A Single Camera Stereo System for Obstacle Detection

Zhigang Zhu*, Xueyin Lin, Dingji Shi, Guangyou Xu
Department of Computer Science and Technology
Tsinghua University, Beijing 100084, China

ABSTRACT

This paper presents a real-time visual obstacle detection system which is novel in both sensors and algorithms. Single-camera stereo vision is realized by using a set of handily placed mirrors, and the optical and electrical identities of binocular image pair are guaranteed. A software gaze transformation method is presented to produce zero disparities for road surface figures, thus facilities the obstacle detection. The proposed non-zero disparity filtering (NZDF) algorithm makes the obstacle detection simple and robust. Problems of feature detection and correspondence are avoided and the system is robust, near real-time, and applicable to natural scene environments. Experimental results are given for outdoor real scene images.

Keywords: Mobile robot, Intelligent vehicle, Obstacle detection, Single-camera stereo, Gaze transformation

1. INTRODUCTION

A mobile robot must be capable of detecting obstacles in real time in order to carry out the autonomous navigation task in an unstructured outdoor environment. Existing obstacle detection systems typically compute the position of a obstacle relative to the mobile robot by using range information [1,2], which may be obtained from a laser ranger finder or various vision-based techniques. Among them the stereo/motion vision systems have advantages of low cost, low power consumption and a high degree of mechanical reliability. However the relatively slow processing speed of the existing stereo vision methods limit their applications in real world problems, mainly due to the long computing time required for establishing stereo correspondences.

We believe that the reconstruction of a detailed 3D world map is tremendously difficult, computationally prohibitive and not necessary in some of the applications, such as obstacle detection. Furthermore, in its most general form, distinguishing a visual target (e.g. an obstacle) is the notorious segmentation problem, and it is not clear how to extract this information from visual signals in all cases. In paper [3], controlled camera movements are used to simplify the visual processing necessary to keep the camera locked on the target. The central idea is that localizing attention in 3D space makes the pre-categorical visual processing sufficient to hold gaze. Converged cameras produce a horopter (surface of zero stereo disparity) in the scene and hence binocular features with no disparity can be located with a simple filter, showing the object's location in the image. Instead of requiring a way to recognize the object, the system relies on active control of the camera movements and binocular fixation to locate the target.

Simple and fast stereo obstacle detection methods have been proposed based on that fact that obstacles mostly lie on the flat ground [4-7]. Badal et al [4] developed a practical obstacle detection and avoidance system for outdoor robot. A fast hierarchical stereo correspondence algorithm produces sparse disparity information and compares it with that of the ground plane. The system can only detect obstacle protruding sufficiently high from the flat ground plane with the requirements of 1D epipolarity and identical cameras and digitizers. The system requires a tedious calibration procedure. The system executes at 2 Hz on a Datacube MV-200 and SGI

* The author is currently on a leave in the Computer Science Department, University of Massachusetts at Amherst, MA 01003. Email: zhu@cs.umass.edu, or zhuzhg@mail.tsinghua.edu.cn. This work was supported by China Advanced Research Project during 1990-1996.

host.

Stojorham et al [5] proposed a computationally simple floor level detection method which can check the presence of obstacle above the ground plane. A geometrical image transformation, namely inverse perspective mapping was introduced to project both original stereo images onto the ground plane so that points on the ground plane are assigned zero disparities. Although this method is computationally efficient, the resolutions of the mapped images are reduced duo to the severe distortion by the inverse perspective transformation.

Zhu & Lin [6] proposed two fast obstacle detection algorithms based on the so-called reprojection transformation, which also projects the original images on the ground plane. The characteristic of the methods is that only one camera is used and motion parallax is utilized. The first algorithm differs from that of Stojorham's in that the accurate motion parameter (the "baseline") between two images is not available so that a search for zero disparity matching of the ground plane is needed. The second algorithm creates the epipolar plane image from the re-projected image sequence and the 3D coordinates of the obstacle are accurately estimated. The algorithms have been implemented in PIPE, a pipelined image processing engine, at 5 Hz and 30 Hz respectively.

Cho et al [7] proposed a computationally efficient stereo vision method for obstacle detection in indoor mobile robot navigation. The original stereo images are geometrically transformed into a common image plane located at the mid-point of the stereo image plane. In the transformed stereo images, the image pixels of the floor plane have zero isodisparities, while the obstacle features above the floor level correspond to non-zero isodisparities. The proposed correspondence scheme of the isodisparity prediction and verification facilitates frame-wise operations without an exhausted searching process for stereo matching, provided that the paired image edges have unique matches by the edge angles and intensity patterns. Camera calibration is needed for the geometrical transformation, which introduces some system error. The isodisparity checking is carried out for every constant disparity value so that the required time for the entire processing in PC is 15 seconds.

The limitations of the aforementioned work are in twofold. First, the cameras should be perfectly calibrated so that the exact zero disparity properties for a given plane can be guaranteed. This requirement raises a tedious problem for the calibration algorithm. Second, The identity assumption is difficult to keep for two physical separated cameras and digitizers. With the emphasis to overcome these limitations, we have developed a near real-time stereo vision system for obstacle detection. The system has the following features:

1) A general model is proposed for the zero disparity gaze transformation. There is no need for the calibration of the internal and external camera parameters in order to realize the gaze transformation. Hence the system is simple and the transformation is accurate.
2) A single-camera stereo equipment is designed so that the electrical and photometric identity can be satisfied for the stereo pair. The system is cheap because only a single camera and a single digitizer are needed.
3) Our work also explores the use of pre-categorical visual cues (prior to object recognition) in order to distinguish the visual target (obstacle) from the surrounding scene (road). However no mechanical control of camera movement and vergence is needed.

This paper is organized as follows. Section 2 describes the principle of the image gaze transformation. The designs of two types of single camera stereo systems are given in Section 3. Section 4 gives the basic obstacle detection algorithm based on the zero-disparity property of the ground plane. A few discussions and concluding remarks are given in the last section.

2. GAZE TRANSFORMATION

Suppose the coordinate system of left and right cameras are $X_1Y_1Z_1$ and $X_2Y_2Z_2$ respectively, where Z_1, Z_2 are their respective optical axes, and the world (robot) coordinate system is *XYZ* (Fig. 1). The relations between the three coordinates are

$$X = R_1 X_1 + T_1 = R_2 X_2 + T_2 \qquad (1)$$

where X_1, X_2 and X are the coordinates of a point in the left camera, right camera and robot coordinate systems respectively with the forms

$$X_1 = (x_1, y_1, z_1)^T,\ X_2 = (x_2, y_2, z_2)^T,\ X = (x, y, z)^T,$$

R_1, R_2 are the rotation matrixes from $X_1Y_1Z_1$ and $X_2Y_2Z_2$ to XYZ respectively, and T_1, T_2 are the respective transitional vectors. The binocular relation is

$$X_2 = R_2^{-1} R_1 X_1 + R_2^{-1}(T_1 - T_2) = RX_1 + T \qquad (2)$$

where

$$R = R_2^T R_1 = R_2^{-1} R_1 \stackrel{\Delta}{=} (r_{ij})_{3\times 3}$$

and

$$T = R_2^{-1}(T_1 - T_2) = R_2^T(T_1 - T_2) \overset{\Delta}{=} (T_x, T_y, T_z)^T \quad (3)$$

Under the pinhole camera model, the left and right image coordinates for the point X are

$$(u_1, v_1) = \left(F\frac{x_1}{z_1}, F\frac{y_1}{z_1}\right),\ (u_2, v_2) = \left(F\frac{x_2}{z_2}, F\frac{y_2}{z_2}\right) \quad (4)$$

respectively, where F is the equivalent focal length.

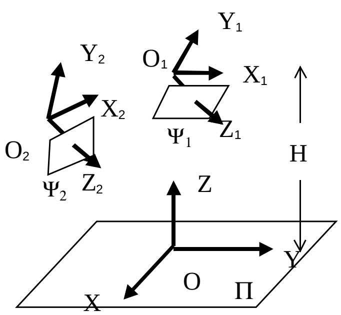


Fig. 1. Coordinate systems and the gaze transformation

From (2) and (4) we have the following relations:

$$\begin{cases} u_2 = F\dfrac{r_{11}u_1 + r_{12}v_1 + Fr_{13} + F\frac{T_x}{z_1}}{r_{31}u_1 + r_{32}v_1 + Fr_{33} + F\frac{T_z}{z_1}} \\ v_2 = F\dfrac{r_{21}u_1 + r_{22}v_1 + Fr_{23} + F\frac{T_y}{z_1}}{r_{31}u_1 + r_{32}v_1 + Fr_{33} + F\frac{T_z}{z_1}} \end{cases} \quad (5)$$

In the above equations, the relationship between (u_1, v_1) and (u_2, v_2) is relevant to the depth (z_1) of the point (x_1, y_1, z_1). If this point lies in a plane Π (e.g. ground plane)

$$p x_1 + q y_1 + r z_1 = 1 \quad (6)$$

the following relation can be derived from (4) and (6)

$$p u_1 + q v_1 + r F = \frac{F}{z_1} \quad (7)$$

If the plane Π coincides with plane XOY, then (p, q, r) can be derived directly from T_1 and the third column of R_1. Substituting Eq. (7) into Eq. (5) we have

$$\begin{cases} u_2 = F\dfrac{r_{11}u_1 + r_{12}v_1 + Fr_{13} + T_x(pu_1 + qv_1 + rF)}{r_{31}u_1 + r_{32}v_1 + Fr_{33} + T_z(pu_1 + qv_1 + rF)} \\ v_2 = F\dfrac{r_{21}u_1 + r_{22}v_1 + Fr_{23} + T_y(pu_1 + qv_1 + rF)}{r_{31}u_1 + r_{32}v_1 + Fr_{33} + T_z(pu_1 + qv_1 + rF)} \end{cases} \quad (8)$$

The cameras' bump with the vehicle only affects (p, q, r) in Eq. (8). When the vehicle moves on a flat plane strictly, all of parameters in Eq. (8) keep unchanged, and therefore Eq. (8) can be simplified as

$$\begin{cases} u_2 = \dfrac{C_1u_1 + C_2v_1 + C_3}{E_1u_1 + E_2v_1 + E_3} \\ v_2 = \dfrac{D_1u_1 + D_2v_1 + D_3}{E_1u_1 + E_2v_1 + E_3} \end{cases} \quad (9)$$

Eq. (9) build up the relationship of the images captured from two viewpoints about a plane Π (e.g. ground plane), which is named as **gaze transformation** about plane Π. If the line passing through optical centers of the left and right cameras is parallel to the plane Π, the gaze transformation is composed of a series of transformations -- a rotation transformation R_1 which projects points of the left image to the plane Π, then a transitional transformation relevant to the baseline between the two cameras, and finally an inverse rotation transformation R_2^{-1} which re-projects the points in plane Π to the right image plane(Fig.1). Generally speaking, the gaze transformation in Eq. (9) is the 2D projective transformation, and can be expresses in the homogeneous form:

$$\begin{pmatrix} k_2u_2 \\ k_2v_2 \\ k_2 \end{pmatrix} = \begin{pmatrix} C_1 & C_2 & C_3 \\ D_1 & D_2 & D_3 \\ E_1 & E_2 & E_3 \end{pmatrix} \begin{pmatrix} k_1u_1 \\ k_1v_1 \\ k_1 \end{pmatrix} \quad (10)$$

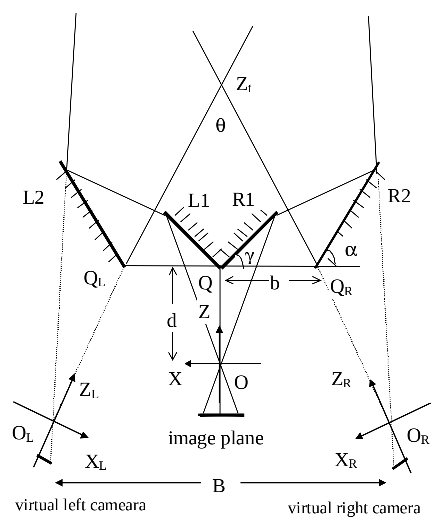
where

$$M = \begin{pmatrix} C_1 & C_2 & C_3 \\ D_1 & D_2 & D_3 \\ E_1 & E_2 & E_3 \end{pmatrix}$$

is a non-singular matrix. From the 2D projective theorem, the gaze transformation can be determined by at least 4 point pairs. Since we do not need to measure any 3D coordinates in the world, the gaze transformation is based on uncalibrated image pairs. If the left image re-projects to the right image using the gaze transformation of Eq. (9), the point in the ground plane will have zero disparities between the reprojected left image and the original right image. This property is very useful for the image-level ground obstacle detection. Comparing to the previous work, only one image, say left image, is reprojected. Furthermore the image resolution is preserved since no severe distortion is introduced by the geometrical transformation even the optical axis of the camera is almost parallel to the ground plane in order to detect distant obstacles.

3. SINGLE CAMERA STEREO

The patented design [8] of an optical based binocular vision system using a single camera stems from the requirement to record the outdoor binocular image sequences by a camcorder frequently and portably. However advantages of a single-camera binocular vision system are far beyond this intuitive fact. First the system is cheap and effective. The stereo image pair can be captured in a snap (simultaneously at the frame rate or field rate). Only a single camera and a single frame grabber are required, and no electrical synchronization is needed. Second, the image pair is identical. Since a single camera is used to capture the stereo pair, no differences of electrical, optical and photometric properties exist, such as iris, focal length, noise level, sensitivity, contrast and etc. The image identity is favorable to the obstacle detection in gray-level image. Third, it can be constructed as a concise binocular vision system for various application, such as active vision and tele-reality. The fixation point, baseline length, pan/tilt angles can be easily controlled by rotating and translating the auxiliary mirrors.



(a) system (top view)

(b) a real image

Fig 2. Basic single-camera binocular equipment: left-right images

Binocular equipment I: right-left images

The basic binocular vision equipment is shown in Fig. 2. Images viewed from left and right viewpoints are projected into the left and right halves of a single image sensor respectively. In the standard setting, the angle between the left (right) primary mirror $L_1(R_1)$ and the image plane is $\gamma = 45^\circ$, and the angle between left (right) auxiliary mirror $L_2(R_2)$ and the image plane is $\alpha \in \left(45^\circ, 90^\circ\right)$. The distance between the optical center (O) of the real camera and the intersecting line of L_1 and R_1 is d, and the nearest distance between left (or right) auxiliary mirror and the optical axis Z is *b*. Half viewing-angle of the camera is β. The single real camera, left/right primary and auxiliary mirrors construct a system of equivalent left and right virtual "cameras" with viewpoints O_L and O_R. The left and right images are projected into the left and right halves of a single sensing plane.

Given d, b, α, the coordinates of left and right optical centers $\left(X_L, Y_L, Z_L\right)$ and $\left(X_R, Y_R, Z_R\right)$ in the camera coordinate system XYZ are

$$\begin{cases} X_L = b - (b+d)\cos 2\alpha \\ Y_L = 0 \\ Z_L = d - (b+d)\sin 2\alpha \end{cases} \tag{11}$$

$$\begin{cases} X_R = -b + (b+d)\cos 2\alpha \\ Y_R = 0 \\ Z_R = d - (b+d)\sin 2\alpha \end{cases} \tag{12}$$

respectively. If the left and right images are rectified through the rotation transformation (similar to gaze transformation) so that the "new" optical axes are parallel, then the baseline length is

$$B = X_L - X_R = 2b + 2(b+d)\cos(\pi - 2\alpha) \tag{13}$$

The fixation angle between the original left and right virtual optical axes is

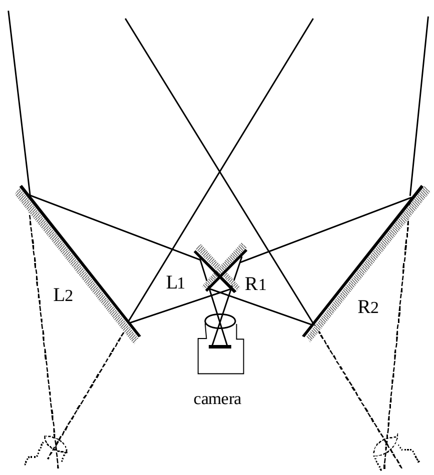
$$\theta = 2\left(90^\circ - 180^\circ + 2\alpha\right) = 4\alpha - 180^\circ \tag{14}$$

and the Z coordinate of the fixation point is

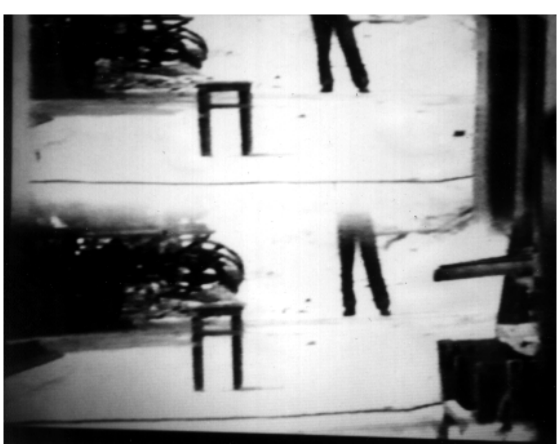
$$Z_f = bctg\,\frac{\theta}{2} + d = b(-tg\,2\alpha) + d \tag{15}$$

Binocular equipment II: up-down images

The modified binocular imaging equipment is designed to satisfy the following requirements of the obstacle detection system. (1) The equivalent image pair form the left and right viewpoints is projected into a single image sensing plane. (2) Viewing angle in the horizontal direction is not reduced so that the system can detect obstacles on a wide road. (3) Binocular images are projected into the up and down halves of a single image plane so that the entire image is fully used, since the up-half of the original monocular image, the projection of the scene above the road, is not important to obstacle detection. Modifying the design in Fig.2 by extending the up-half of the left primary mirror and the down-half of the right primary mirror, the left and right virtual cameras can see the entire field of view (FOV) in the horizontal direction (Fig.3). In order that the up- and down-half of the FOV form the common FOV, the left and /or right auxiliary mirrors should tilt slightly down and/or up respectively. The properties of binocular imaging geometry, baseline length and the fixation point are similar to those of the basic binocular equipment I. The difference is that in the modified equipment the optical centers O_L, O_R of the two virtual cameras are not located in the same height, i.e. $Y_R < 0, Y_L > 0$.



(a) system (top view)



(b) a real image

Fig 3. Modified single-camera binocular equipment: up-down images

Fig. 2(b) and Fig. 3(b) show the real images captured by the basic and modified binocular system respectively.

4. REAL-TIME OBSTACLE DETECTION

The basic real-time obstacle detection (RTSOD) system consists of three parts: the sensor, the gaze transformation and the non-zero disparity filtering algorithm. The sensor is the single-camera binocular imaging equipment given in Section 2. The gaze transformation has been described in Section 1. In this section the non-zero disparity filtering algorithm is given and the RTSOD implementation and performance are discussed.

Basic algorithm

The non-zero disparity filter (NZD filter) is the heart of the RTSOD algorithm. In the existing binocular active vision system [3], the 3D shape of the horopter (surface of zero stereo disparity) is rather complicated (and at worst may not be well defined), since it involves vertical, as well as horizontal, disparities. In the RTSOD system, the gaze transformation creates a zero disparity plane for all the features on the ground plane. In this way the obstacle can be detected easily by the NZD filter. The image difference operation separates points with non-zero disparities from those with zero disparities by examining difference values. Theoretically, all the pixels of the ground plane features will have zero values in the difference image. However in practice this property can not be kept strictly due to the photo-metric and geometric bias. So the NZD Filter is applied to throw away points whose difference values are lower than a certain threshold, and then to filter out the small areas which indicate small disparities. The grouping procedure gathers the non-disparity points into each possible obstacle regions.

The basic algorithm consists of six steps:

Step 1. *Binocular imaging*: An image pair is captured simultaneously by the single camera binocular imaging equipment and a frame grabber.

Step 2. *Gaze transformation*: The reprojected left image L_{pi} is created by re-projecting the left image L_i to the plane of the right image R_i

Step 3. *Image difference*: Difference image D_i is produced by calculate the absolute difference of each corresponding points in L_{pi} and R_i

Step 4. *Non-zero disparity filtering*: Non-zero disparity image B_i is created by first throwing away the points in the D_i whose value is smaller than a certain threshold T , then filtering out the small areas which indicate small disparities.

Step 5, *Obstacle grouping*: Set of obstacle regions $\{O_i, i = 1,\cdots N\}$ is obtained by grouping the connected or nearby points into each region O_i. Sizes of obstacles and distances between obstacles are considered in the grouping procedure.

Step 6: *Parameter Computing*. The location (x, y)□width w, height h is calculated for each obstacle (region) and an uncertainty value is estimated for each one .

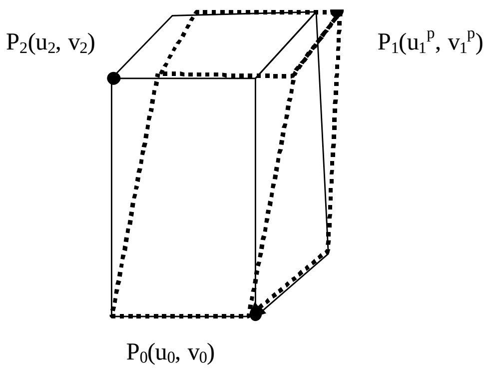


Fig. 4. 3D interpretation for the obstacle region

3D measurements using representative points

Each obstacle region O is the result of binocular difference. In order to avoid the difficult correspondence problem, three representative points are selected to roughly estimate the 3D parameters (location, width and height) of the obstacle. Fig. 4 shows the 3D interpretation for the obstacle region in the difference image. The lowest point $P_0(u_0, v_0)$ of the region O is considered as the ground point. In practice it may be slightly high up from the ground due to the zero-disparity filtering process. This factor can be included to estimate the 3D parameters of the obstacle more accurate. The up-left point $P_2(u_2, v_2)$ is the up-left contour point of the obstacle in the image of the right camera, while the up-right point $P_1(u_1^p, v_1^p)$ is reprojected from the up-right contour point of the obstacle in the image of the left camera. P_1 should be projected back to the original left image coordinates (u_1, v_1) by an inverse gaze transformation before the 3D computing. The 3D coordinates of these three points (x_i, y_i, z_i)(i=0,1,2) can be easily calculated by using the relations between the cameras' and the robot's coordinate system. The location, height and width can be derived then.

Real-time implementation

The RTSOD algorithm is implemented in a PC 486/66MHz with a frame grabber PC-Video 100. For a pair of 256*256*8 gray level image, the total time for a detection period is about 0.5-0.6 seconds, among which 250 ms is for 512*512 image capture and data transfer, 110ms for the gaze transformation, image difference and non-zero disparity filtering, 130-230ms (which varies with number of regions) for region grouping, and 10ms for parameter calculation and image display. The total processing time is reduced under 0.4 seconds in a Pentium/75MHz PC with the same frame grabber. The gaze transformation is carried out by using look-up tables of two-value functions $u_2 = U(u_1, v_1)$ and $v_2 = V(u_1, v_1)$ in case that the vehicle moves smoothly and the normal vector (p, q, r) of the ground plane keeps unchanged. The system still works properly if the normal vector (p, q, r) does not change severely. The work condition is analyzed in the following subsection and the algorithm has been developed for the case when the bump of the vehicle cannot be neglected [9].

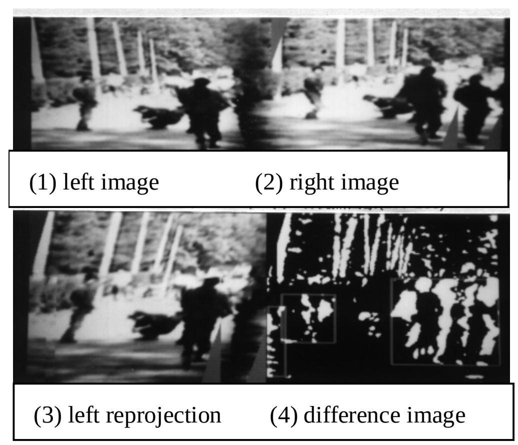


Fig. 5. Example 1: heavy shadows

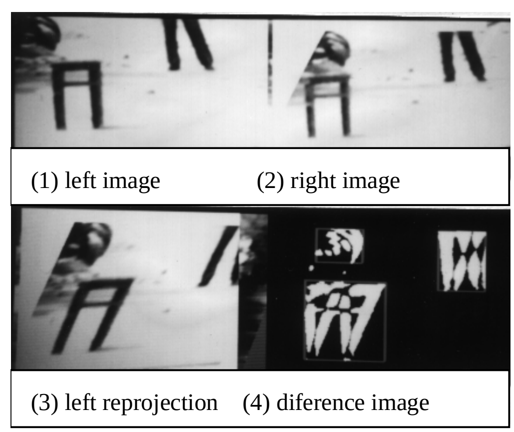


Fig. 6. Example 2: multiple obstacles

Fig.5-7 give some of the examples for the obstacle detection in the outdoor natural scene.. Fig. 5 shows the experimental results of the image in Fig. 2(b) where heavy shadows cover the road. Fig. 6 shows the results of the image in Fig. 3(b) where three obstacles are detected correctly. Fig. 7 shows one of the examples during the real-time operation of the RTSOD system.

Performance analysis

Detectable height: It is interesting to know the minimum height of an obstacle that the system can detect. Suppose at distance D, the detectable height of an obstacle is h. We will analyze the manifest of the point $(x,y,z)=(0,D,h)$ in the left, right and the reprojected images. Explicit results cannot be easily derived under the general case. Without the loss of generality, we suppose the optical axes of both left and right camera are parallel to the ground, and their image planes lie in the same plane ($T_z=0$), then we have $R=I,\ \ T=(T_x,0,0)^t$. Assume the distance from each of the optical center to the ground is H. It can be proved that the "disparity" of the point $(0,D,h)$ in the right image and the reprojected left image is

$$\Delta u = du - du^0 = \frac{FT_x}{DH}h,$$

where du is the actual disparity of $(0,D,h)$ before the gaze transformation and du^0 is the disparity of the ground point $(0,D,0)$. Hence we build up the relationship between the obstacle's detectable height h with the distance D, focal length F and camera's height H:

$$h=\frac{DH}{FT_x}\Delta u \qquad (16)$$

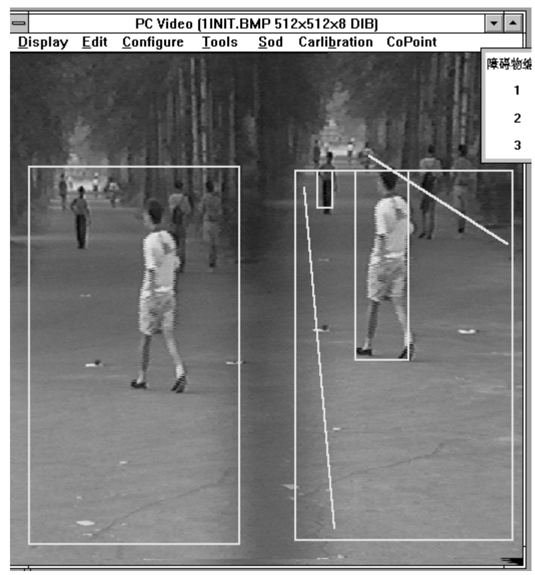
It should be noticed that the magnitude of "disparity" Δu is the measure of the distinctness of an obstacle. For example, given

$$D=30m, H=2.5m, T_x=1m, F=\frac{256\,pixel}{8.8mm}\times 25mm$$

if the minimum distinguished "disparity" is $\Delta u = 2\ pixel$, then the minimum height is about $0.20\,m$. We have the following conclusions.

(1) The obstacle's detectable height is proportion to the image resolution of the discrete image. The smaller is the region in the difference image that can be distinguished, the lower is the detectable height of the obstacle.

(a)Left image, right image, left reprojected image and obstacle regions in the difference image (two obstacle are detected)



(b)The results (two small rectangles) overlay in the original right image. The large rectangles in both the left and right images indicate the areas of interests (AOIs). Obstacles are detected in the intersection region of the two AOIs and within the road region indicated by two white lines in the right image

障碍物编号	位置(米)	距离(米)	宽度(米)	高度(米)	SAD
1	-0.11	22.34	0.69	1.38	100.0
2	-0.94	82.08	0.66	1.13	37.6
3	×	×	×	×	×

(c)Obstacle parameter list: Number, position (m), distance (m), width (m), height (m) and the belief measure (%) estimated from the sum of absolute difference (SAD) in each obstacle region. The actual distance and height of the first obstacle (a pedestrian) are 23.0 m and 1.7m respectively. The baseline is 0.34 m.

Fig. 7. An operation example of RTSOD system

(2) For the same height h, the farther be the distance D, the smaller is the "disparity" Δu. Therefore small obstacle is easier to detect in near-distance.

(3) For the same height h, the disparity Δu increases by the increasing of the focal length and /or the baseline length T_x, however the (common) field of view of the two images is reduced.

(4) It is interesting to find that the obstacle's minimum detectable height is also proportional to the camera's height H. The higher is the height H, the higher is the minimum detectable height. For this reason the camera should be installed as low as possible.

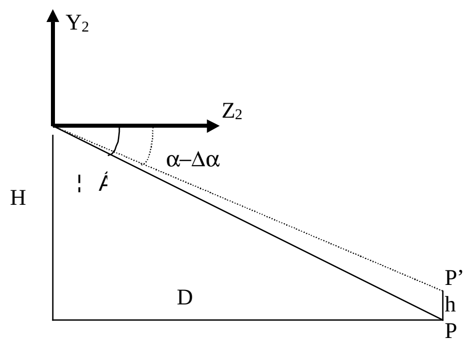


Fig.8. Tolerance of the vehicle's bump

Tolerance of the vehicle's bump: We also consider the simplified model, i.e. the optical axes of the cameras are parallel to the ground plane. For a ground point P at distance D, the angle between the projection ray O_2P and the optical axis Z_2 of the right camera is $\alpha = \tan^{-1}\frac{H}{D}$. Suppose the camera tilts down an angle $\Delta\alpha$ due to the bump of the vehicle, it is approximately equivalent to the situation that the point P moves up to point P' of height h, so that the angle between O_2P' and Z_z becomes $\alpha - \Delta\alpha$, therefore we have

$$h = H - D\tan(\alpha - \Delta\alpha) = D[\tan\alpha - \tan(\alpha - \Delta\alpha)]$$

The tolerance tilted angle range $\Delta\alpha$ can be estimated from $h \le h_{\min}$ (the minimum detectable height)

$$|\Delta\alpha| \approx |\sin\Delta\alpha| \le h_{\min}\cos^2\alpha = \frac{D}{D^2 + H^2}h_{\min} \qquad (17)$$

Given $H = 2.5m, D = 30m, h_{\min} = 0.40m$, we have $|\Delta a| \le 0.8^\circ$. It means that if the tilted angle is larger than the tolerance, the figures (e.g. shadows on the road) on the road may be misdeemed as obstacles. However if orientation of the ground textures are perpendicular to the road direction (i.e. x axis of the camera), the negative effect of tilted bump of the vehicle will greatly reduced.

5. DISCUSSIONS AND CONCLUDING REMARKS

In the basic RTSOD algorithm, the operation condition is that the vehicle moves smoothly, i.e. the tilted angle due to the vehicle's bump is smaller than the tolerance value. So that the gaze transformation can be realized with look-up tables with known camera posture parameters. However mistakes may occur if the vehicle bumps severely on a textured road surface. In this case the tilted angle should be measured whenever necessary in order to modify the parameters of the gaze transformation. We have proposed the dynamic gaze transformation algorithm, which can be implemented in near real time [9]. The basic idea is to establish the relation between the tilted angle and the gaze parameters, and find the best parameters for zero-disparities of the ground plane.

The single-camera stereo equipment and the obstacle detection system can be used in smart cars, traffic monitoring, surveillance, virtual reality and medical imaging. The novel gaze transformation is a powerful pre-processing method for 3D scene modeling[10] and motion analysis. It can greatly simplify the correspondence problem of stereo vision and the tracking problem in visual motion.

6. ACKNOWLEDGMENTS

The authors would like to thank Prof. Edward M. Riseman and Prof. Allen R. Hanson of University of Massachusetts at Amherst for their valuable discussions. Thanks are also given to Dr. Hong Heather Yu of Princeton University for her help in the presentation of the manuscript.

7. REFERENCES

[1] Thorpe C, et al. Vision and Navigation for the CMU Navlab. *IEEE Trans. PAMI*, 1988, 10(3): 362 ~ 376

[2] Kayaalp A, Eckman J L, Near real-time range detection using a pipeline architecture, *IEEE Trans SMC*, vol 20, no 6, Nov/Dec 1990.

[3] Coombs D and Brown C, Real-time binocular smooth pursuit, *Int J Computer Vision*, vol. 11, no. 2, 1993: pp 147-164.

[4] Badal S, Ravela S, Draper B, Hanson A, A practical obstacle detection and avoidance system, In: *Proc. IEEE Workshop on Applications of Computer Vision*, 1994.

[5] Storjoham K, et al, Visual obstacle detection for automatically guided vehicles. In: *Proc IEEE ICRA*, pp761-766, 1990.

[6] Zhu Z G. Lin X Y, Realtime algorithms for obstacle avoidance by using reprojection transformation. In: *Proc IAPR Workshop on Machine Vision Application*, 1990. 393-396

[7] Cho Y C, Cho H S, A Stereo Vision-based Obstacle Detection method for mobile robot navigation, *Robotica*, 12,1994:203-216

[8] Shi D J, Zhu Z G, Chen Y K, Real-time stereo imaging equipment, *Chinese Novel Practical Patent* No ZL95211525.5, 1995.12.

[9] Zhu Z G, *Environment Modeling for Visual Navigation*, Ph.D. Dissertation , Tsinghua University, May 1997 (in Chinese)

[10] Lin X Y, Zhu Z G, Den W, "A Stereo Matching Algorithm Based on Shape Similarity for Indoor Environment Model Building", *Proc.* 13th *IEEE International Conference Robotics and Automation*, 1996: pp 765-770