

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.

For more information visit www.intechopen.com



Stereo Matching Method and Height Estimation for Unmanned Helicopter

Kuo-Hsien Hsia¹, Shao-Fan Lien² and Juhng-Perng Su²

¹*Far East University*

²*National Yunlin University of Science & Technology
Taiwan*

1. Introduction

The research and development of autonomous unmanned helicopters has lasted for more than one decade. Unmanned aerial vehicles (UAVs) are very useful for aerial photography, gas pollution detection, rescue or military applications. UAVs could potentially replace human beings in performing a variety of tedious or arduous tasks. Because of their ubiquitous uses, the theory and applications of UAVs systems have become popular contemporary research topics. There are many types of UAVs with different functions. Generally UAVs can be divided into two major categories, fixed-wing type and rotary-wing type. The fixed-wing UAVs can carry out long-distance and high-altitude reconnaissance missions. However, flight control of fixed-wing UAVs is not easy in low-altitude conditions. Conversely, rotary-wing UAVs can hover in low altitude while conducting surveys, photography or other investigations. Consequently in some applications, the rotary-wing type UAVs is more useful than the fixed-wing UAV. One common type of rotary-wing type UAVs is the AUH (Autonomous Unmanned Helicopter). AUHs have characteristics including of 6-DOF flight dynamics, VTOL (vertical taking-off and landing) and the ability to hover. These attributes make AUHs ideal for aerial photography or investigation in areas that limit maneuverability.

During the past few years, the development of the unmanned helicopter has been an important subject of research. There have been a lot of researches interested in a more intelligent design of autonomous controllers for controlling the basic flight modes of unmanned helicopters (Fang et al., 2008). The controller design of AUHs requires multiple sensor feedback signals for sensing states of motion. The basic flight modes of unmanned helicopters are vertical taking-off, hovering, and landing. Because the unmanned helicopter is a highly nonlinear system, many researchers focus on the dynamic control problems (e.g. Kadmiry & Driankov, 2004; C. Wang et al., 2009). Appropriate sensors play very important roles in dynamic control problems. Moreover, the most important flight mode of autonomous unmanned helicopter is the landing mode. In consideration of the unmanned helicopter landing problem, the height position information is usually provided by global positioning system (GPS) and inertial measurement unit (IMU). The system of the autonomous unmanned helicopter is a 6-DOF system, with 3-axis rotation

information provided by IMU and 3-axis moving displacement information provided from GPS.

Oh et al. (2006) brought up the tether-guided method for autonomous helicopter landing. Many researches used vision systems for controlling helicopter and searching landmark (Lin, 2007; Mori, 2007; C.C. Wang et al., 2009). In the work of Saito et al. (2007), camera-image based relative pose and motion estimation for unmanned helicopter were discussed. In the works of Katzourakis et al. (2009) and Xu et al. (2006), navigation and landing with the stereo vision system was discussed. Xu et al. used the stereo vision system for estimating the position of the body. From the work of Xu, it was shown that the stereo vision does work for the position estimation.

For unmanned helicopter autonomous landing, the information of the height is very important. However, the height error of GPS is in general about from 5 to 8 meters, which is not accurate enough for autonomous landing. For example, the accuracy of Garmin GPS 18-5Hz is less than 15 meters (GPS 18 Technical Specifications, 2005). After many times of measurement, the average error of this GPS was obtained to be around 10 meters. Since the height error range of GPS is from 5 to 8 meters, to conquer the height measurement error of GPS, the particular stereo vision system is designed for assisting GPS, and the measurement range of this system is set to be at least 6 m.

Image systems are the common guiding sensors. In the AUHs control problems, image systems are usually collocated with IMU and GPS in the outdoor environment. The image system has been used on vehicles for navigation, obstacle avoidance or position estimation. Doehler & Korn (2003) proposed an algorithm to extract the edge of the runway for computing the position of airplane. Bagen et al. (2009) and Johnson et al. (2005) discussed the image-guided method with two or more images for guiding the RC unmanned helicopter approaching to the landmark. Undoubtedly multiple-camera system measurement environment is an effective and mature method. However, the carrying capacity of a small unmanned helicopters has to be considered. Therefore the image systems are the smaller the better. A particular stereo vision system is developed for reducing the payload in our application.

In this chapter, we focus on the problem of estimating the height of the helicopter for the landing problem via a simple stereo vision system. The key problem of stereo vision system is to find the corresponding points in the left image and the right image. For the corresponding problem of stereo vision, two methods will be proposed for searching the corresponding points between the left and right image. The first method is searching corresponding points with epipolar geometry and fundamental matrix. The epipolar geometry is the intrinsic projective geometry between two cameras (Zhang, 1996; Han & Park, 2000). It only depends on the camera internal parameters and relative position. The second method is block matching algorithm (Gyaourova et al., 2003; Liang & Kuo, 2008; Tao et al., 2008). The block matching algorithm (BMA) is provided for searching the corresponding points with a low resolution image. BMA will be compared with epipolar geometry constraint method via experimental results.

In addition, a particular stereo vision system is designed to assist GPS. The stereo vision system composed of two webcams with resolutions in 0.3 mega pixels is shown in Figure 1. To simplify the system, we dismantled the cover of the webcams. The whole system is very

light and thin. The resolution of cameras will affect the accuracy of height estimation result. The variable baseline method is introduced for increasing the measuring range. Details will be illustrated in the following sections.



Fig. 1. The stereo vision system composed of two Logitech® webcams

2. Design of stereo vision system

2.1 Depth measuring by triangulation

In general, a 3D scenery projected to 2D image will lose the information of depth. The stereo vision method is very useful for measuring the depth. The most common used method is triangulation.

Consider a point $P=(X, Y, Z)$ in the 3D space captured by a stereo vision system, and the point P projected on both left and right images. The relation is illustrated in Figure 2. In Figure 2, the projected coordinates of point P on the left and the right images are (x_l, y_l) and (x_r, y_r) respectively. The formation of the left image is:

$$\frac{X}{Z} = \frac{x_l}{f} \Rightarrow X = \frac{x_l}{f} Z \quad (1)$$

and the formation of the right image is:

$$\frac{X-b}{Z} = \frac{x_r}{f} \quad (2)$$

From (1) and (2), we have

$$Z = \frac{fb}{(x_l - x_r)} = \frac{fb}{\Delta x} \quad (3)$$

where f is focal length, b is the length of baseline and $\Delta x = (x_l - x_r)$ is the disparity. From (3), the accuracy of f , b and Δx will influence the depth measuring. In the next section, the camera will be calibrated for obtaining accurate camera parameters.

There are three major procedures for stereo vision system design. Firstly, the clear feature points in image need be extracted quickly and accurately. The second procedure is searching for corresponding points between two images. Finally, computing the depth using (3).

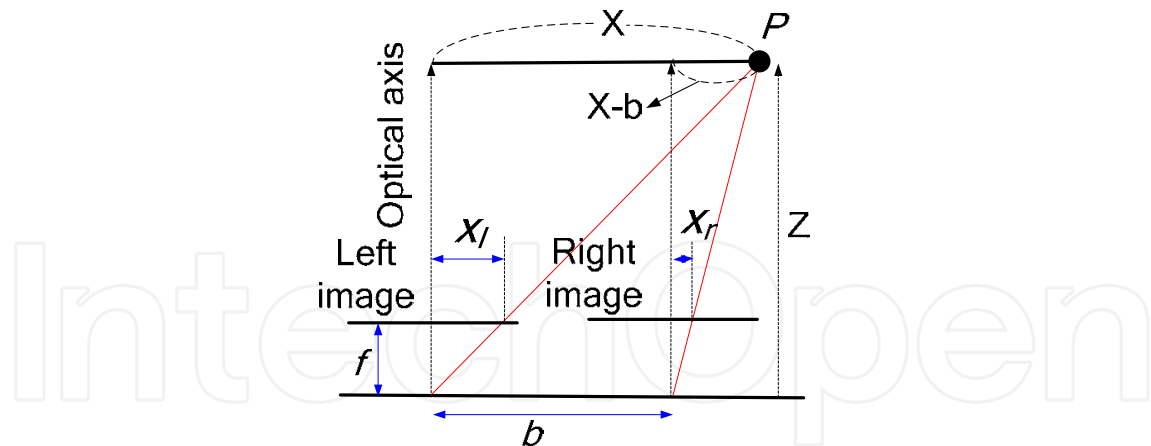


Fig. 2. Geometric relation of a stereo vision system.

2.2 Depth resolution of stereo vision system

The depth resolution is a very important factor for stereo vision system design (Cyganek & Siebert, 2009). The pixel resolution will reduce with the depth. The relations of depth resolution is illustrated in Figure 3.

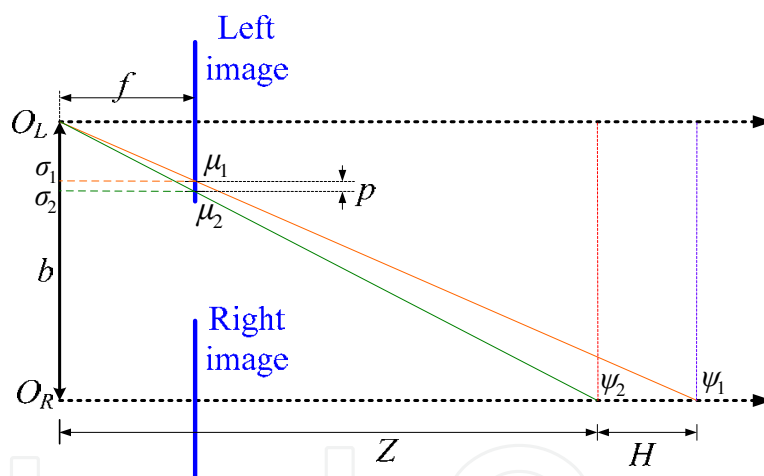


Fig. 3. Geometry relation of depth resolution.

From Figure 3, with the similarity of triangle $\Delta O_L \mu_1 \sigma_1$ to $\Delta O_L \psi_1 O_R$ and $\Delta O_L \mu_2 \sigma_2$ to $\Delta O_L \psi_2 O_R$, we can have the following relations:

$$\frac{\overline{O_L O_R}}{\overline{O_R \psi_1}} = \frac{\overline{O_L \sigma_1}}{\sigma_1 \mu_1}, \frac{\overline{O_L O_R}}{\overline{O_R \psi_2}} = \frac{\overline{O_L \sigma_2}}{\sigma_2 \mu_2} \quad (4)$$

where $\overline{O_L O_R} = b$, $\overline{O_L \sigma_1} = \overline{O_L \sigma_2} = f$, $\overline{O_R \psi_2} = Z$ and $\overline{O_R \psi_1} = H$. Then we obtain

$$p = \overline{O_L \sigma_2} - \overline{O_L \sigma_1} = \frac{fb}{Z} - \frac{fb}{Z+H} \quad (5)$$

where p is the width of a pixel on image. Next, we can have the following equation by rearranging (5):

$$H = \frac{pZ^2}{fb - pZ} \quad (6)$$

where H is the depth-change when a pixel change in the image, and is called the pixel resolution. Assuming $fb / Z \gg H$, the following approximation will be obtained,

$$H \approx \frac{pZ^2}{fb} \quad (7)$$

In addition, the (6) is true with the condition:

$$fb \neq pZ \quad (8)$$

Therefore the limit value of Z is

$$Z = \frac{fb}{p} \quad (9)$$

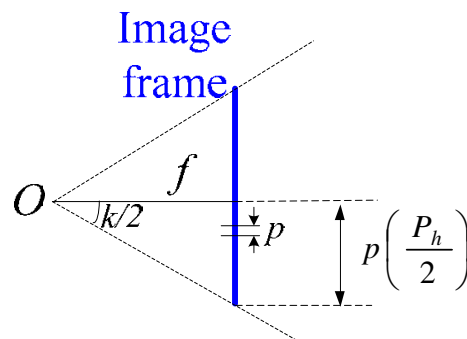


Fig. 4. Geometry relation of f and p .

For single image, the f , b and p are all constants. Thus there is no depth information from a single image. Furthermore, consider Figure 4, and we will have:

$$f = p \left(\frac{P_h}{2 \tan(k/2)} \right) \quad (10)$$

where k is the horizontal view angle, P_h is the horizontal resolution of the camera.

Combining (6) with (10), we will have:

$$H = \frac{Z^2}{\left[P_h b / 2 \tan(k/2) \right] - Z} \quad (11)$$

From (11), the relation between baseline and pixel resolution are shown in Figure 5. Obviously, the pixel resolution and baseline are in a nonlinear relation. Moreover, they are almost in

inverse proportion. The accuracy of system depends on choosing an appropriate baseline. In general, if small pixel resolution is expected, one should choose a larger baseline.

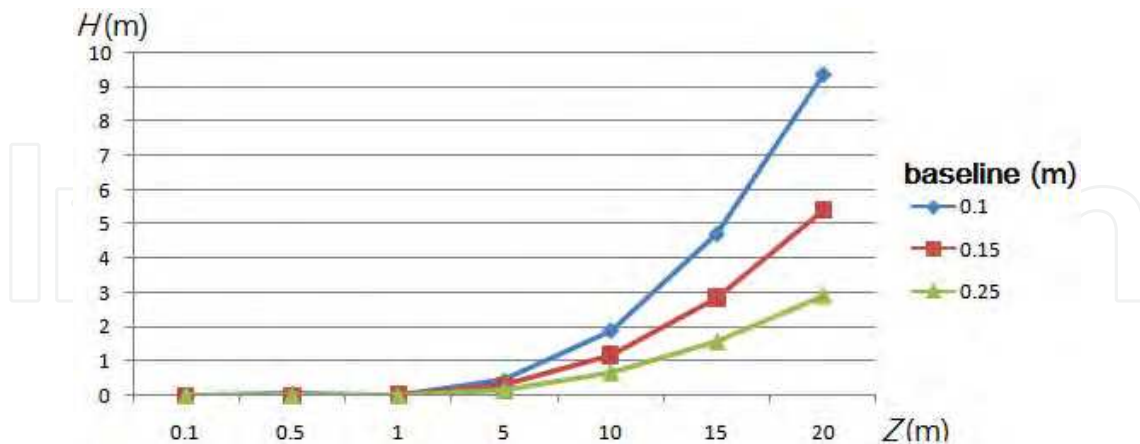


Fig. 5. The pixel resolution H with different baseline for stereo vision system setup.

3. Searching for corresponding points

The stereo vision system includes matching and 3D reconstruction processes. The disparity estimation is the most important part of the stereo vision. The disparity is computed by matching method. Furthermore, the 3D scene could be reconstructed by disparity. The basic idea of disparity estimation is using the pixel intensity of a point and its neighborhood on an image as a matching template to search the most matching area on another image (Alagoz, 2008; Wang & Yang, 2011). The similarity measurement between two images is defined by correlation functions. Based on different matching unit, there are two major categories of matching method which will be discussed. They are area-based matching method and feature-based matching method.

3.1 Area-based matching method

A lot of area-based matching methods have been proposed. Using area-based matching methods, one can obtain the dense disparity field without detecting the image features. Generally, the matching method has good results with flat and complex texture images. Template matching method and block matching method are relatively prevalent methods of the various area-based matching methods. Hu (2008) proposed the adaptive template for increasing the matching accuracy. Another example is proposed by Siebert et al. (2000). This approach uses 1D area-based matching along the horizontal scanline. Figure 6 illustrates the 1D area-based matching. Bedekar and Haralick (1995) proposed the searching method with Bayesian triangulation. Moreover, Tico et al. (1999) found the corresponding points of fingerprints with geometric invariant representations. Another case is area matching and depth map reconstruction with the Tsukuba stereo-pair image (Cyganek, 2005, 2006). In this case, the matching area is 3×3 pixels and the image size is 344×288 pixels (download from <http://vision.middlebury.edu/stereo/eval/>). The disparity and depth map are reconstructed and the depth information in the 3D scene are obtained. The results are illustrated in Figure 7.

However, there are still some restrictions for area-based matching method. Firstly, the matching template is established with pixel intensity, therefore the matching performance

are dependent on brightness, contrast, and textures. If the brightness is changed a lot or textures are monotonous, the matching performance will not be good. Secondly, the matching results will not be good when the image with depth discontinuity or masking. Last, the computational complexity is very high. Therefore, the feature-based matching methods are developed for improving the defects of area-based matching method.

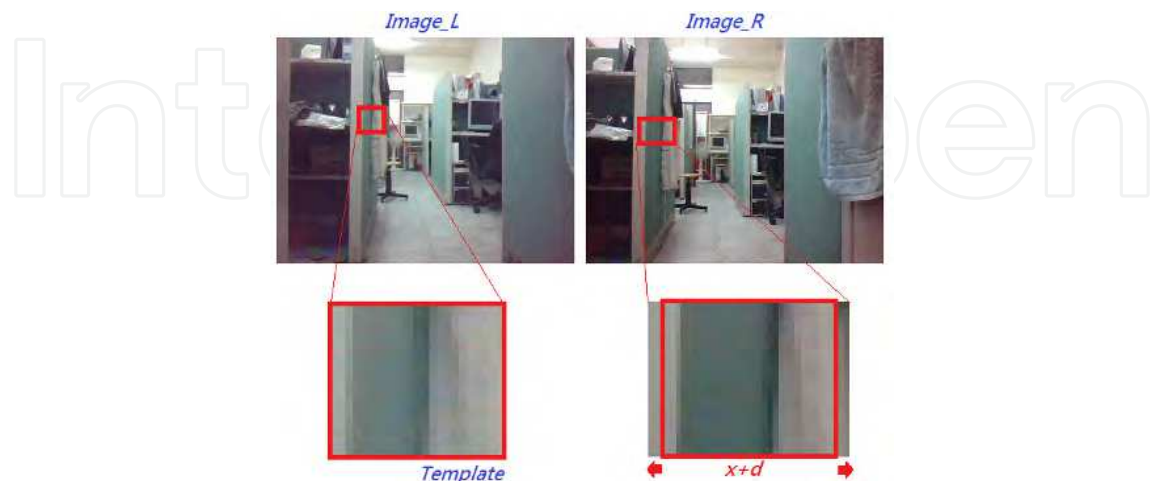


Fig. 6. 1D area-based matching along the horizontal scanline.

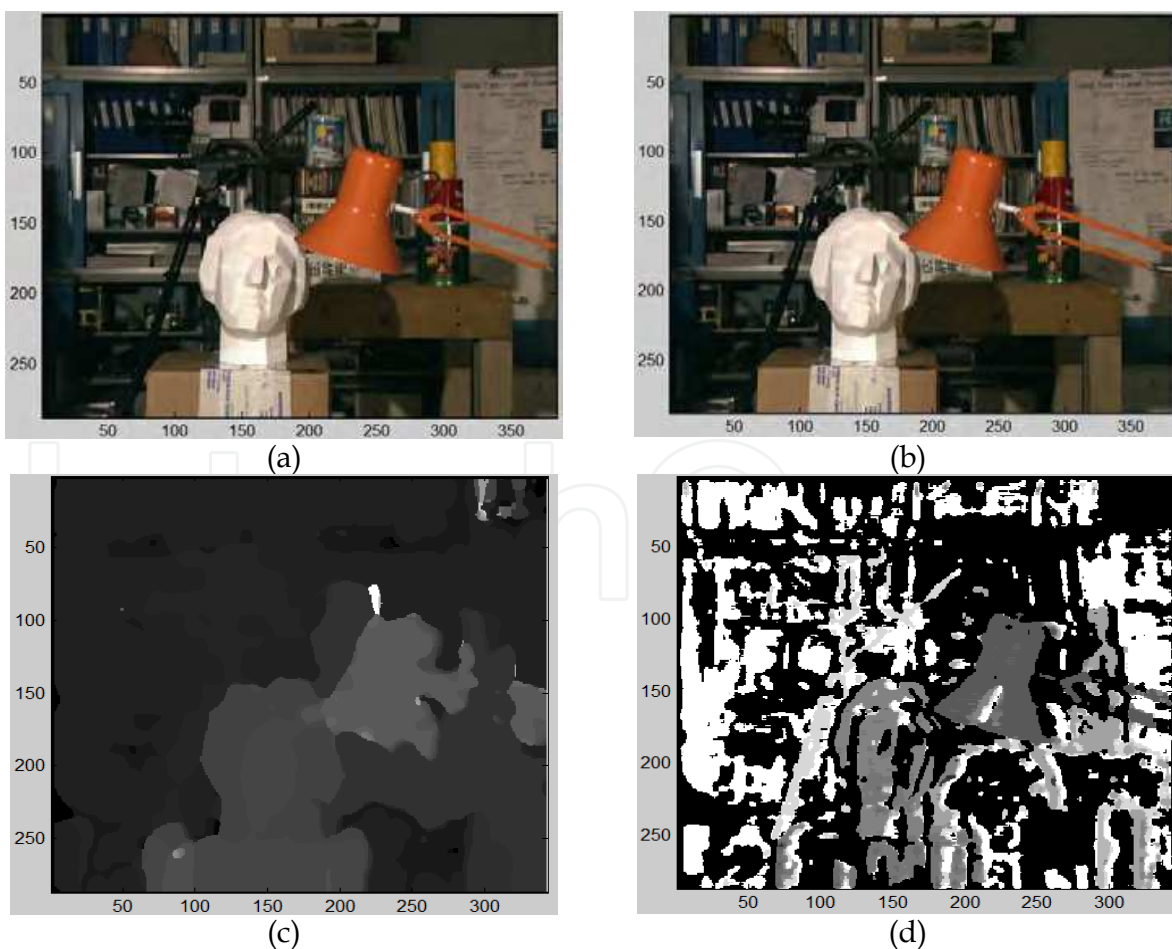


Fig. 7. (a) Left image. (b) Right image. (c) Disparity map. (d) Depth map.

3.2 Feature-based matching method

The feature-based matching method is matching the corresponding points on image with the features of the scene. To highlight the information of space, the features are more easily than the pixel intensity of area. Moreover, feature-based matching method is more robust than area-based method for brightness changing. There are two steps for the feature-based matching method. They are feature extraction and feature matching. The features are usually the lines, corners or planes in the image. The specific operators are utilized for extracting the features.

Many feature-based matching methods have been proposed for searching the feature correlation between the right and left images. Both the intensity and orientation of the features could be the matching templates for searching the correspondence of the features. Therefore, for the depth discontinuity or masking problems, the feature-based matching method can obtain better matching result. In addition, feature-based matching method computes only for the features instead of all pixels, hence the computing load is smaller than area-based matching method. Olson (2002) proposed the matching method based on statistics. This method extracted a few eigenvectors as the matching templates, and it used the maximum-likelihood for template matching.

Moreover, the phase-based image matching is also a very accurate matching method (Muquit et al., 2006). The images are transformed into frequency domain by 2D Discrete Fourier Transforms (2D DFTs). The best matching vector is obtained by computing the phase correlation function. Figure 8 is a simulation of phase-based image matching. The example image is “Cristo Redentor” in Brazil and the image sizes are both 119×127 pixels. Figure 8(c) is the phase correlation of Figures 8(a) and 8(b). From Figure 8(c), we can see that the peak is located at (111, 126), and hence the motion vector is (8, 1).

Similarly, feature-based matching method have two restrictions. Firstly, the dense disparity field could not be obtained. Therefore, it is not easy to reconstruct the complex 3D scene. Secondly, the matching performance is affected by feature extraction results directly. In other words, if the features are too sparse, the matching results will not be good.

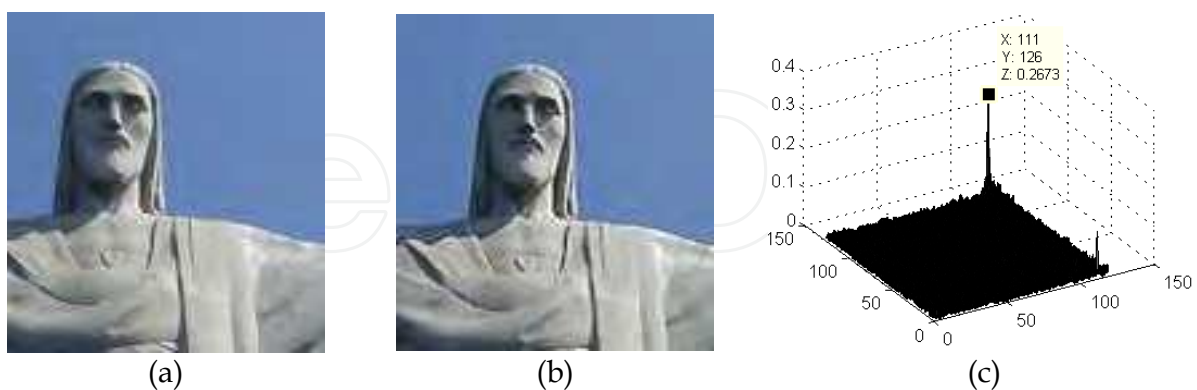


Fig. 8. Case simulation of phase-based image matching.

Since both the area-based and feature-based matching methods have some restrictions, the hybrid matching algorithm has been proposed in recent years. For example, Fuh et al. (1989) combined the optic flow and block matching algorithm for increasing the matching performance. In this chapter, we will combine the feature points and epipolar geometry constraint for reducing the computation.

3.3 Feature points detection

There are usually millions pixels in an image, therefore how to extract the significant feature points is the interesting research topic. Including edge detecting method (Canny, 1986), Tabu search algorithm (Glover, 1989, 1990), neural network (NN) (Takahashi et al., 2006) or Hough transform (Duan et al., 2010) are useful methods for extracting the special features from an image. However, the point sets of lines or edges are still too large. In addition, the searching speeds of most feature extracting algorithms, such as Tabu search, are not fast enough for real-time stereo vision systems. Consequently, the Harris corner detector (Nixon & Aguado, 2008) is proposed for detecting the feature points.

The main principle of Harris corner detector is using the Gaussian filter to detect the cornersresponse of each pixel in the image. Gaussian filter can not only enhance the significant corners, but also remove the undesirable corners. Moreover, it can reduce the probability of misjudgment. Although the Harris corner detector is a very useful tool, it requires a lot of computing time. Therefore the corner detection operations are applied only for the basic rectangle to reduce computation time.

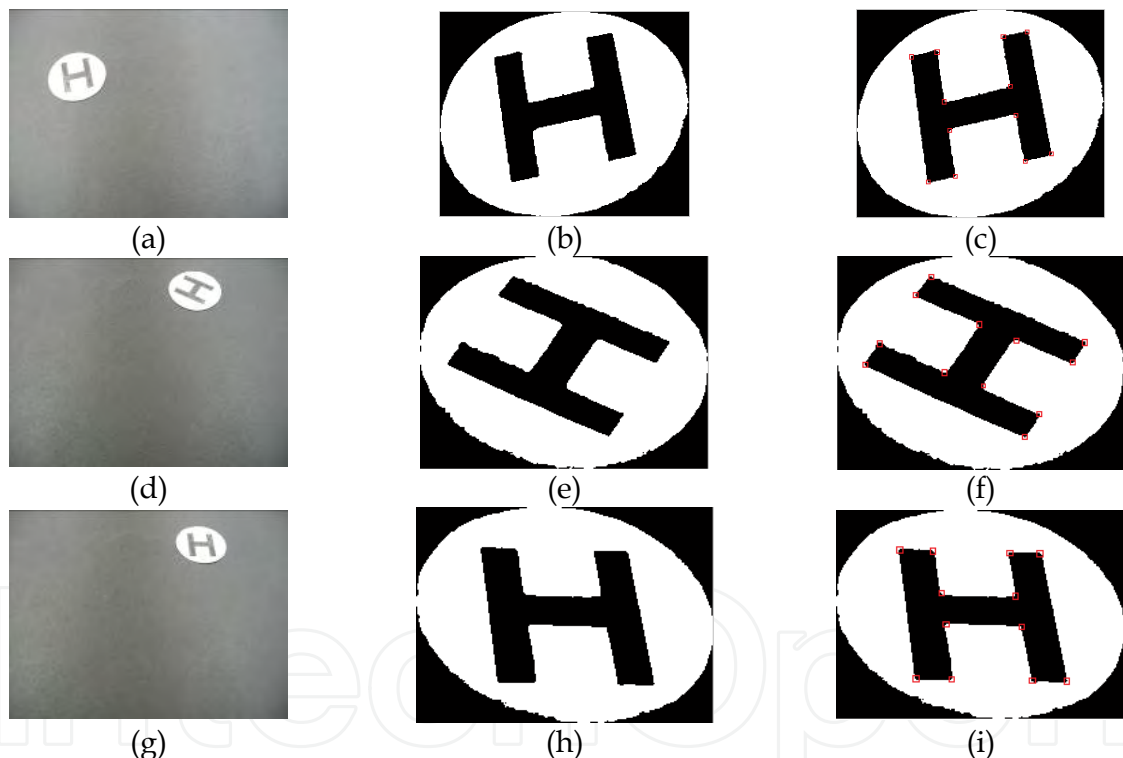


Fig. 9. The simulations of the image processing and feature extraction results. (a) The test image 1. (b) The binary image and the basic rectangle of test image 1. (c) The corner detection result of test image 1. (d) The test image 2. (e) The binary image and the basic rectangle of test image 2. (f) The corner detection result of test image 2. (g) The test image 3. (h) The binary image and the basic rectangle of test image 3. (i) The corner detection result of test image 3.

The following demo example shows the results of landmark image corner detection. The procedure is described as follows.

Step 1. Detect the corners of the image in the basic rectangle.

Step 2. Detect the convex corners of the label 'H'.

Step 3. Label the convex corners and extract the four most outside corners of the label 'H'.

Step 4. Find the intersection of the diagonals, and designate it as the approximate center of the landmark image.

Some examples for demonstrating the Harris corner detector are illustrated in Figure 9. Several advantages can be summarized from Figure 9. The Harris corner detector is very robust for corner detection. Moreover the Harris corner detector could not only detect the edges but also the corners of the object. The detection procedure with the basic rectangle segmentation can greatly enhance the detecting efficiency.

3.4 Epipolar geometry constraints

The epipolar geometry is the intrinsic projective geometry between two cameras. It only depends on the camera internal parameters and relative position. The fundamental matrix F is the algebraic description of the epipolar geometry. Epipolar geometry between two views is illustrated in Figure 10.

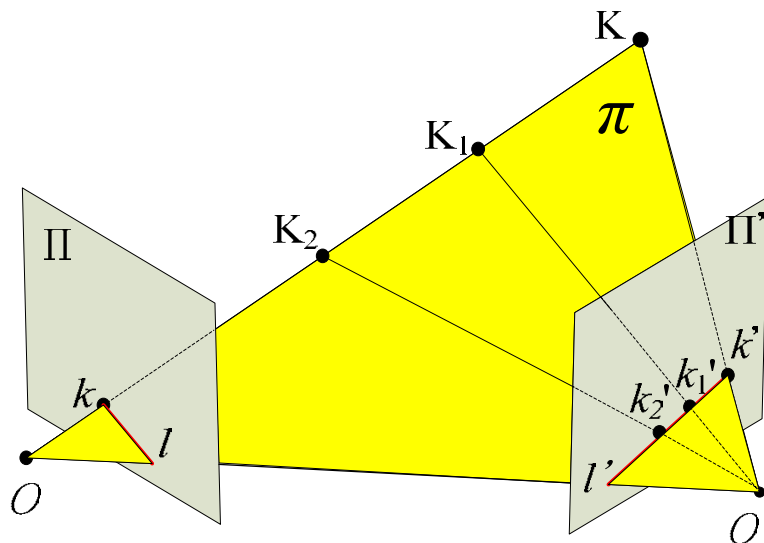


Fig. 10. Epipolar geometry.

From Figure 10, a point K in the 3D space is projected on Π and Π' respectively. The point k in Π corresponds to an epipolar line l' line on Π' , and it can be represented as:

$$l' = Fk \quad (12)$$

where F is fundamental matrix, $k = (x, y)$ and $l' = (a \ b \ c)^T$. The point $k' = (x', y')$ lies on l' , then:

$$k'^T l' = k'^T Fk = 0 \quad (13)$$

Rewrite equation (13) as:

$$[x' \ y' \ 1] \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0 \quad (14)$$

Based on Hartley's 8-point algorithm (Hartley, 1995), (14) can be further represented as:

$$xx'f_{11} + xy'f_{21} + xf_{31} + yx'f_{12} + yy'f_{22} + yf_{32} + x'f_{13} + y'f_{23} + f_{33} = 0 \quad (15)$$

Rewrite (15) as:

$$\begin{bmatrix} x_1x'_1 & x_1y'_1 & x_1 & y_1x'_1 & y_1y'_1 & y_1 & x'_1 & y'_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_nx'_n & x_ny'_n & x_n & y_nx'_n & y_ny'_n & y_n & x'_n & y'_n & 1 \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix} = 0 \quad (16)$$

With 8 points, (16) can be solved.

Under epipolar geometry constraint, searching of the corresponding points will be reduced from 2D image to one line.

3.5 Block matching algorithm (BMA)

The key point of the stereo vision is how to search the corresponding points quickly and effectively. Epipolar geometry constraint is the typical skill for finding the corresponding points. However, movement of stereo vision systems will cause the images blurred. Moreover, since the resolution of webcam is pretty low, the task of searching the corresponding points becomes difficult. Here we will apply the block matching algorithm (BMA) for searching the corresponding points.

BMA is a standard technique for encoding motion in video sequences. It aims at detecting the similar block between two images. The matching efficiency depends on the chosen block size and search region. It is not easy to choose an appropriate block size. Usually, bigger blocks are less sensitive to the noise but will spend more computation time. Fast matching methods for searching the corresponding points have been proposed in some representative studies (e.g. Tao, 2008). Tao's method is to match the reference block in the left image and the candidate block on the epipolar line in the right image. Here we need more accurate matching results for finding out the corresponding points. Therefore, the full-search (FS) is used for achieving better searching results.

In Figure 11, the sum of absolute difference (SAD) is used for the block similarity measuring. The first pixel of the chosen block on left image is (x, y) and the block size is $N \times N$. The search region on the right image is defined to be a rectangular with the image width as the width and height of $2k$. Since the left and right cameras of the stereo vision system are placed on the same line, the k can be small in order for reducing the computation time.

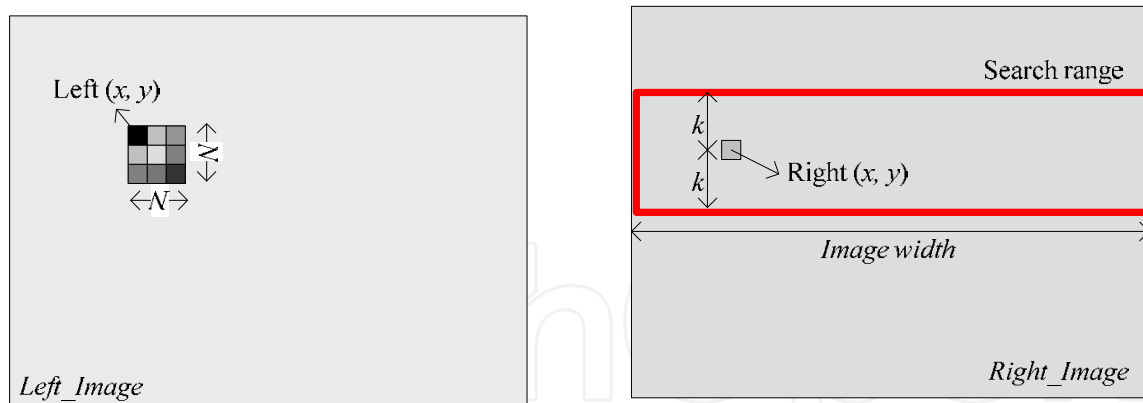


Fig. 11. Block template and the search region.

3.6 Matching cost function

The matching cost function is a norm to represent the degree of correctness of a match. The smaller the matching cost, the higher the correctness. The sum of squared differences (SSD) and normalized cross correlation (NCC) are frequently used matching cost functions other than SAD. Functions of SAD, SSD, and NCC are illustrated as (17-19).

$$SAD(x, y, r, s) = \sum_{i=0}^{i=m} \sum_{j=0}^{j=n} \left| R_Image_{(x+i, y+j)} - L_image_{((x+r)+i, (y+s)+j)} \right| \quad (17)$$

$$SSD = \sum_{i=0}^{i=m} \sum_{j=0}^{j=n} \left(R_Image_{(x+i, y+j)} - L_image_{((x+r)+i, (y+s)+j)} \right)^2 \quad (18)$$

$$NCC = \frac{\sum_{i=0}^{i=m} \sum_{j=0}^{j=n} \left(R_Image_{(x+i, y+j)} \right) \left(L_image_{((x+r)+i, (y+s)+j)} \right)}{\sqrt{\sum_{i=0}^{i=m} \sum_{j=0}^{j=n} \left(R_Image_{(x+i, y+j)} \right)^2 \sum_{i=0}^{i=m} \sum_{j=0}^{j=n} \left(L_image_{((x+r)+i, (y+s)+j)} \right)^2}} \quad (19)$$

where m and n are the length and width of the block, (x, y) is the position of the block on right image and (r, s) denotes the motion vector.

Comparing these matching cost functions to norms in algebra, the SAD is analogous to an 1-norm in algebra, and the SSD is analogous to a 2-norm in algebra. And the NCC uses an inner-product-like operation. Obviously the computation complexity of SAD is lower than the other matching functions. The SAD is the most frequently used matching cost function in applications since it is one of the more computationally efficient methods. The advantage of SAD has been mentioned in literatures (Humenberger, 2010; Point Grey Research Inc., 2000; Bradski, 2010). In our application, the computation time is an important factor, therefore the SAD will be the matching cost function for searching the correspondence.

4. Camera calibration

A camera calibration is done to build the relationship between the world coordinates and their corresponding image coordinates. Consider a pinhole camera model (David & Ponce, 2002) as shown in Figure 12. The camera parameters can be distributed into intrinsic parameters and extrinsic parameters (Hartley & Zisserman, 2003).

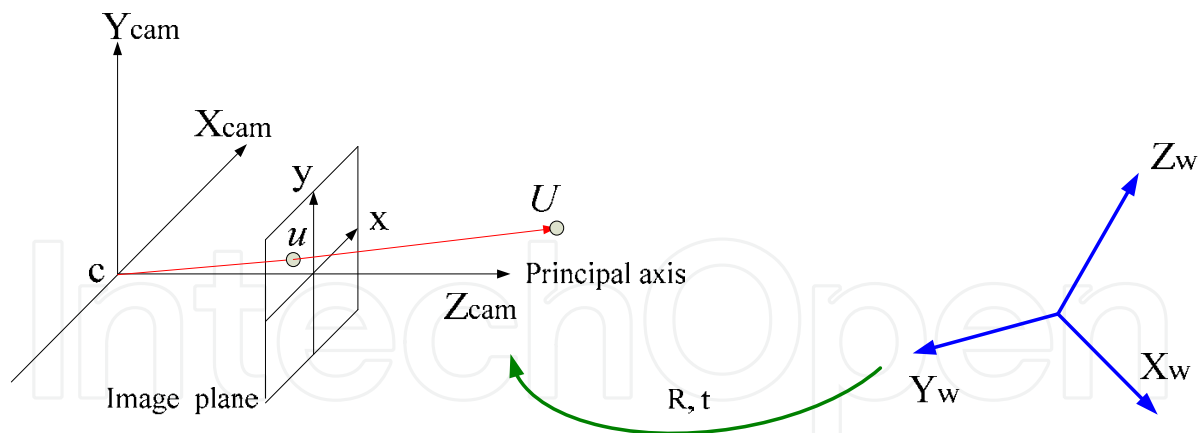


Fig. 12. Pinhole camera model and projective transformation between world and camera coordinate system.

Based on the collinearity equation and the pinhole camera model (Luhmann et al., 2007), the transformation between image point and reference point can be represented as

$$\begin{aligned} q &= \frac{L_1X + L_2Y + L_3Z + L_4}{L_9X + L_{10}Y + L_{11}Z + 1} \\ \tilde{q} &= \frac{L_5X + L_6Y + L_7Z + L_8}{L_9X + L_{10}Y + L_{11}Z + 1} \end{aligned} \quad (20)$$

where (q, \tilde{q}) is a reference point on the image. Re-arrange equation (20), we will have:

$$\begin{aligned} L_1X + L_2Y + L_3Z + L_4 - qL_9X - qL_{10}Y - qL_{11}Z &= q \\ L_5X + L_6Y + L_7Z + L_8 - \tilde{q}L_9X - \tilde{q}L_{10}Y - \tilde{q}L_{11}Z &= \tilde{q} \end{aligned} \quad (21)$$

Equation (21) is equivalent to

$$\begin{bmatrix} X & Y & Z & 1 & 0 & 0 & 0 & 0 & -qX & -qY & -qZ \\ 0 & 0 & 0 & 0 & X & Y & Z & 1 & -\tilde{q}X & -\tilde{q}Y & -\tilde{q}Z \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ L_3 \\ L_4 \\ L_5 \\ L_6 \\ L_7 \\ L_8 \\ L_9 \\ L_{10} \\ L_{11} \end{bmatrix} = 0 \quad (22)$$

The coefficients $L_1 \sim L_{11}$ are called DLT parameters. In order to solve (22), at least 6 control points are required.

The special orientation of camera can be reconstructed by DLT parameters. The principal point (q_0, \tilde{q}_0) is

$$\begin{aligned} q_0 &= J^2(L_1L_9 + L_2L_{10} + L_3L_{11}) \\ \tilde{q}_0 &= J^2(L_5L_9 + L_6L_{10} + L_7L_{11}) \end{aligned} \quad (23)$$

where $J = -\left(\sqrt{K_9^2 + K_{10}^2 + K_{11}^2}\right)^{-1}$, and the principal distance is

$$\begin{aligned} c_q &= \left[J^2(L_1^2 + L_2^2 + L_3^2) - q_0^2 \right]^{1/2} \\ c_{\tilde{q}} &= \left[J^2(L_5^2 + L_6^2 + L_7^2) - \tilde{q}_0^2 \right]^{1/2} \end{aligned} \quad (24)$$

The elements of the rotation matrix R are

$$\begin{aligned} r_{11} &= \frac{J(q_0L_9 - L_1)}{c_q} & r_{12} &= \frac{B(\tilde{q}_0L_9 - L_5)}{c_{\tilde{q}}} & r_{13} &= JL_9 \\ r_{21} &= \frac{J(q_0L_{10} - L_2)}{c_q} & r_{22} &= \frac{B(\tilde{q}_0L_{10} - L_6)}{c_{\tilde{q}}} & r_{23} &= JL_{10} \\ r_{31} &= \frac{J(q_0L_{11} - L_3)}{c_q} & r_{32} &= \frac{B(\tilde{q}_0L_{11} - L_7)}{c_{\tilde{q}}} & r_{33} &= JL_{11} \end{aligned} \quad (25)$$

The position of the camera center will be given by

$$\begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} = - \begin{bmatrix} L_1 & L_2 & L_3 \\ L_5 & L_6 & L_7 \\ L_9 & L_{10} & L_{11} \end{bmatrix}^{-1} \begin{bmatrix} L_4 \\ L_8 \\ 1 \end{bmatrix} \quad (26)$$

From (23) -(26), the camera matrix C is obtained.

$$C = KP \quad (27)$$

where K , called four-parameter, is given by:

$$K = \begin{bmatrix} f_x & p_x \\ & f_y & p_y \\ & & 1 \end{bmatrix} \quad (28)$$

and P , called camera matrix, is given by:

$$P = (R | t) \quad (29)$$

where t is the translation matrix.

There are many ways to solve the camera matrix P . Least square method, SVD or pseudo-inverse method can be used in the case of an over-determined system. For example, Hartley (1997) used the specific form of Kruppa's Equation and explicitly in terms of singular value decomposition (SVD) of fundamental matrix for calculating the focal length of camera; Zhang (1999) proposed the camera calibration procedure by using specifically model plane; and Heikkilä (2000) proposed the 4-step camera calibration procedure to solve the projective relation between the model plane and the image plane. Here the calibration of the webcam is base on Zhang's procedure for solving camera matrix P .

The Matlab® based toolbox developed by Bouguet (2008) is used for camera calibration. The image size is 640×480 pixels and the intrinsic parameters of the webcams and camera matrix are shown in Table 1.

Focal Length	Left camera		Right camera	
	f_x	f_y	f_x	f_y
	688.92	690.84	690.62	690.30
Principal point	293.70	237.91	289.32	242.02

Table 1. Intrinsic parameters of the webcams.

5. Experimental results

In our applications, the stereo system will provide the real-time height information for AUHs. It is necessary that the method should be simple and fast. The local search BMA and epipolar geometry constraints are utilized for searching the stereo corresponding points. There are three parts in this section. In the first part, corresponding points with fundamental matrix are searched and the BMA sreaching are verified. Next, the simulations of height estimation for AUHs are illustrated. The third part of this section demonstrates comparison of our methods with some other methods.

5.1 Measurement results of epipolar geometry constraints and BMA

From Table 1, the corresponding points can be obtained with BMA and epipolar geometry constraint from the obtained parameters of the webcams. As mentioned above, the fundamental matrix can be solved by the left and right camera matrices. An example shows the two images under the epipolar geometry constraint, and the results are shown in Figure 13. Right images are the reference images, and the corresponding points are lying on the epipolar lines of the left images.

Figure 14 demonstrates the corresponding points searching results with grid board. Almost all matching points between two images in specific area have been found out. The estimation results of different distances are shown in Figure 15. In the simulations, the measurement range is from 50 cm to 600 cm. Figure 15 shows that the error of estimation is less than 10 cm when the distance is 225 cm and baseline is 10 cm. When the distance is 300 cm with baseline being 15 cm, the estimation error is less than 10 cm. When the distance is 425 cm with baseline being 25 cm, the error of estimation is less than 10 cm. When the distance is 475 cm with baseline being 25 cm, the estimation error is less than 10cm. So we can conclude that the wider the baseline is, the further the measurement distance is.

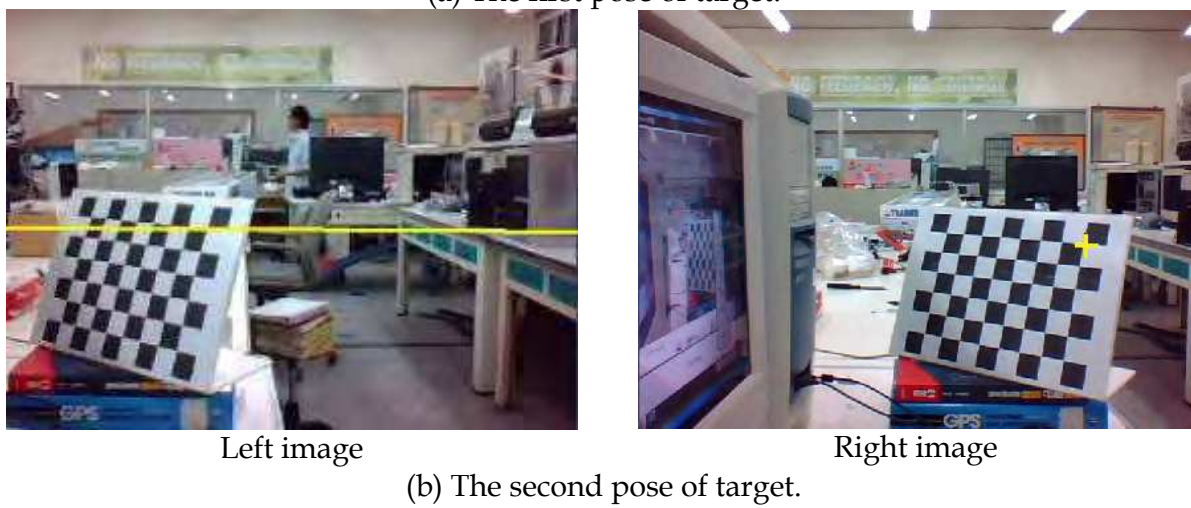


Fig. 13. Epipolar geometry constraint.

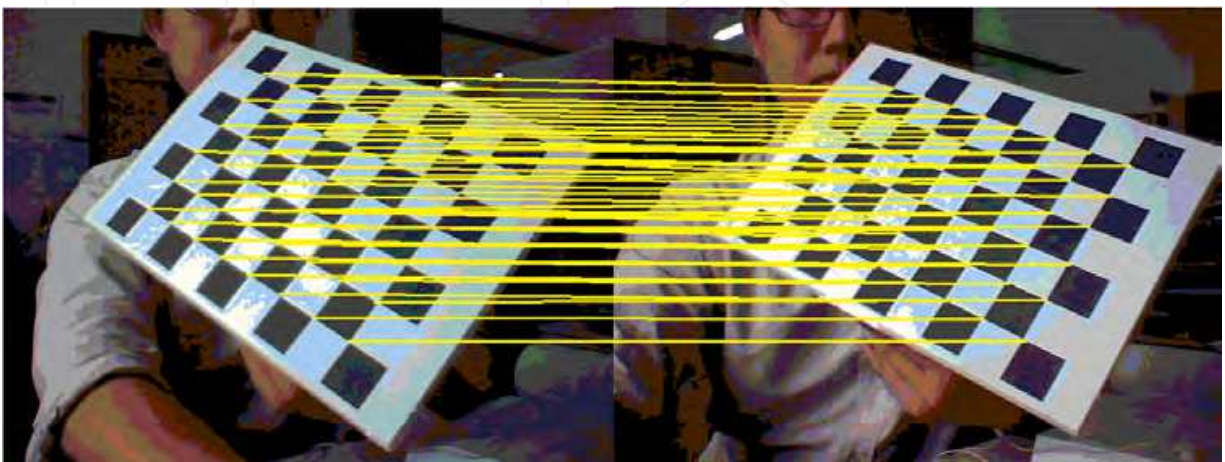


Fig. 14. Searching results of corresponding points with epipolar geometry constraint.

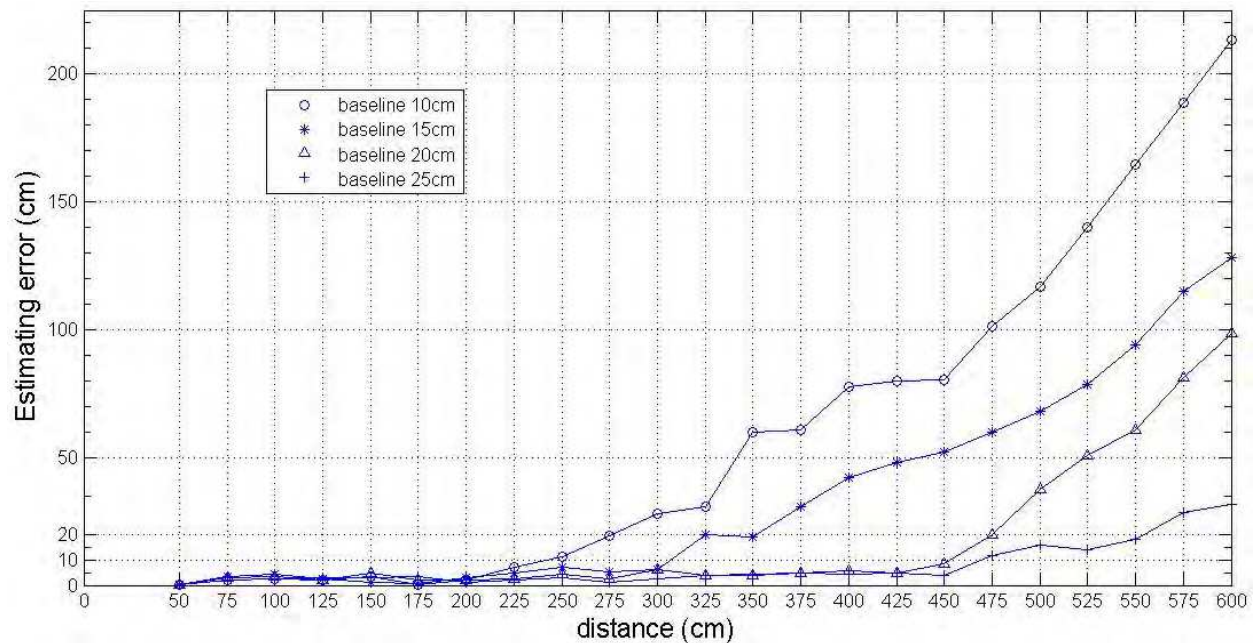
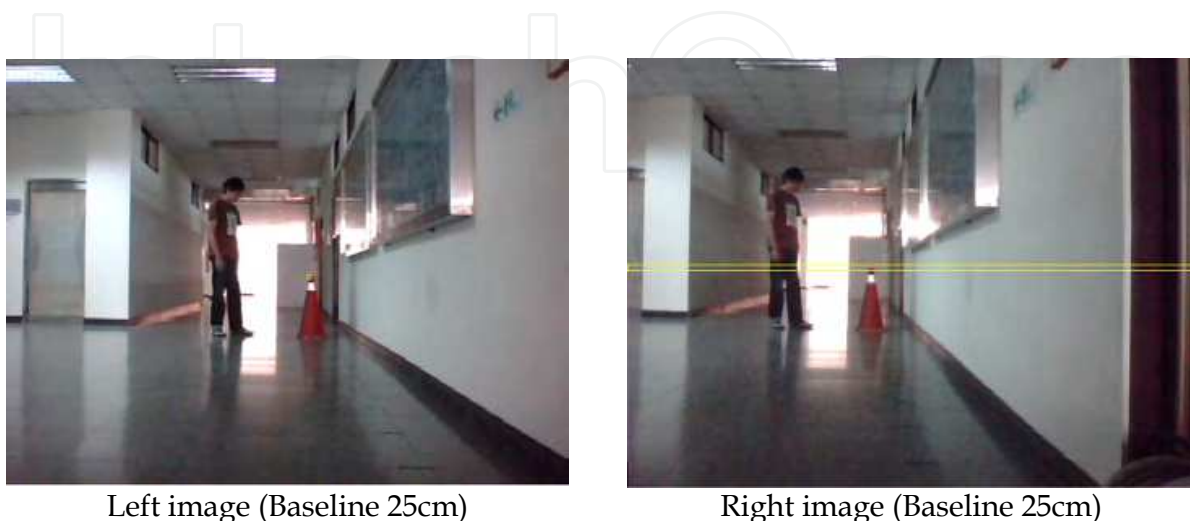


Fig. 15. Estimation errors with epipolar geometry constraint method.

For the BMA, Figure 16 shows that the size of the template block on left image is 5×5 pixels, the searching range is 640×10 pixels and the distance of target is 700 cm. The estimation results of different distances by BMA are shown in Figure 17.

From Figures 15 and 17, we can conclude that the measurement distance increase with baseline increasing. We can also conclude that the range of measurement distance of BMA is more than that of epipolar geometry constraint method, in the sense of the same error tolerance. However, as the measuring range increasing, BMA searching results almost in the same range because of the low resolution of the image, and this causes that the measurement error increases quickly.



Left image (Baseline 25cm)

Right image (Baseline 25cm)

Fig. 16. Searching corresponding points with BMA.

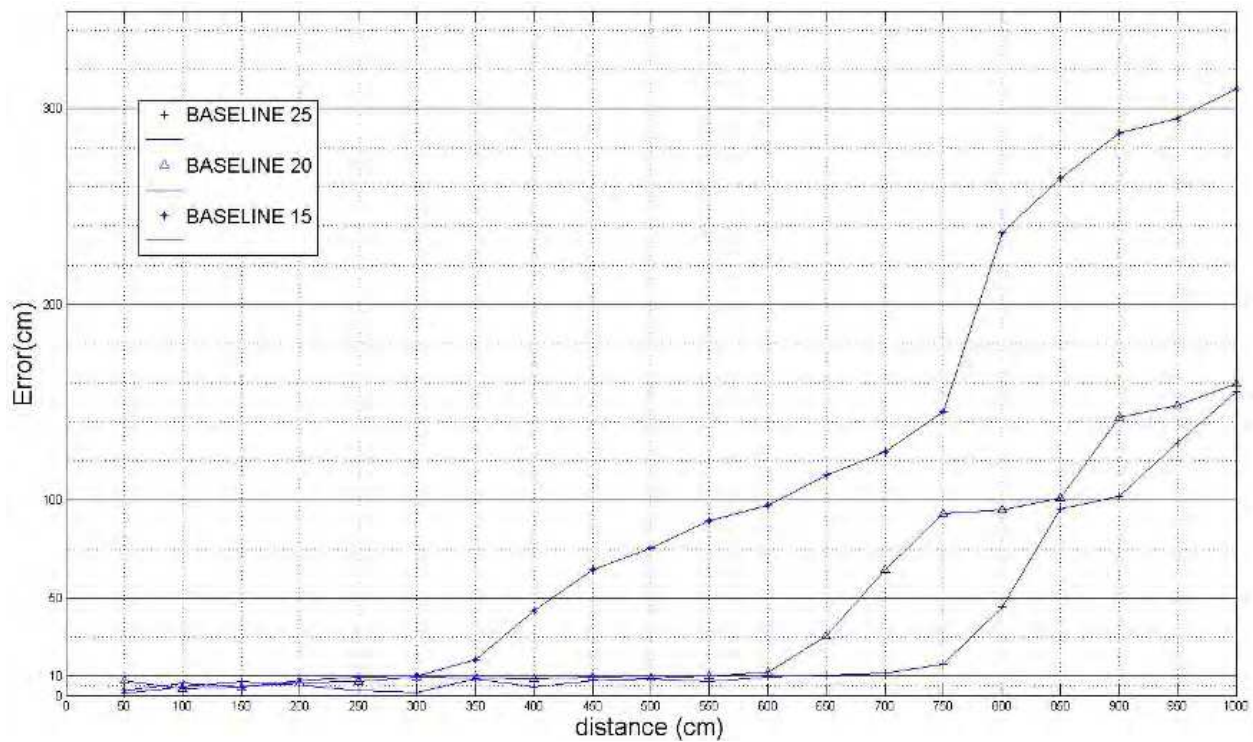


Fig. 17. Estimation errors with BMA.

5.2 The simulations of height estimation for AUHs

In this section, some stereo pairs of aerial photographs are captured for demonstrating our methods. Figures 18 through 21 illustrate the stereo pairs captured at different heights including 10.1m, 8.2m, 5.3m and 3.7m with baselines 10cm, 15cm, 20cm and 25cm. The Figure (c) of each figure group is an example for demonstrating the height estimation. Figure 22 illustrates the results of image processing. Figure 22(b) is the edge image of the landmark and Figure 22(c) is the disparity map.

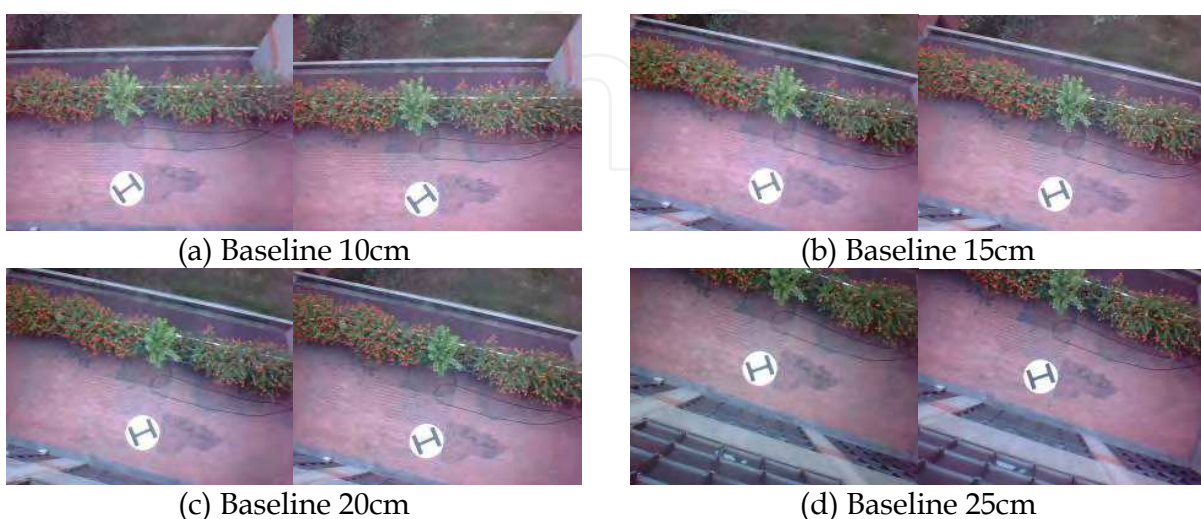


Fig. 18. Stereo pictures captured at height 10.1m.

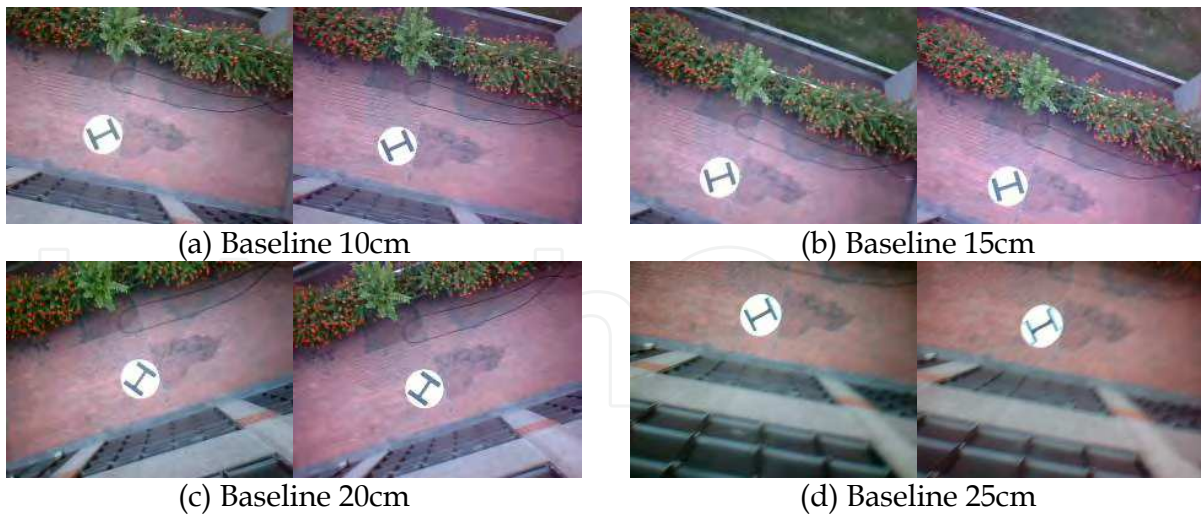


Fig. 19. Stereo pictures captured at height 8.2m.

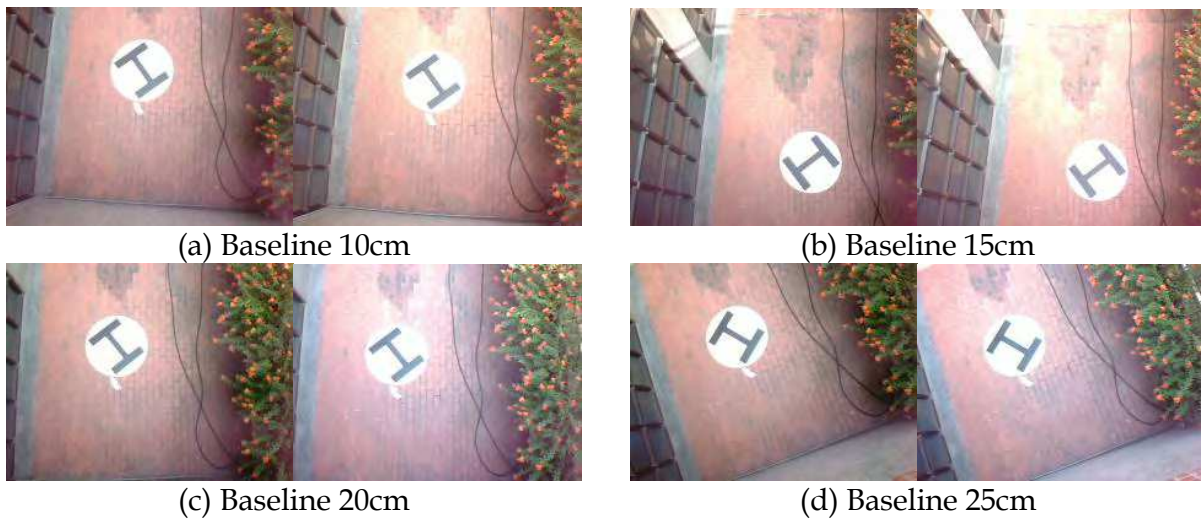


Fig. 20. Stereo pictures captured at height 5.3m.

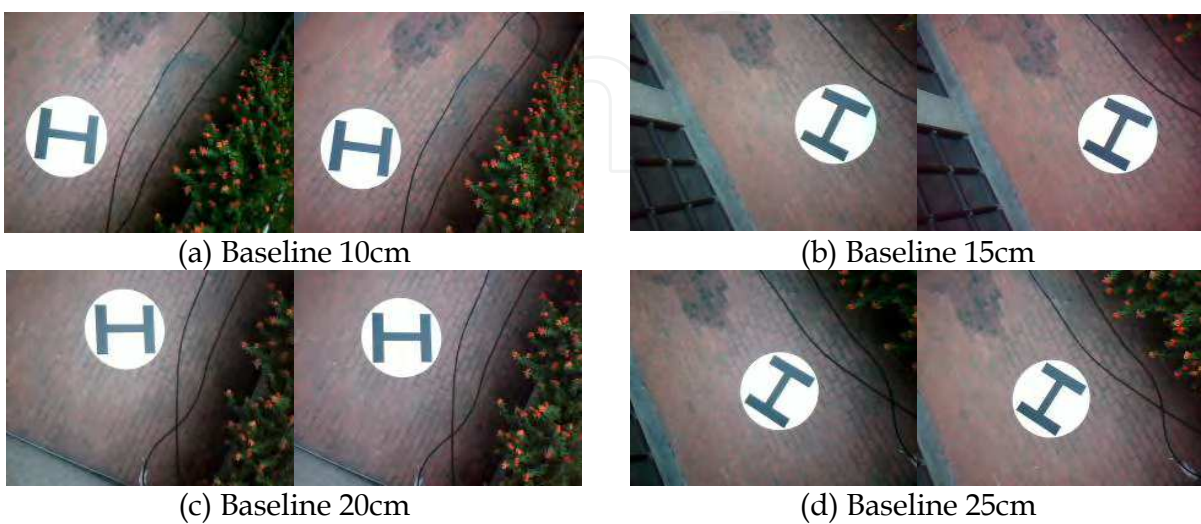


Fig. 21. Stereo pictures captured at height 3.7m.

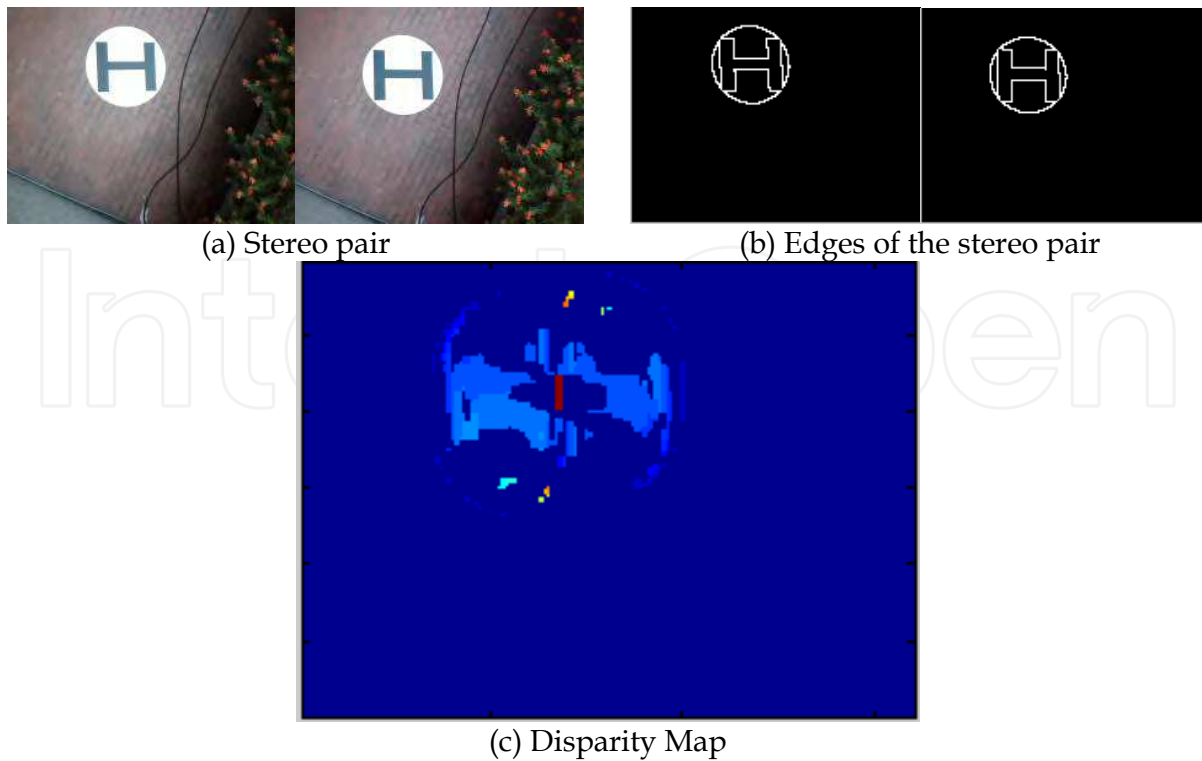


Fig. 22. Image processing and disparity computing results of Figure 21(b).

The estimating results of the simulations are illustrated in Figure 23. In Figure 23, the x-axis are the length of baselines, the y-axis are the estimation errors, and all the quantities in this figure are in meters. We can find from the figure that the estimation error is decreasing as the baseline increasing. And as the height growing with the baseline unchanged, the estimation error is increasing. When the height is 10.1m and the baseline is 10cm, both the errors of BMA and epipolar geometry constraint are over 2.5m. We can also conclude that the estimation errors by BMA will be less than those by epipolar geometry at the same condition.

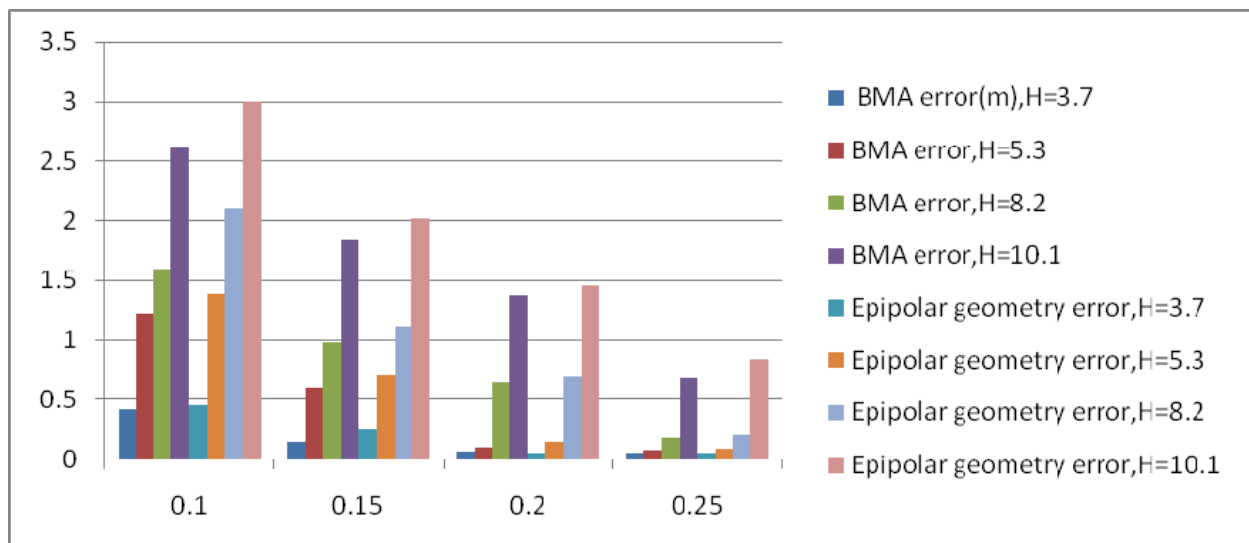


Fig. 23. Estimation results of landing simulations via BMA and epipolar geometry constraint.

5.3 Comparison of our methods with other methods

In Humenberger's work (Humenberger et al., 2010), a comparison of processing speed of some real-time stereo vision systems has been made. The proposed methods in this chapter are computed on the platform of CPU. The methods of Point Grey Research Inc. (Point Grey Research Inc., 2000), and Bradski (Bradski, 2010) are also computed on the platform of CPU. Table 2 shows the processing speed of these two systems comparing to our methods with Fig. 22(a) as the test image and SAD as the matching cost function. The size of the test image is 640×480. The processing speed is given in frames per second (fps).

Reference	Frames per second (fps)
Point Grey Research Inc.	34.5
Bradski	21.1
Our method 1 (Epipolar Geometry Constraint)	24.6
Our method 2 (BMA)	20.1

Table 2. Computation time of different matching methods.

6. Conclusion

For general purpose of helicopter autonomous flight, GPS is very useful. However, position information provided by GPS is not accurate enough for autonomous landing of an helicopter. The stereo vision system is designed to assist GPS while helicopter is autonomous landing. For small unmanned helicopters, the effective measuring range of 8 m is enough for landing control. The stereo vision system is a very competent sensor for height estimation. On the helicopter autonomous landing problem, stereo vision system could estimate the height of the helicopter. In this chapter, we proposed a low cost stereo vision system which is much cheaper than a GPS. The proposed system can provide acceptably accurate height information for the unmanned helicopter landing control system in certain range.

From the simulation results, it is evident that different baselines will produce different measurement results. The wider the baseline is, the longer that the system can be used for height estimation with acceptable range of error. Comparing the height estimation error of GPS, we can conclude that the system indeed provides more accurate information of height, and it is more useful for the helicopter autonomous landing. To increase the measurement range, one should use cameras of higher resolution and/or increase the baseline.

There are three major works need to be concerned in the future study. Firstly, maybe it is necessary to increase the number of cameras for expanding the camera range. In recent years, multi-view 3D construction technology had made significant progress. The 3D topographical construction with multi-view images that appears to be feasible. The second, the BMA should be further improved. A lot of search methods have been proposed for speeding-up or ameliorating the matching performance. Therefore our approach will be more improved. Finally, the other matching method (e.g. region matching) will be attempted for better matching performance.

7. References

- Alagoz, B.B. (2008). Obtaining depth maps from color images by region based stereo matching algorithms, *Computer Vision and Pattern Recognition*, Vol. 8, No. 4. pp. 1-13.
- Bagen, W., Hu, J. & Xu, Y. (2009). A vision-based unmanned helicopter ship board landing system, *International Congress on Image and Signal Processing*, pp. 1-5.
- Bedekar, A.S. & Haralick, R.M. (1995). A Bayesian method for triangulation and its application to finding corresponding points, *International Conference on Image Processing*, Vol. 2, pp. 362 - 365.
- Bradski, G. & Kaehler, A. (2008). *Learning OpenCV, Computer Vision with the OpenCV Library*, O'Reilly, 555 pages.
- Bouguet, J.Y. (2008). *Camera calibration toolbox for Matlab*, available from: http://www.vision.caltech.edu/bouguetj/calib_doc/
- Canny, J.F. (1986). A computational approach to edge detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 8, No. 6, pp. 679 - 698.
- Cyganek, B. (2005) Machine efficient adaptive image matching based on the nonparametric transformations. In: *Computational Science - ICCS 2005, 5th International Conference, Atlanta, GA, USA, May 22-25, 2005, Proceedings, Part I. Volume 3514 of Lecture Notes in Computer Science*, Sunderam, V.S., van Albada, G.D., Sloot, P.M.A., & Dongarra, J. eds., pp. 757 - 765, Springer, ISBN 978-3-540-26032-3, New York.
- Cyganek, B. (2006) Matching of the multi-channel images with improved nonparametric transformations and weighted binary distance measures. In: *Combinatorial Image Analysis - Proc. of the 11th International Workshop on Combinatorial Image Analysis (IWCIA 2006), volume 4040 of Lecture Notes in Computer Science*, Reulke, R., Eckardt, U., Flach, B., Knauer, U., & Polthier, K. eds., pages 74-88. Springer, ISBN 978-3-540-35153-5, New York.
- Cyganek, B. & Siebert, J.P. (2009). *An Introduction to 3D Computer Vision Techniques and Algorithms*, Wiley, 483 Pages.
- David, A. F. & Ponce, J. (2002). *Computer Vision: A Modern Approach*, Prentice Hall, 693 Pages.
- Doehler, H.-U. & Korn, B. (2003). Robust position estimation using images from an uncalibrated camera, *Digital Avionics Systems Conference*, Vol. 2, pp. 9.D.2-1-9.D.2-7.
- Duan D., Xie M., Mo, Q., Han, Z. & Wan, Y. (2010). An improved Hough transform for line detection, *International Conference on Computer Application and System Modeling*, Vol. 2, pp. V2-354 - V2-357.
- Fang, Z., Wu, J. & Li, P. (2008). Control system design and flight testing for a miniature unmanned helicopter. *Proceedings of the 7th World Congress on Intelligent Control and Automation*, pp. 2315-2319.
- Fuh, C.S. & Maragos, P. (1989). Region-based optical flow estimation, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp.130 - 133.
- Glover, F. (1989). Tabu search - Part I, *ORSA Journal on Computing*, Vol. 1, No. 3, pp. 190 - 206.
- Glover, F. (1990). Tabu search - Part II, *ORSA Journal on Computing*, Vol. 2, No. 1, pp. 4 - 32.
- Gyaourova, A., Kamath C. & Cheung S. (2003). Block matching for object tracking, *Technical Report on Lawrence Livermore National Laboratory*, Report No. UCRL-TR-200271, 13 Pages.
- Han, J.H. & Park, J.S. (2000). Contour matching using epipolar geometry, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.22, pp. 358 - 370.

- Hu W.C. (2008). Adaptive template block-based block matching for object tracking, *Eighth International Conference on Intelligent Systems Design and Applications*, Vol. 1, pp. 61 - 64.
- Hartley, R.I. & Zisserman, A. (2003). *Multiple View Geometry in computer vision*, Cambridge, 634 Pages.
- Hartley, R.I. (1995). In defence of the 8-point algorithm, *5th International Conference on Computer Vision*, Cambridge, pp. 1064-1070.
- Hartley, R.I. (1997). Kruppa's equations derived from the fundamental matrix, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 2, pp. 133 - 135.
- Humenberger, M., Zinner, C., Weber, M., Kubinger, W. & Vincze, M. (2010), A fast stereo matching algorithm suitable for embedded real-time systems, *Computer Vision and Image Understanding*, Vol. 114, No.11, pp. 1180 - 1202.
- Heikkilä, J. (2000). Geometric Camera Calibration Using Circular Control Points, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 10, pp. 1066 -1077.
- Johnson, A., Montgomery, J. & Matthies, L. (2005). Vision guided landing of an autonomous helicopter in hazardous terrain, *IEEE International Conference on Robotics and Automation*, pp. 3966 - 3971.
- Kadmiry, B. & Driankov, D. (2004). A fuzzy gain-scheduler for the attitude control of an unmanned helicopter, *IEEE Transactions on fuzzy systems*, Vol. 12, No. 4, pp. 502 - 515.
- Katzourakis, D., Vitzilaios, N.I. & Tsourveloudis, N.C. (2009). Vision aided navigation for unmanned helicopters, *17th Mediterranean Conference on Control & Automation*, pp. 1245-1250.
- Liang, T. & Kuo, P. (2008). A novel fast block-matching algorithm for motion estimation using adaptively asymmetric patterns. *International Journal of Innovative Computing, Information and Control (IJICIC)*, Vol. 4, No. 8, pp. 2011 - 2024.
- Lin, F., Chen, B. M. & Lum, K.Y. (2007). Integration and implementation of a low-cost and vision-based UAV tracking system, *26th Chinese Control Conference*, pp. 731 - 736.
- Luhmann, T., Robson, S., Kyle, S. & Harley, I. (2007). *Close Range Photogrammetry: Principles, Techniques and Applications*, Wiley, 528 pages.
- Mori, R., Hirata, K. & Kinoshita, T. (2007). Vision-based guidance control of a small-scale unmanned helicopter, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2648-2653.
- Muquit, M.A., Shibahara, T. & Aoki, T. (2006). A high-accuracy passive 3D measurement system using phase-based image matching, *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, Vol. E39-A, No. 3, pp. 686-697.
- Nixon, M.S. & Aguado A.S. (2008). *Feature Extraction & Image Processing*, Elsevier, 424 pages.
- Oh, S.R., Pathak, K., Agrawal, S.K., Pota, H.R. & Garratt, M. (2006). Approaches for a tether-guided landing of an autonomous helicopter, *IEEE Transactions on Robotics*, Vol. 22, No. 3, pp. 536 - 544 .
- Olson, C.F. (1997). Maximum-likelihood image matching, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 6, pp. 853 - 857.
- Saito, S., Bao Y., & Mochizuki T. (2007). Autonomous flight control for RC helicopter using camera image, *SICE Annual Conference*, pp. 1536 - 1539.
- Siebert, J.P. & Marshall, S.J. (2000). Human body 3D imaging by speckle texture projection photogrammetry, *Sensor Review*, Vol. 20, No. 3, pp. 218 - 226.
- Takahashi, Y., Karungaru, S., Fukumi, M. & Akamatsu, N. (2006). Feature point extraction in face image by neural network, *SICE-ICASE International Joint Conference*, pp. 3783 - 3786.

- Tao, T., Koo, J.C. & Choi, H.R. (2008). A fast block matching algorithm for stereo correspondence, *IEEE Conference on Cybernetics and Intelligent Systems*, pp. 38 - 41.
- Tico, M., Rusu, C. & Kuosmanen, P. (1999). A geometric invariant representation for the identification of corresponding points, *International Conference on Image Processing*, Vol. 2, pp. 462 - 466.
- Wang, C., Lei X., Liang, J., Wu, Y. & Wang, T. (2009). An adaptive system identification method for a micro unmanned helicopter robot. *IEEE International Conference on Robotics and Biomimetics*, pp.1093 - 1098.
- Wang, C.C., Lien, S.F., Hsia, K.H. & Su, J.P. (2009). Image-guided searching for a landmark, *Artificial Life and Robotics*, Vol. 14, No. 1, pp. 95 - 100.
- Wang, L. & Yang R. (2011). Global stereo matching leveraged by sparse ground control points, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3033 - 3040.
- Xu, C., Qiu, L., Liu, M., Kong, B. & Ge, Y. (2006). Stereo vision based relative pose and motion estimation for unmanned helicopter landing, *IEEE International Conference on Information Acquisition*, pp. 31 - 36.
- Zhang, Z. (1996). On the epipolar geometry between two images with lens distortion, *13th International Conference on Pattern Recognition*, Vol.1, pp. 407 - 411.
- Zhang, Z. (1999). Flexible camera calibration by viewing a plane from unknown orientation, *7th IEEE International Conference on Computer Vision*, pp. 666 - 673.
- GPS 18 Technical Specifications, Revision D (2005), Garmin International, 33 Pages.
- Triclops, *Technical Manual* (2000), Point Grey Research Inc., available from: <http://www.ptgrey.com/products/triclopsSDK/triclops.pdf>

IntechOpen



Machine Vision - Applications and Systems

Edited by Dr. Fabio Solari

ISBN 978-953-51-0373-8

Hard cover, 272 pages

Publisher InTech

Published online 23, March, 2012

Published in print edition March, 2012

Vision plays a fundamental role for living beings by allowing them to interact with the environment in an effective and efficient way. The ultimate goal of Machine Vision is to endow artificial systems with adequate capabilities to cope with not a priori predetermined situations. To this end, we have to take into account the computing constraints of the hosting architectures and the specifications of the tasks to be accomplished, to continuously adapt and optimize the visual processing techniques. Nevertheless, by exploiting the low-cost computational power of off-the-shelf computing devices, Machine Vision is not limited any more to industrial environments, where situations and tasks are simplified and very specific, but it is now pervasive to support system solutions of everyday life problems.

How to reference

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

Kuo-Hsien Hsia, Shao-Fan Lien and Juhng-Perng Su (2012). Stereo Matching Method and Height Estimation for Unmanned Helicopter, Machine Vision - Applications and Systems, Dr. Fabio Solari (Ed.), ISBN: 978-953-51-0373-8, InTech, Available from: <http://www.intechopen.com/books/machine-vision-applications-and-systems/stereo-matching-method-and-height-estimation-for-unmanned-helicopter>

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

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen