Research Article

Development and Application of the Stereo Vision Tracking System with Virtual Reality

Chia-Sui Wang,¹ Ko-Chun Chen,² Tsung Han Lee,³ and Kuei-Shu Hsu²

¹ Department of Information Management, Chia Nan University of Pharmacy & Science, Tainan 717, Taiwan

² Department of Applied Geoinformatics, Chia Nan University of Pharmacy & Science, Tainan 717, Taiwan

³ Project Coordination and Liaison Section, Planning and Promotion Department, Metal Industries Research and Development Center, Taipei 10075, Taiwan

Correspondence should be addressed to Kuei-Shu Hsu; ksksu888@gmail.com

Received 28 June 2014; Accepted 9 September 2014

Academic Editor: Stephen D. Prior

Copyright © 2015 Chia-Sui Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A virtual reality (VR) driver tracking verification system is created, of which the application to stereo image tracking and positioning accuracy is researched in depth. In the research, the feature that the stereo vision system has image depth is utilized to improve the error rate of image tracking and image measurement. In a VR scenario, the function collecting behavioral data of driver was tested. By means of VR, racing operation is simulated and environmental (special weathers such as raining and snowing) and artificial (such as sudden crossing road by pedestrians, appearing of vehicles from dead angles, roadblock) variables are added as the base for system implementation. In addition, the implementation is performed with human factors engineered according to sudden conditions that may happen easily in driving. From experimental results, it proves that the stereo vision system created by the research has an image depth recognition error rate within 0.011%. The image tracking error rate may be smaller than 2.5%. In the research, the image recognition function of stereo vision is utilized to accomplish the data collection of driver tracking detection. In addition, the environmental conditions of different simulated real scenarios may also be created through VR.

1. Introduction

The VR application games on computers and mobile phones are currently very popular on internet. As the user simulates vehicle-running situations by means of VR scenario, multiple cameras may be utilized to observe the user [1]. Moreover, for condition analysis occurs in vehicle traveling, the situational simulation may be utilized in addition to environmental factors and external variable conditions to test the response, action variation, and so forth of the user [2]. It may also become the reference for developing equipment and parts for vehicles. The custom of the user may be obtainable through direct operation of VR [3].

In the research, the application of stereo vision system on VR system is the topic. The computer vision system may be roughly divided into planar vision and stereo vision [4–6]. The difference lies in the capability of estimating the depth

information of the target object in the image, also referred to as depth perception, such that the computer vision system is more practical. There are many practical applications, such as remote sensing and monitoring, medical image processing, robot vision system, military reconnaissance, mineral exploration, cartography, and the like [7–9].

Additionally, due to the advance of technology, the application of 3D stereo vision is prosperous, such that the visual enjoyment of human beings is improved. In order for people to perform test with VR scenarios in various different environments and experience real feeling, in addition to creating environmental conditions simulating real scenarios, we will employ Unity, software, to make a VR system and use in conjunction with various scenarios. Subsequently, it may be promoted to smart phone platforms for use [10, 11]. With VR for simulation, not only it is safe and more convenient, but also the test accuracy will not be influenced by the environmental factors introduced by real test [12, 13].

2. Research Method

How to obtain the depth of an object in the image? The main issue is the possibility to find out the corresponding points in the stereo image. What is a stereo image? An image is called a set or a pair of stereo images as long as it is taken by extracting simultaneously for the same object or target with two or more cameras mounted in different locations. What are corresponding points? They are the projections of a certain point of an object at different locations in a 3D space. The location difference of two corresponding points in paired images is referred to as disparity. The disparity is associated with the locations of the corresponding points in space, orientation, and physical properties of cameras. If the parameters of cameras are known, the depth of the object may be calculated from the image. At first, we explain how the points in the space are projected onto the image plane. Assume the coordinate value of any point P in the space relative to CCD center is (X, Y, Z) , it is imaged to P' point in the image after projection. Its coordinate value relative to the image center is (x, y) , while the center point of the image relative to CCD coordinate value is $(0, 0, f)$, where f is the distance from the center point of CCD to the sensing field. Concept View of Stereo Vision Image Mechanism Space shows a concept view of the stereo vision system space. The research utilizes the stereo vision system as principal axis. Two lenses are used to determine image depth. This part is a feature that the single-eye vision cannot achieve. As shown in Figure 1, the single-eye vision can only scan target object for a plane without knowing the image depth of the target object, such that the stereo vision is necessary for enhancement by using two-eye focusing function to calculate the corresponding angle for obtaining the image depth of the target object:

$$\begin{aligned} \begin{bmatrix} x_k \\ y_k \\ z_k \end{bmatrix} &= R \left(\begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + T \right), \quad \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} = \begin{bmatrix} -T_x \\ -T_y \\ -T_z \end{bmatrix}, \\ R &= R_x R_y R_z R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{bmatrix}, \\ R_y(\theta) &= \begin{bmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{bmatrix}, \quad R_z(\theta) = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned} \quad (1)$$

2.1. Eye Recognition Analysis. Eye recognition generally comprises pretreatment, feature extraction, sample learning, recognition, and so forth. There are various methods to implement eyeball recognition technology as follows.

- (1) Projection method: in the projection method, the eye location is detected according to the distribution feature of image projected in a certain direction. The projection method is a statistical method, which utilizes the gray level information of eyes to detect the ordinate and abscissa of pupils by means of horizontal and vertical projections, respectively. Thereby,

FIGURE 1: Concept view of stereo vision image mechanism space.

accurate positioning for human eyes is available. Integral projection function (IPF) and mean projection function are common projection functions.

- (2) MPF and variance projection function (VPF)—Hough transformation method: Hough transformation transforms the image from spatial domain to parametric domain. The curve in image is illustrated in a certain parametric form satisfied by most boundary points. The pupil is used as a standard circle. The location of the eye pupil may be positioned accurately with the Hough transformation by means of the standard equation of circle: $(x - a)^2 + (y - b)^2 = r^2$. The robustness of the Hough transformation is enhanced considerably because of apparent geometric analyticity.
- (3) AdaBoost classification method: the AdaBoost algorithm is a highly effective iterative operation algorithm in robotic learning field. It trains different weak classifiers for the same training set, followed by collecting these weak classifiers together to compose a strong classifier. This algorithm is advantageous for higher classification precision and faster human eye recognition. However, the effectiveness of such algorithm depends on the choice of classifiers. There is a significantly important application with respect to fast human eye detection.
- (4) Sample matching method: the location of the pupil may be searched dynamically in the image window from left to right and from up to down by using circular sample according to the pupil shape. The sample matching searches for a small image from a larger image. It identifies the target location by using the most similar location as matching point by means of the similarity calculation for sample and matching zones. The sample matching algorithm is within the scope of robotic learning field and is an effective eye recognition algorithm.

2.2. Sobel Edge Detection. The Sobel operator comes from the combination of differential operation and low pass operation. It has the effect of edge detection in addition to the advantage of noise reduction. Because the derivative will reduce the

FIGURE 2: Schematic view of entire system structure (a). Schematic view of entire system structure (b).

noise intensity, the feature of Sobel operator to filter noise is particularly advantageous. The derivative of Sobel operator mask is shown in the following formulas:

$$\nabla f = |Gx| + |Gy|,$$

$$\begin{bmatrix} P1 & P2 & P3 \\ P4 & P5 & P6 \\ P7 & P8 & P9 \end{bmatrix}, \quad gx = \begin{bmatrix} P1 & P2 & P3 \\ P4 & P5 & P6 \\ P7 & P8 & P9 \end{bmatrix},$$

$$gy = \begin{bmatrix} P1 & P2 & P3 \\ P4 & P5 & P6 \\ P7 & P8 & P9 \end{bmatrix}, \quad (2)$$

$$\begin{bmatrix} f(x-1, y-1) & f(x, y-1) & f(x+1, y-1) \\ f(x-1, y) & f(x, y) & f(x+1, y) \\ f(x-1, y+1) & f(x, y+1) & f(x+1, y+1) \end{bmatrix}.$$

The input for said method is a deep image corresponding to 1280×1024 RGB colorful image, which is the information that may be provided by stereo vision. The skin color will be further detected from the generated maximum skin colorful fleck. From conservative estimation, the method for calculating the spatial expansion in eyes includes a circular mask expansion, whose radius is $r = 15$. In view of the previous tracking of estimated 3D location, the 3D point of skin color is within a predetermined depth range (25 mm). The estimation is reserved, while other depths are set to zero:

$$E(h, O) = D(O, h, C) + \lambda_k * Kc(h),$$

$$D(O, h, C) = \frac{\sum_{\min} (|o_d - r_d|, d_M)}{\sum (o_s \vee r_m) + \varepsilon} \quad (3)$$

$$+ \lambda \left(1 - \frac{2 \sum (o_s \wedge r_m)}{\sum (o_s \wedge r_m) + \sum (o_s \vee r_m)} \right).$$

2.3. Random Optimization by Means of Pixel Group. Global optimization: update equation reevaluating the velocity and location of each pixel:

$$V_{k+1} = w(v_k + c_1 r_1 (P_k - x_k) + c_2 r_2 (G_k - x_k)), \quad (4)$$

$$X_k + 1 = X_k + V_k + 1.$$

Such objective function is optimized and one single frame is assumed for eye location. Therefore, such method and

the sequence optimization question necessary for tracing eyes of a person obtain the characteristic value for each point. The spatial continuity is utilized because it depends on the sampling frequency of the expected observation motion image.

3. Experiment of VR Driver Tracking Recognition System

The research belongs to tentative research. 50 operators who have ever driven vehicles are sampled randomly. There are total 100 operators invited to perform the field test of the “VR driver behavior test system.” Thereby, the influence and idea of the VR driver behavior test system with respect to drivers are studied. In the program, the 3D virtual image technique is applied to test vehicle drivers. The building model of this city is created using Google SketchUp modeling software, followed by employing Unity to make the VR application environment of our city, to present the image with high similarity relative to the entity, and to simulate scenes in various places and scenes from various angles realistically. The entire system structure is shown in Figure 2.

As seen in Figure 3, by combining 3D stereo virtual image technique with space plan and scene design, it can simulate the change of various weather conditions and the four seasons, climate conditions and water scene, mist effect, and so forth. Alternatively, it may improve existent scenario by simulating the situation in the plan realistically come to the design. Further, it is studied and evaluated with respect to practicability. Not only various different visual effects are available easily, but also the error rate is reduced effectively. Moreover, the reliability and reality of the program are improved significantly. With such a technology, the vehicle-driving test, vehicle accessory development, vehicle driving environmental factors, and the like may also be combined to simulate and evaluate different situation designs.

For the vehicle VR stereo reality tracking system of the research, in simulating driving process, the computer is used to aid in the modeling software to create a virtual field and create real street and virtual environment in advance. The model implementation is presented in VR system for it to compare different driving behaviors and environmental variables. In addition, the VR system is added

FIGURE 3: Structure view of the research.

FIGURE 4: Stereo vision pupil recognition analysis (a). Stereo vision pupil recognition analysis (b).

with various sudden condition simulations to compare the data of simulated environment and the real vehicle traveling data mutually, to provide developers with data collection of driver behavior and improvement reference. An integrated simulation is performed using Unity 3D game engine. By means of the operation data and motion behavior in subject and scenario, followed by applying various different environmental variables, and the situation of accidental occurrence of driver in driving vehicle for general vehicle traveling process, the impact of various environmental variables on driver behavior and the impact of visual angle of driver from different environmental variables are analyzed, to improve data analysis of vehicle driver behavior in order for increased development benefit and cost. Figure 5 shows the possible impacts of sudden motorcycles and pedestrians in driving. The stereo vision is utilized to track the concentration points of the eyes of drivers and the data is collected for analysis, as shown in Figure 4.

4. Conclusions

The stereo image vision determination determines target objects by subtracting images in conjunction with marginalization. From the data and figures, it slightly shows that there might have large one time instant misjudgment. Besides marginalization, special color recognition is also used. Since the color recognition results from subtraction of 2D matrices between simple RGBs, there is no significant impact for image processing speed. By eliminating the colors outside the desired target object under tracking, from such filtering and judgment, the possibility of misjudgment of image depth measurement will be reduced for the target object. It does the best to exclude the issue of marginal intervening interference.

Because of the popularity of smart products in present society, in response to such social phenomena, the program develops street scenes of various different test environments, different climate environments, and different road

FIGURE 5: Impacts of motorcycles on vehicle turning in driving.

test environments. It not only increases the test stability, reliability, and practicability substantially but also improves driver verification because the stereo vision is used for tracking in VR. The system platform structure is simple, with project VR scenes, in conjunction with stereo vision system, the implementation of VR scenes is rapid and easy, and various uncertain factors may be added in the game engine. Once the VR scenarios and interactive interface development are finished, the entire system may be completed quickly. In addition, the provided real test information may be used for product development evaluation. Compared to general ones, which can perform field test only when the product is developed completely, the time and cost are reduced substantially without safety issues. Through the simulation test done by the driver of the system, the obtained result may be used as an evaluation platform for driver behavior.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Authors' Contribution

Chia-Sui Wang wrote the paper, Ko-Chun Chen and Tsung Han Lee carried out the experiment of VR driver tracking recognition system, and Kuei-Shu Hsu carried out programmed check of the system. All authors read and approved the final paper.

Acknowledgments

This research is supported by the National Science Council, Taiwan, under Contracts nos. NSC 102-2221-E-041-007 and NSC 102-2622-E-041-002-CC3.

References

- [1] R. Mahony and T. Hamel, "Image-based visual servo control of aerial robotic systems using linear image features," *IEEE Transactions on Robotics*, vol. 21, no. 2, pp. 227–239, 2005.
- [2] L. M. Bergasa, P. F. Alcantarilla, and D. Schleicher, "Non-linearity analysis of depth and angular indexes for optimal stereo SLAM," *Sensors*, vol. 10, no. 4, pp. 4159–4179, 2010.
- [3] N. Kobayashi and M. Shibata, "Visual tracking of a moving object using a stereo vision robot," *Electronics and Communications in Japan*, vol. 91, no. 11, pp. 19–27, 2008.
- [4] M. Kim, J. G. Jeon, J. S. Kwak, M. H. Lee, and C. Ahn, "Moving object segmentation in video sequences by user interaction and automatic object tracking," *Image and Vision Computing*, vol. 19, no. 5, pp. 245–260, 2001.
- [5] L. Palopoli, L. Abeni, G. Bolognini, B. Allotta, and F. Conticelli, "Novel scheduling policies in real-time multithread control system design," *Control Engineering Practice*, vol. 10, no. 10, pp. 1091–1110, 2002.
- [6] Y.-P. Hung, C.-S. Chen, Y.-P. Tsai, and S.-W. Lin, "Augmenting panoramas with object movies by generating novel views with disparity-based view morphing," *The Journal of Visualization and Computer Animation*, vol. 13, no. 4, pp. 237–247, 2002.
- [7] P. Kaewtrakulpong and R. Bowden, "A real time adaptive visual surveillance system for tracking low-resolution colour targets in dynamically changing scenes," *Image and Vision Computing*, vol. 21, no. 10, pp. 913–929, 2003.
- [8] C.-M. Lai, H.-M. Huang, S.-S. Liaw, and W.-W. Huang, "A study of user's acceptance on three-dimensional virtual reality applied in medical education," *Bulletin of Educational Psychology*, vol. 40, no. 3, pp. 341–362, 2009.
- [9] T. Monahan, G. McArdle, and M. Bertolotto, "Virtual reality for collaborative e-learning," *Computers and Education*, vol. 50, no. 4, pp. 1339–1353, 2008.
- [10] G. Tsechpenakis, K. Rapantzikos, N. Tsapatsoulis, and S. Kollias, "A snake model for object tracking in natural sequences," *Signal Processing: Image Communication*, vol. 19, no. 3, pp. 219–238, 2004.
- [11] Z. Pan, A. D. Cheok, H. Yang, J. Zhu, and J. Shi, "Virtual reality and mixed reality for virtual learning environments," *Computers & Graphics*, vol. 30, no. 1, pp. 20–28, 2006.
- [12] Y. Fang, W. E. Dixon, D. M. Dawson, and P. Chawda, "Homography-based visual servo regulation of mobile robots," *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics*, vol. 35, no. 5, pp. 1041–1050, 2005.
- [13] J. Kim, I. Fisher, A. Yezzi, M. Çetin, and A. S. Willsky, "A nonparametric statistical method for image segmentation using information theory and curve evolution," *IEEE Transactions on Image Processing*, vol. 14, no. 10, pp. 1486–1502, 2005.