary, the influence of distance and illumination on F_1 , F_2 , F_3 and F_6 is negligible.

Given the above restriction, the control strategy is as follows. Initially, the Approach behavior and the Roll behavior occur simultaneously in the approaching stage, in which the camera is moved toward the target. This process is iteratively performed until $|\delta F_1|$, $|\delta F_2|$, $|\delta F_3|$ and $|\delta F_6|$ are below the specified limiting values, ε_1 , ε_2 , ε_3 and ε_6 , respectively. The fine positioning strategy is then activated. In the fine positioning stage, Yaw and Pitch run concurrently to adjust the orientation of the camera. Approach and Roll are then executed again. These two processes iteratively run in turns until $|\delta F_i|$ is less than ε_i , $i=1, 2, \dots, 6$. Finally, the basic operation Catch is inspired. Fig. 9 shows the behavior-based look-and-move control structure with rough motion transformation. Fig. 10 presents the control strategy flowchart.

For each step of motion, the motion command relative to the camera frame is transformed to a command relative to the end-effector frame using the proposed rough motion transformation. The inaccuracy of the hand-eye relationship can be neglected by adjusting on-line the singletons of the consequent parts in the fuzzy rules. The adjustment is implemented by the back-propagation algorithm that exploits a gradient descent method to reduce image feature errors.

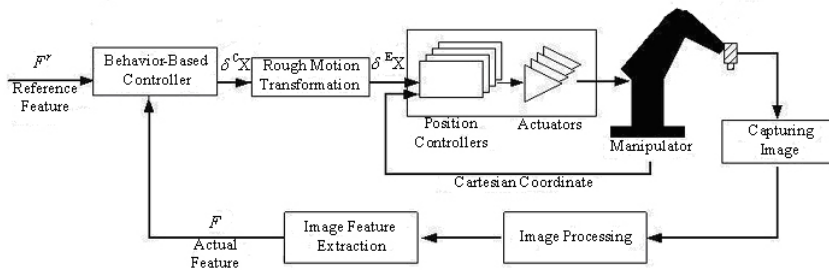


Fig. 9. Behavior-based look-and-move control structure.

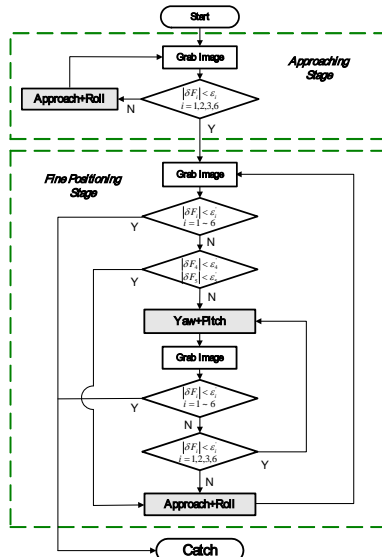


Fig. 10. Control strategy flowchart for manipulation.

4. Experimentation

4.1 Experimental Setup

As shown in Fig. 7, a CCD camera is mounted on the end-effector of the manipulator in such a way that each coordinate axis in the camera frame is approximately parallel to that in the end-effector frame. A ruler and the naked eyes reveal that the distance between ${}^c X$ and ${}^E X$, dx , is about 73 mm. dz is also estimated to be around 80 mm. Therefore, $dk = \sqrt{(dx)^2 + (dz)^2}$ is approximately 107 mm. The end-effector employed to grasp the workpiece is a two-finger gripper with parallel motion of the fingers, and is attached to the end of the manipulator. The gripper's fingers are installed approximately parallel to the ${}^E X$ axis of the end-effector frame. The gripper is always open, except when the end-effector is commanded to grasp the workpiece from the side of the workpiece. The manipulator's end-effector will pick up a workpiece that is a rectangular parallelepiped whose volume is $68\text{mm} \times 30\text{mm} \times 55\text{mm}$. The zoom of the camera is adjusted so that

horizontal angle of view is around 39° . However, gripper's fingers are still not in the field of vision. An auto-focus function is exploited to maintain the sharpness of the workpiece image. The mobile base is stopped and fixed next to the workstation, which is roughly parallel to the surface of the workstation to verify the proposed control strategy. As presented in Fig. 11, the workpiece is placed in six different positions, which are separated by 15 cm, to simulate the possible position and orientation errors that arise in the application stage. In each position, the workpiece is pointed in three directions. It is placed at Pos2 and Pos5 in 0° and -45° directions, and is tilted by 3° and 6° to the station surface to simulate non-flat ground in the application stage. Before the manipulation is performed, one reference image is captured by applying the teach-by-showing method. The end-effector is first driven by a teaching box to a location so that the gripper can grasp the workpiece. Then, the end-effector is driven 10 cm above by a teaching box along the negative- Z axis to the target location. Notice that the target location of the end-effector in reference with the object frame is always fixed and is expected to reach. Then, the corresponding reference image features are extracted. Given the uncalibrated CCD camera and the imprecisely known hand-eye configuration, (F'_1, F'_2) does not necessarily correspond to the center pixel of the image plane, and F'_6 does not always equal zero degree. All of these three feature values depend on the captured reference image. However, the reference image feature F'_3 , which presents the distance between the camera and the workpiece, is set to zero at all times. The other two feature values, F'_4 and F'_5 , depend on the captured reference image.

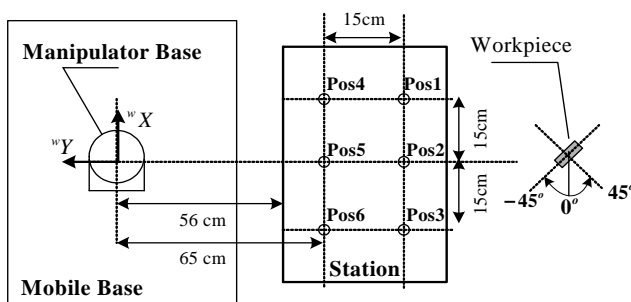


Fig. 11. Possible locations of the workpiece to be picked up.

The parameters of the above experimental setup to evaluate the positioning performance of the eye-in-hand manipulator are set as follows. Table 1 lists the initial parameters of the fuzzy membership functions, including the mean-points and the standard deviations of the Gaussian curves, and the learning rate of each DOF control in the designed neural fuzzy controllers. The developed control strategy is divided into two stages. The approaching stage continues until the errors in the image features δF_1 , δF_2 , δF_3 and δF_6 are below $\varepsilon_1 = 6$, $\varepsilon_2 = 6$, $\varepsilon_3 = 0.1$ and $\varepsilon_6 = 2$, respectively. In neural fuzzy controllers, the corresponding energy values E_1 , E_2 , E_3 and E_6 are less than 18, 18, 0.005 and 2, respectively. As to the following fine positioning stage, it never terminates unless the errors in the image features δF_1 , δF_2 , δF_3 , δF_4 , δF_5 and δF_6 are less than the specified values $\varepsilon'_1 = 2$, $\varepsilon'_2 = 2$, $\varepsilon'_3 = 0.01$, $\varepsilon'_4 = 0.01$, $\varepsilon'_5 = 0.01$ and $\varepsilon'_6 = 0$. The corresponding energy values $E_1 \sim E_6$ are below 2, 2, 0.00005, 0.00005, 0.00005 and 0, respectively.

| Rule | Rule Parameter | NB (j=1) | NM (j=2) | NS (j=3) | ZO (j=4) | PS (j=5) | PM (j=6) | PB (j=7) | η_i |
|------------------|---------------------|----------|----------|----------|----------|----------|----------|----------|----------|
| R_1^j (i=1) | C_1^j pixel | -90 | -60 | -30 | 0 | 30 | 60 | 90 | 0.0005 |
| | σ_1^j pixel | 11.339 | 11.339 | 11.339 | 11.339 | 11.339 | 11.339 | 11.339 | |
| | w_1^j mm | -50 | -10 | -3 | 0 | 3 | 10 | 50 | |
| R_2^j (i=2) | C_2^j pixel | -90 | -60 | -30 | 0 | 30 | 60 | 90 | 0.0005 |
| | σ_2^j pixel | 11.339 | 11.339 | 11.339 | 11.339 | 11.339 | 11.339 | 11.339 | |
| | w_2^j mm | -50 | -10 | -3 | 0 | 3 | 10 | 50 | |
| R_3^j (i=3) | C_3^j | -0.75 | -0.5 | -0.25 | 0 | 0.25 | 0.5 | 0.75 | 500 |
| | w_3^j | 0.0945 | 0.0945 | 0.0945 | 0.0945 | 0.0945 | 0.0945 | 0.0945 | |
| | w_3^j mm | -150 | -80 | -25 | 0 | 25 | 80 | 150 | |
| R_4^j (i=4) | C_4^j | -0.0276 | -0.0207 | -0.0138 | 0 | 0.0138 | 0.0207 | 0.0276 | 0.05 |
| | σ_4^j | 0.0026 | 0.0026 | 0.0026 | 0.0026 | 0.0026 | 0.0026 | 0.0026 | |
| | w_4^j degree | -1.5 | -1 | -0.5 | 0 | 0.5 | 1 | 1.5 | |
| R_5^j (i=5) | C_5^j | -0.0081 | -0.0054 | -0.0027 | 0 | 0.0027 | 0.0054 | 0.0081 | 0.05 |
| | σ_5^j | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | |
| | w_5^j degree | -1.5 | -1 | -0.5 | 0 | 0.5 | 1 | 1.5 | |
| R_6^j (i=6) | C_6^j degree | -90 | -60 | -30 | 0 | 30 | 60 | 90 | 0.01 |
| | σ_6^j degree | 11.339 | 11.339 | 11.339 | 11.339 | 11.339 | 11.339 | 11.339 | |
| | w_6^j degree | -90 | -60 | -30 | 0 | 30 | 60 | 90 | |

Table 1. Initial parameters of neural fuzzy controllers.

4.2 Results

In experiments performed to evaluate the positioning performance of the eye-in-hand manipulator, the end-effector of the manipulator is firstly driven to the top location, presented as ${}^w[-80, -750, 470, 180, 0, 90]$ in the world frame. Then, the end-effector is visually guided to grasp the workpiece according to the proposed control strategy with the preset parameters. Fig. 12 displays the images in the course of approaching the workpiece.

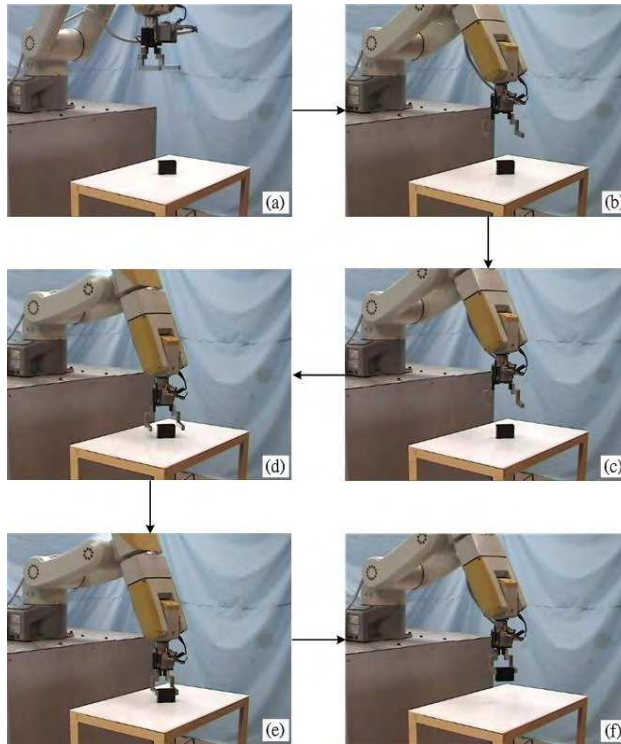


Fig. 12. (a)~(f) Images in the course of approaching the workpiece, based on the proposed control strategy; (a)~(b) approaching stage; (b)~(c) fine positioning stage; (d)~(f) catch operation.

The position and orientation errors are defined below to describe the positioning performance of the manipulator numerically. The location of the end-effector in each step, T_o^{eStep} , is recorded during the task manipulation. The coordinate transformation matrix between each step and desired locations of the end-effector can be written as follows.

$$T_{eDesired}^{eStep} = T_o^{eStep} (T_o^{eDesired})^{-1} = \begin{bmatrix} R_{eDesired}^{eStep} & t_{eDesired}^{eStep} \\ 0 & 1 \end{bmatrix}$$

$$\Rightarrow T_{eDesired}^{eStep} = \begin{bmatrix} n_x & t_x & b_x & X \\ n_y & t_y & b_y & Y \\ n_z & t_z & b_z & Z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (9)$$

where T_o^{eStep} corresponds to the location of the end-effector with respect to the world frame in each step, and $T_o^{eDesired}$ corresponds to the desired location for the end-effector with respect to the world frame. In each test, (10) gives the position error and (11) specifies the orientation error as the angle of rotation between the two coordinate frames about the principle axis.

$$Error_{position} = \sqrt{X^2 + Y^2 + Z^2} \quad (10)$$

$$Error_{orientation} = \tan^{-1} \left(\frac{\sqrt{(t_z - b_y)^2 + (b_x - n_z)^2 + (n_y - t_x)^2}}{n_x + t_y + b_z - 1} \right) \quad (11)$$

Table 2 presents the resulting positioning errors and the number of execution steps in each test. The number of steps is that required for the end-effector to travel from the top location to the target location. Executing this behavior-based control strategy takes a minimum of around 11 steps and a maximum of 17 steps in cases in which the workpiece is arranged without a tilt. The initial orientation of the workpiece only weakly influences the total execution time. In the cases in which the tilt angle is 3° or 6°, about 22 steps to 35 steps are required. In all of the tests, the final position error is less than 2.9 mm, and the final orientation error is less than 1.5°. Additionally, as the tilt angles are increased, more steps are needed for positioning the end-effector.

| Workpiece Locations | $Error_{Position}$ (mm) | $Error_{Orientation}$ (degree) | Number of Steps |
|------------------------------|-------------------------|--------------------------------|-----------------|
| Pos1 No Rotation | 1.0531 | 0.1384 | 14 |
| Pos1 Rotation 45° | 0.7573 | 0.1909 | 16 |
| Pos1 Rotation -45° | 1.7142 | 0.2180 | 11 |
| Pos2 No Rotation | 1.0337 | 0.1160 | 13 |
| Pos2 Rotation 45° | 0.5197 | 0.0767 | 15 |
| Pos2 Rotation -45° | 1.1542 | 0.0956 | 13 |
| Pos3 No Rotation | 1.2159 | 0.1732 | 16 |
| Pos3 Rotation 45° | 0.9185 | 0.1447 | 12 |
| Pos3 Rotation -45° | 1.0466 | 0.5370 | 12 |
| Pos4 No Rotation | 0.7296 | 0.0600 | 16 |
| Pos4 Rotation 45° | 1.4728 | 0.3718 | 14 |
| Pos4 Rotation -45° | 2.8471 | 0.5503 | 13 |
| Pos5 No Rotation | 0.6244 | 0.0591 | 16 |
| Pos5 Rotation 45° | 0.8570 | 1.1380 | 14 |
| Pos5 Rotation -45° | 1.5124 | 0.1721 | 16 |
| Pos6 No Rotation | 0.7469 | 0.0516 | 16 |
| Pos6 Rotation 45° | 1.4360 | 0.4218 | 13 |
| Pos6 Rotation -45° | 0.8015 | 0.0649 | 17 |
| Pos2 No Rotation Tilted 3° | 1.6559 | 1.1362 | 22 |
| Pos2 No Rotation Tilted 6° | 1.5903 | 0.3740 | 32 |
| Pos2 Rotation -45° Tilted 3° | 0.2003 | 0.1174 | 25 |
| Pos2 Rotation -45° Tilted 6° | 1.8498 | 0.3057 | 35 |
| Pos5 No Rotation Tilted 3° | 2.4739 | 0.4196 | 24 |
| Pos5 No Rotation Tilted 6° | 2.4876 | 0.3764 | 29 |
| Pos5 Rotation -45° Tilted 3° | 2.4000 | 1.4623 | 24 |
| Pos5 Rotation -45° Tilted 6° | 2.2709 | 0.3534 | 31 |

Table 2. Positioning errors of the eye-in-hand manipulator and the number of steps in each test.

4.3 Discussion

Although these experiments were performed with the mobile base fixed on the ground such that the base of the manipulator was approximately parallel to the surface of station, the workpiece was placed in six positions with various orientations and two tilt angles. These experimental data simulate the possible locations of the workpiece in relation to the base of the manipulator, when the mobile manipulator is driven to the station using a guidance control system of the mobile base in the application stage. Notably, a workpiece placed in various positions is used to simulate the position errors of the mobile base in the traveling and lateral directions. The workpiece points in different directions to simulate the orientation errors of the mobile base, and the setups with the different tilt angles of the workpiece are used to simulate the non-horizontality of the ground.

In the application stage, if the workpiece is placed in the middle position between Pos2 and Pos5 and pointed in the 0° direction, then the manipulator base is supposed to be driven by the mobile base to stop next to it. The experimental results confirm that the eye-in-hand manipulator can still pick up the workpiece, even when the position errors of the mobile base in the direction of travel and in the lateral direction are $\pm 150\text{mm}$ and $\pm 75\text{mm}$, respectively. Moreover, the manipulator can tolerate a workpiece rotation of $\pm 45^\circ$ and a non-horizontality angle of the ground of 6° .

In the other experiments, the workpiece is shifted by manually pushing the workstation in the approaching stage to test the advantages of the reactive arbitration. Nevertheless, the end-effector can correct its path to reach the desired pose without a loss, and pick up the workpiece. Furthermore, in all of the experimental tests, the workpiece was picked up successfully. Therefore, the probability of successful grasping was about 100%.

5. Conclusion

Mobile robots commonly need a guidance control system to navigate the mobile base and a method for performing the pick-and-place operations. The guidance control system and the non-planar ground inevitably cause the position and orientation errors of the mobile base. Therefore, this study proposes a behavior-based control strategy that employs an uncalibrated eye-in-hand vision system to control the end-effector of the manipulator to approach and grasp the target workpiece. All the designed behaviors are defined from the perspective of the camera and are mediated through fuzzy rules with a look-and-move control structure. The presented neural fuzzy controllers map image features in image space to relative motion commands in the camera frame. These motion commands are then transformed to the end-effector frame by the proposed rough motion transformation.

This work differs from the references (Wasik & Saffiotti, 2002, 2003) as follows. (1) A back-propagation algorithm is used to reduce image feature errors through the adjustment of the singletons of the consequent parts in the fuzzy singleton rules. (2) Two additional image features, the ratios of the lengths of the two pairs of opposite sides on the quadrangular image, are extracted to guide the rotation of the camera when the workpiece has a tilt angle. (3) The control values defined in the camera frame generated from the controller in (Wasik & Saffiotti, 2003) are transformed to joint controls by a Jacobian transformation, revealing that either a hand-eye calibration process has been implemented or the hand-eye relationship is known beforehand. However, this study proposes a rough motion transformation to replace the time-consuming hand-eye calibration. Only a general ruler and naked eye are required to estimate hand-eye configuration. The inaccuracy of this estimation can be ignored by adjusting on-line the singletons of the consequent parts in the fuzzy singleton rules. (4) A more human-like control strategy is

planned. The end-effector approaches the workpiece and the image is simultaneously maintained in the reference pixel of the image plane. The path from the initial location of the end-effector to the location where the workpiece can be grasped is shortened.

6. References

- Anglani, A.; Taurisano, F.; De Giuseppe, R. & Distanto, C. (1999). Learning to Grasp by using Visual Information, *Proceedings of IEEE International Conference on Computational Intelligence in Robotics and Automation*, pp. 7-14, ISBN: 0-7803-5806-6, Monterey, CA, USA, Nov. 1999.
- Bien, Z. & Park, J.. (1993). Hybrid Fuzzy Self-organizing Controller for Visual Tracking, In: *Fuzzy Logic: State of the Art*, Lowen, R & Roubens, M., (Ed.), page numbers (569-578), Kluwer Academic Publishers, ISBN: 0-7923-2324-6, Netherlands.
- Hutchinson, S.; Hager, D. G. & Corke, P. I. (1996). A Tutorial on Visual Servo Control. *IEEE Transactions on Robotics and Automation*, Vol. 12, No. 5, (Oct. 1996) page numbers (651-670), ISSN: 1042-296X.
- Kim, C. S.; Seo, W. H.; Han, S. H. & Khatib, O. (2001). Fuzzy Logic Control of a Robot Manipulator Based on Visual Servoing, *Proceedings of IEEE International Symposium on Industrial Electronics*, pp. 1597-1602, ISBN: 0-7803-7090-2, Pusan, Korea, June 2001.
- Kim, J. G.; Cha, D. H.; Cho, K. S. & Kim, S. H. (1995). An Auto Tuning Fuzzy Rule-Based Visual Servoing Algorithm for a Slave Arm, *Proceedings of IEEE International Symposium on Intelligent Control*, pp. 177-182, ISBN: 0-7803-2722-5, Monterey, CA, USA, Aug. 1995.
- Lin, C. T. & Lee, C. S. G. (1996). *Neural Fuzzy Systems: A Neural-Fuzzy Synergism to Intelligent Systems*, Prentice-Hall, Inc., ISBN: 0-13-261413-8, Upper Saddle River, NJ, U.S.A.
- Nomura, H.; Hayashi, I. & Wakami, N. (1992). A Learning Method of Fuzzy Inference Rules by Descent Method, *Proceedings of IEEE International Conference on Fuzzy Systems*, pp. 203-210, ISBN: 0-7803-0236-2, San Diego, CA, USA, Mar. 1992.
- Suh, I. H. & Kim, T. W. (1994). Fuzzy Membership Function Based Neural Networks with Applications to the Visual Servoing of robot Manipulators. *IEEE Transactions on Fuzzy Systems*, Vol. 2, No. 3, (Aug. 1994) page numbers (203-220), ISSN: 1063-6706.
- Ting, Y.; Chen, Y. H.; Lin, M.; Dai, S. C. & Kang, Y. (1997). A Study on the Inaccuracy of Vision System of Mobile Robots Causing the Failure of Pick-and-place Tasks, *Proceedings of 1997 Florida Conference on Recent Advances in Robotics*, pp. 43-46, Miami, FL, USA, Apr. 1997.
- Wasik, Z. & Saffiotti, A. (2002). A Fuzzy Behavior-Based Control System for Manipulation, *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1596-1601, ISBN: 0-7803-7398-7, EPFL, Switzerland, Oct. 2002.
- Wasik, Z. & Saffiotti, A. (2003). A Hierarchical Behavior-Based Approach to Manipulation Tasks, *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 2780-2785, ISBN: 0-7803-7736-2, Taipei, Taiwan, Sept. 2003.



Mobile Robotics, Moving Intelligence

Edited by Jonas Buchli

ISBN 3-86611-284-X

Hard cover, 586 pages

Publisher Pro Literatur Verlag, Germany / ARS, Austria

Published online 01, December, 2006

Published in print edition December, 2006

This book covers many aspects of the exciting research in mobile robotics. It deals with different aspects of the control problem, especially also under uncertainty and faults. Mechanical design issues are discussed along with new sensor and actuator concepts. Games like soccer are a good example which comprise many of the aforementioned challenges in a single comprehensive and in the same time entertaining framework. Thus, the book comprises contributions dealing with aspects of the Robotcup competition. The reader will get a feel how the problems cover virtually all engineering disciplines ranging from theoretical research to very application specific work. In addition interesting problems for physics and mathematics arises out of such research. We hope this book will be an inspiring source of knowledge and ideas, stimulating further research in this exciting field. The promises and possible benefits of such efforts are manifold, they range from new transportation systems, intelligent cars to flexible assistants in factories and construction sites, over service robot which assist and support us in daily live, all the way to the possibility for efficient help for impaired and advances in prosthetics.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

T. I. James Tsay and Y. F. Lai (2006). Intelligent Pose Control of Mobile Robots Using an Uncalibrated Eye-in-Hand Vision System, Mobile Robotics, Moving Intelligence, Jonas Buchli (Ed.), ISBN: 3-86611-284-X, InTech, Available from:

http://www.intechopen.com/books/mobile_robotics_moving_intelligence/intelligent_pose_control_of_mobile_robots_using_an_uncalibrated_eye-in-hand_vision_system

INTECH

open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2006 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.