International Journal of Automation Mechatronics & Robotics – IJAMR
Volume 1 : Issue 1 [ISSN 2374 – 1546]
Publication Date : 09 January 2014

Design and Implementation of Vision-Guiding Controller for Autonomous Cart

Ying-Shing Shiao, Po-Hsiang Yu, Ching-Wei Wu, Chieh-Wen Liao

Abstract*—This paper studies the image processing algorithms to recognize the face features and hands gestures for the motion control of an autonomous cart. The camera installed on a cart is used to grab the images of its task scene, from which to recognize the commander and his or her hands gestures, which is combined with the cart motions to determine the distances and directions of the target for control cart directions and speeds to achieve the purpose of image navigation. The experimental results show the proposed approaches make the cart with the abilities to interactive with its commander. *(Abstract)

***Keywords*—face identification, gesture recognition, cart navigation.**

I. Introduction *(Heading 1)*

The visual servo method was proposed by Hill and Park. In their study, the images grabbed by cameras are treated as human eyes processing the non-contact environment measuring [1]. The robot get images which provides environmental information as the motion references of navigation or moving planning to get the autonomous control.

A cart is combined with machine vision and motion control by image processing to figure out the specific targets, calculate the target position for trajectory control. These image identification methods get the implementation of a cart with visual guidance [2].

Many studies explored image processing methods for face recognition. The conventional methods of face recognition are based on the face features [3,4] and the skin color identification [5,6]. M. Turk and A. Pentland adopted the principal component analysis (PCA) for face alignment [7]. A two complementary face detection strategies based on color was used in [8]. In the first strategy, color information is used to display facial feature pattern and the geometry constraints is use to filter out the specific pattern. In second strategy, an adjustable window searches the possible face candidates.

Eyes and mouth are the specific features of a face because of the nature of unique geometry. The feature of eyes and mouth is an isosceles triangle. The mouth area, pixels, is larger than the eye's. Apply the statistical data of these features to be the constraints; the candidate face can be determined. In addition to, the rotated face image and the different face expression have to be considered as the feature variations for validation [8, 9]. In case of features insufficient cause that the skin color-based face feature extraction method is unable to obtain the candidate faces. It is evidence that the face postures are changed drastically. This problem is solved by an adjustable search window. This method uses binary processing for split face color blocks. Search window is used to distinguish a face or other objects with similar face color, and then the PCA used for construct the basis of features for face detection [10].

Besides, there are relevant methods for face feature extraction. S. Y. Lee, Y. K. Ham, and R. H. Park according to the knowledge-based method proposed the feature extraction method for finding out the face contour, and using Neuro-Fuzzy algorithm for confirming the face recognition results [11]. C. C. Han, H. Y. Mark Liao, G. J. Yu, and L. H. Chen proposed morphology-based technology to recognize the eyes image for identifying the candidate faces, and then the back-propagation neural network is used for confirm [12]. In some studies, the contour shapes with specific parameters such as circular, polygonal, elliptical, etc. are treated as the shape features to create the template for making out the gestures [13, 14].

In this paper, an image processing algorithm is proposed for recognizing the face and gestures of a commander to identify its destination which is combined with cart motion and path planning, so as to navigate a cart. Consequently, the cart is equipped with a camera to achieve visual servo control, in which images used for object identification and motion path planning to implement an autonomous wheeled robot.

II. Vision Guiding

The proposed image processing algorithms for guiding a cart is shown in Fig. 1. The camera on cart grabs a scene images used for recognition commander and gesture. The results of face recognition are used to determine the commander's position. The coordinate of a commander is the reference for direction control. The commander's gesture is the destination of the cart. The cart equipped a camera which grab images are treated for motion control to get the ability of interactive with commander.

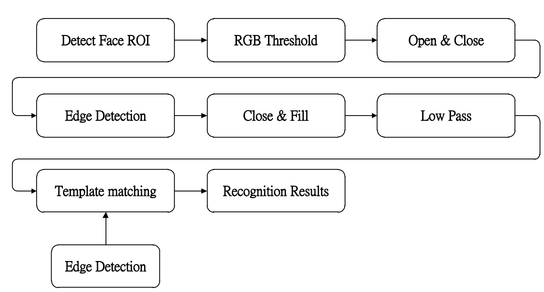


Figure. 1 Face recognition

The commander's face recognition algorithm is shown in Fig. 1. The face recognition based on skin color is used to select a ROI including eyes image. A RGB threshold

Ying-Shing Shiao, Po-Hsiang Yu, Ching-Wei Wu, Chieh-Wen Liao
National Changhua University of Education
Taiwan
Gn03314903@hotmail.com

processing gets its image binary. The morphology open and close are used to repair and to emphasize the eyes pattern. The edge detection method is used to get the contour of eyes image. The morphological fill is used to fill the eyes contour. The low-pass filter is used to retain the eye pattern for comparing with the commander eyes template.

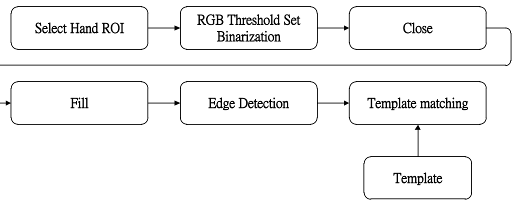


Figure.2 Gesture recognition

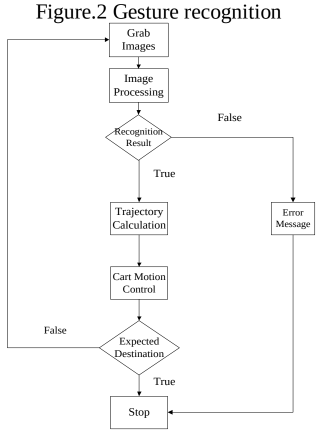


Figure. 3 The processing flow of a vision-guiding cart.

he image processing algorithm for gesture recognition is shown in Fig. 1. The methods used in the procedure of gesture recognition are same as the face recognition's. The significant difference is the hand contours used for gesture recognition and the eyes template used for commander recognition. The processing of gesture recognition is morphological fill followed edge detection. Contrariwise, the processing of commander recognition is edge detection follow morphological fill. First, the hand image is selected and after RGB threshold processing get hand binary image which follow the morphological close to remove the redundant. The morphological fill is used to close the contours in the hand image. By the geometric template matching, the correct gestures are confirmed.

The key method for recognition are the template matching and morphological operations that are respectively explained as the follows.

The template matching is used for face and gesture recognition that is illustrated in Fig. 4. The compared procedure of the template with an image is a convolution operation expressed as:

$$C(i,j)=\sum_{x=0}^{K-1}\sum_{y=0}^{L-1} w(x,y)\,f(x+i,y+j) \qquad (1)$$

Where the template size is $K\times L$, image size is $M\times N$, and $i=0,1.....M-1,\ j=0,1.....N-1$. The maximum of is the location of the template in the image. In accordance with this method, the position of the commander and his or her gesture in the image is determined.

A structure is chosen for erosion image. Not only the object smaller than the structure is eroded, but also the boundary pixels of the main objects are removed. Dilation is opposite to erosion. The objects in image are expanded by the structure. After dilation the space inside the objects or the gaps on its boundary smaller than structure are filled. The morphological erosion and dilation are used for open, close, and fill processing in Fig. 1.

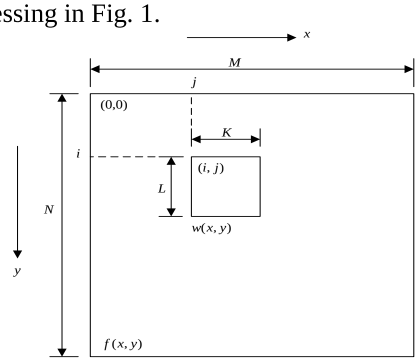


Figure. 4 Template matching.

The morphological open is combined with morphological dilation and erosion. The noise such as bright spots can be removed and the object boundary can make smoother by the morphological open. The morphological open operation for an image *I* is obtained by erosion follow dilation as :

$$\text{Opening}(\,I\,)=\text{dilation}(\,\text{erosion}(\,I\,)\,) \qquad (2)$$

The morphological close operation is obtained by dilation follow erosion that can fill the holes of object image and make the object more uniform.

$$\text{Closing}(\,I\,)=\text{erosion}(\,\text{dilation}(I)\,) \qquad (3)$$

Fill the holes are convenient for making objects with more specific features that reduce the processing burden for recognition. Only identify the commander and his or her gestures are not for controlling a cart. The camera model is needs to deal with the relationship of target and its image. The geometric optical of lens used for determination the coordinate of a stood commander and the destination at which the gestures point. The image of cart task scene to be identified obtains the commander location and gestures direction which are converted to the control parameters for cart motion control.

In order to identify the destination for a cart that using the pin hole model to obtain the depth information [18, 19]. The relationship of an object and its image are shown as:

$$x=f\cdot\frac{X}{Z} \qquad (4)$$

where is the focal length of a lens, is distance of an object in front of camera, is the object height. The methods of Fig. 1 used to identify the commander height (pixels) in the image. If the commander height is known, the distance between the commander the camera is known.

III. Cart Motion Control

The coordinate system for the cart moving control is shown in Fig. 5.

O_I and O_{car} are the origin of the world coordinate and the cart, respectively. ${}^{I}X_{car}$, ${}^{I}Y_{car}$ and θ are used to describe a cart in the coordinate system. The destination in world coordinates is expressed by $P_T(X_I,Y_I)$. The vision system measuring the destination in motion coordinate system is represented by $P_{car}(X_{car}, Y_{car})$. P_T and P_{car} is related by the following equation:

$$P_T = A \cdot P_{car} \tag{5}$$

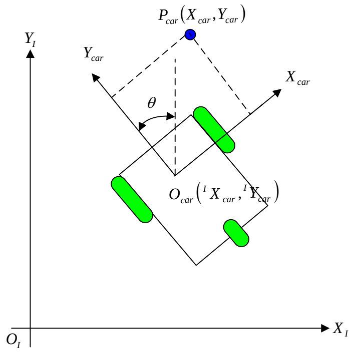


Figure. 5 The relationship of a cart and its destination.

where is the transformation matrix of world coordinate and cart coordinate. Using the cart origin in the world coordinates and angle between cart axis and world axis , then Eq.(5) can be represented by the follow:

$$\begin{bmatrix} X_I \\ Y_I \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & {}^I X_{car} \\ -\sin\theta & \cos\theta & {}^I Y_{car} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_{car} \\ Y_{car} \\ 1 \end{bmatrix} \tag{6}$$

Shown in Fig. 6, ϕ is angle between the destination and a cart coordinate axis Y_{car}. The direction controllable range is $-90° < \phi < 90°$, and the cart speed is limit for safety. In axis Y_{car}, the distance is the distance from the destination for cart keeps a safe distance D_{safe} in front of the destination. In axis X_{car}, dx is the distance between cart and it's destinations. The errors of dy, dx and ϕ are used to control cart motion to move to the destination. Finally, the cart and the destination will keep at a safe distance D_{safe}. In control, the cart forward to the destination, the moving distance dy and dx can be expressed by the follow:

$$dy = (Y_{car} - D_{safe}) \cdot \cos(d\phi) \text{ and } dx = X_{car} \tag{7}$$

To control the cart directions $d\phi$, the movement distance dy and dx is represented by cart coordinate. (6) is used for coordinate transformation shown as:

$$\begin{bmatrix} dx_I \\ dy_I \\ 1 \end{bmatrix} = \begin{bmatrix} \cos(d\phi) & \sin(d\phi) & {}^I X_{car} \\ -\sin(d\phi) & \cos(d\phi) & {}^I Y_{car} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} dx \\ dy \\ 1 \end{bmatrix} \tag{8}$$

Where dx_I, dy_I is cart moving distances represented by using coordinate X_I and Y_I.

The steering control of a cart is shown in Fig. 7. $F_r(x_r, y_r)$ is the center point of the rear wheel axle. This point locate in the centerline of cat body, point $F_r(x_r, y_r)$ is at the intersection of the front wheel axle and the body center line, angle θ between the rear wheel axle and body center line, angle $d\phi$ between the cart center line and the reference axis X_I, l is the distance between the front wheel and the rear wheel, radius r for the rear wheel turning, radius k for the front wheel turning, v_r is the rear wheel speed, and v is the speed of the cart. By the geometric relationship of Fig. 7 obtains the cart dynamic equation is:

$$k = \frac{l}{\tan(d\phi)}, r = \frac{l}{\sin(d\phi)}, \text{ and } \frac{dx_r}{dy_r} = \tan(d\phi - \theta) \tag{9}$$

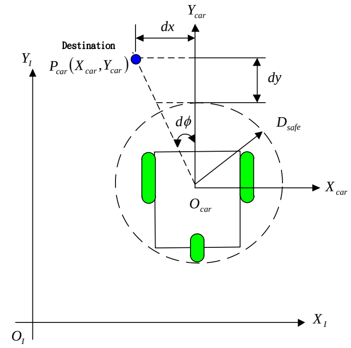


Figure. 6 Cart moving distance dy and moving angle $d\phi$.

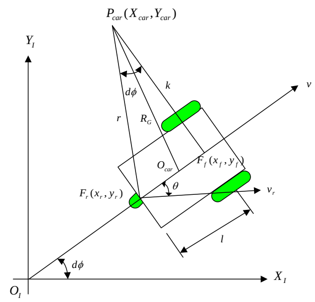


Figure. 7 The cart steering

In the cart front wheel center

$$\frac{dy_f}{dx_f} = \tan(d\phi) \tag{10}$$

In case of cart wheels movement with no slip, the front wheel coordinates(x_f, y_f) can be inferred from the rear wheel coordinates (x_r, x_r) as:

$$\begin{aligned} y_f &= y_r + l \cdot \sin\theta \\ x_f &= x_r + l \cdot \cos\theta \end{aligned} \tag{11}$$

If the rear wheel speed v_r and the steering angle θ are known, the cart dynamic equation can be written as:

$$\dot{y}_r = v_r \sin(d\phi)\cos\theta, \; \dot{x}_r = v_r \cos(d\phi)\cos\theta, \; d\dot{\phi} = v_r \frac{\sin\theta}{l} \tag{12}$$

Where $\dot{y}_r$ and $\dot{x}_r$ are cart speed component in the Y_I and X_I axis direction, respectively. $d\dot{\phi}$ is steering variation. For a front wheel derived cart, the relationship of rear wheel speed v_r and cart speed v is:

$$v_r = \frac{v}{\cos\theta} \tag{13}$$

The (12) are rewritten as:

$$\dot{y}_r = v \cdot \sin(d\phi), \; \dot{x}_r = v \cdot \cos(d\phi), \; d\dot{\phi} = v \cdot \frac{\tan\theta}{l} \tag{14}$$

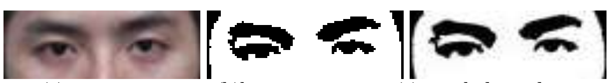
International Journal of Automation Mechatronics & Robotics – IJAMR
Volume 1 : Issue 1 [ISSN 2374 – 1546]
Publication Date : 09 January 2014

IV. Experimental Results

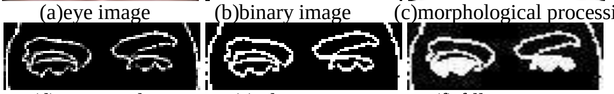
This paper designs an autonomous cart shown in Fig. 8. The autonomous cart equipped with a camera to grab the images of the task environment send to a note book PC. The image processing algorithms proposed in II are implemented to recognize the commander and his or her hands gestures. The recognition results are used for the cart motion control. The control algorithms are shown in section III.

Figure. 8 An autonomous cart.

Fig. 9 and Fig. 10 are the image processing results of recognizing the commander and his hand gesture, respectively. There are a variety of gabbed images by the camera on the cart. The different view angles need different templates for matching. From various angles grab 3 images of a commander face. Each commander with total of 21 images used to create his templates. In the experiments, the cart camera grabbing image is processed to match with the 21 templates for looking the commander. Approximated 1/3 commander images cannot be recognized. The main problem is that the 21 templates are not sufficient to treat the variations of eye features. More templates are required for increasing the recognition results. In addition, illumination is the key factor affects the results of binary image so as to impact the resulting features.



(a)eye image (b)binary image (c)morphological processing

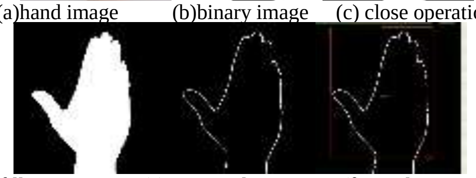


(d) contour detection (e) close operation (f) fill operation

(g) eyes template (h) template matching

Figure 9 Facial feature recognition

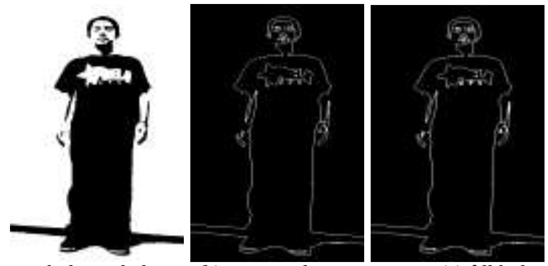
(a)hand image (b)binary image (c) close operation



(d) fill operation (e)contour detection (f)template matching

Fig. 10 Hand gestures recognition.

Fig. 11 is an illustration for determining the commander height in pixels. Because the commander height is known in meters, the camera pinhole model is used to get the commander distance in front of the cart. The hand gesture combined with the commander distance for the cart planning its motion paths.



(a) morphological close (b) contour detection (c) fill holes

Fig.11 A commander image

Fig. 12 (a) the commander location (1m, 3m)

Fig. 12 (b) the commander location (-1m, 3m)

The experimental results are shown in Fig. 12. The autonomous cart in the middle position which is the origin (0,0) of cart coordinates. The commander located in front and right of the cart 3 meters and 1 meter (1000mm, 3000mm). The cart move to reach the destination (1010mm, 2880mm). The trajectory and the location errors are shown in Fig. 12(a). Fig. 12 (b) is the experimental results of commander stand in front and left of the cart 3 meters and 1 meter (-1000mm, 3000mm). Finally, the cart reach the position (-890mm, 2900mm).

In the experiments, the cart autonomously moves to the destination. There are the errors because of wheel slips and distance estimation errors of image processing.The template is designed so that author affiliations are not repeated each time for multiple authors of the same affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization). This template was designed for two affiliations.

V. Conclusion

This paper studied image identification methods which are applied to an autonomous cart. The proposed algorithms have the function to correctly detect the commander and his or her gestures and locations. The experimental results have shown that in a non-obstacle space the cart with camera and the proposed image processing algorithms that are successfully for self-navigation. The wheel slip, environmental illumination, and face feature variations are not consideration that cause cart moving errors and recognition fails. It is worth for further study.

References

[1] J. Hill and W. T. Park, “Real Time Control of a Robot with a Mobile Camera,”Proceeding of the 9th International Symposium on Industrial Robots (ISIR), 1979, pp. 222-246.

[2] Nai-Wen Wu, “A Potential-field-based Method for Path Planning of Autonomous Guided Vehicles,” M.S. thesis, Dept. Elect. Eng., National Yunlin Univ. of Science and Technology, Yunlin, Taiwan, 2004.

[3] C. H. Lin and K. C. Fan, “Triangle-Based Approach to the Detection of Human Faces,” Pattern Recognition, Vol. 34, No. 6, Jun. 2001.

[4] I. S. Hsieh and K. C. Fan, “A Statistic Approach to the Detection of Human Faces in Color Nature Scene,” Pattern Recognition, Vol. 35, No. 7, 2002, pp. 1583-1596.

[5] Ming-Feng Huang, “The study of Monitoring and Control System Using Human Face Tracking Methods,” M.S. thesis, Dept. Computer Science & Information Engineering, National Cheng Kung Univ., Tainan, Taiwan, 2003.

[6] Wen-Jun Chen, “Face Recognition and Its Applications to Human-Robot Interaction,” M.S. thesis, Dept. Elect. Eng., National Chiao Tung Univ., Hsinchu, Taiwan, 2003.

[7] M. Turk and A. Pentland, “Eigenfaces for Recognition,” Journal of Cognitive Neuro-science, Vol. 3, No. 1, 1991, pp. 71-86.

[8] Yi-Tsung Chien, “A Complementary Strategy for Skin-color-based Face Detection,” M.S. thesis, Dept. Elect. Eng., National Central Univ., Taoyuan, Taiwan, 2001.

[9] Weng-Hang Tam, “A Human Face Tracker System Design Study Using the Technology of Digital Image Processing,” M.S. thesis, Dept. Elect. Eng., National Cheng Kung Univ., Tainan, Taiwan, 2000.

[10] Yao-Zhang Xie, “A Study of Real-Time Face Tracking with an Active Camera,” M.S. thesis, Dept. Mechanical and Electro-Mechanical Eng., National Sun Yat-Sen Univ., Kaohsiung, Taiwan, 2005.

[11] S. Y. Lee, Y. K. Ham, and R. H. Park, “Recognition of Human front Faces using Knowledge-based Feature Extraction and Neuro-fuzzy Algorithm,” Pattern Recognition, Vol. 29, No. 11, 1996, pp. 1863-1867.

[12] C. C. Han, H. Y. Mark Liao, G. J. Yu, and L. H. Chen, “Fast Face Detection Via Morphology-based Pre-processing,” in Proc. 9th International Conference on Image Analysis and Processing, Florence, Italy, 17-19 Sep. 1997, Lecture Notes in Computer Science, Vol. 1311, pp. 469-476.

[13] Pei-Ying Chen, “A Robust Visual Servo System for Tracking an Arbitrary-Shaped Object by a New Active Contour Method,” M.S. thesis, Dept. Elect. Eng., National Taiwan Univ., Taipei, Taiwan, 2003.

[14] Tien-Pao Wu, “A Visual Servoing System for Object Tracking Applications of Mobile Robots Based on Shape Features,” M.S. thesis, Dept. Mechanical and Computer-Aided Eng., St. John's Univ., Taipei, Taiwan, 2005.

Ying-Shing Shiao received the Ph.D. in Electrical Engineering from National Cheng Kung University in 1992. He is a Professor in Department of Electrical Engineering, National Changhua University of Education, Changhua, Taiwan. His research interest includes computer vision, digital image processing, control engineering, and intelligent robot.

Po-Hsiang Yu received the B.S. degree in Industrial Education and Technology from National Changhua University of Education in 2012. He is pursuing the M.S. degree in Department of Electrical Engineering, National Changhua University of Education, Changhua, Taiwan. His current research interests include visual servo control and computer vision.

Ching-Wei Wu received the M.S. in Electrical Engineering from National Changhua University of Education in 2009. He is pursuing the Ph.D. degree in Department of Electrical Engineering, National Changhua University of Education, Changhua, Taiwan. His current research interests include visual control and computer vision.

Chieh-Wen Liao received the B.S. degree in Department of Computer and Communication from Shu-Te University in 2011. He is pursuing the M.S. degree in Department of Electrical Engineering,

National Changhua University of Education, Changhua, Taiwan. His current research interests include visual control and computer vision.