Research Paper Engineering

# Horizontal Line Based Stereo Matching Method

* Khyati N. Patel ** Mrs. Sameena Zafar

* P.G. Student, Patel College Of Science & Technology,Ratibad, Bhopal, Rajiv Gandhi Proudyogiki Vishwavidyalaya

** Head Of The Department, Patel College Of Science & Technology, Ratibad, Bhopal, Rajiv Gandhi Proudyogiki Vishwavidyalaya

ABSTRACT

*Obtaining reliable disparity maps, indicating distance of surface from the stereo camera pair, have importance in robotic applications and autonomous systems. Stereo vision is one of methods that can yield depth information of the scene. It uses stereo image pairs from two cameras to produce disparity maps that can be easily turn into depth maps. Reliability of depth maps and computational cost of algorithm is key issue for implementing real time robust applications.*
*Matlab R7b has been chosen for implementing different Stereo Matching Algorithms. Horizontally Line-Based gives good disparity Map. Stereo Matching Algorithms are tested on standard images like Tsukuba, Cravon, we can try for the same images and prepare the experimental real time setup for matching the images. Quality metrics use for evaluating the performance of stereo correspondence algorithms and the techniques used for acquiring our image data sets and ground truth estimates and got comparisons of all algorithm.*

Keywords :

**Introduction**
In recent years, it is an important research direction applying visual image technology for underwater target and environment detection. With the development of the theory of binocular vision, underwater binocular vision technology has been used in many fields, such as marine resources exploration, underwater target detection, marine ecological environment.

Binocular vision, which is inspired by humanvisual process, computes the disparity between correspondence points in images captured by two cameras for distance measurement, and then recovers the depth information of the object. Considering the features of underwater environment, applying binocular vision technology can better perceive underwater environment information. It provides the theory basis for robot understanding the underwater environment and realizing the navigation and positioning. The image matching is one of the key technologies to realize underwater binocular vision. And the result of the matching would affect directly the precision of object recognition and 3D scene reconstruction.

Image matching is a process of seeking the corresponding feature points in two different images which are in the same scene. Currently, there are many research results of under water image matching. stereo matching methods are divided into region-based matching, phase-based matching and feature-based matching. It can obtain a dense disparity map by region-based matching method. But it is rarely used because of its large amount of computation. The parallax image obtained by phase based matching method can reach sub-pixel accuracy. However, the method is more sensitive to the distortion, and it is difficult to choose the size of the matching window precisely. Feature-based matching method can greatly reduce the amount of the matching process calculation. Since it is more sensitive to the position changes, the precision of matching is higher. But there exists miss detection phenomenon. In the three matching methods, the more classical algorithms are Absolute Balance Search algorithm, normalized cross-correlation matching algorithms, image moments matching algorithm, the matching algorithms based on Harris corner points and so on.

Binocular vision is the process of recovering depth from two images with the same height, the same direction and a certain distance, similar to human vision principle. During the process, stereo matching is the key point, which means to find the correspondence pixels of the same physical spatial point on both images. Binocular stereo matching algorithm research falls into two categories. One is based on sparse points, and the other is based on dense points. The latter one is more accurate on image matching. There are many binocular dense-point matching methods in which the representatives are Birch Field algorithm and Yoon algorithm. Birch Field algorithm is used to match two gray-scale images. It applies dynamic programming into matching epipolar lines. It is efficient but inaccurate. Subsequent improvement method of Birch Field is based on each pixel's eight neighborhoods and improves the matching accuracy to some extent. In the study of color image matching, Yoon's method of adaptive window algorithm is a milestone.

Obtaining reliable depth maps, indicating distance of surface from the stereo camera pair, have importance in robotic applications and autonomous systems. Intelligent systems, which can move around by itself, could be developed by obtaining dept information from the sensors. Stereovision is the one of methods that can yield dept information of the scene. It uses stereo image pairs from two cameras to produce disparity maps that can be easily turn into depth maps.

**Brief Review**
Stereo matching is another technique that is well known for measure and it is very easy to understand and program. With stereo matching we can get the exact position of a target by two stereo pictures. However the traditional stereo matching needs a strict condition about camera, such as the axis of stereo camera (a pair of camera with the same properties) must be kept parallel and the height of cameras will be the same. These conditions limited the application of stereo matching. The position of a point P in 3D space can be measured by traditional stereo matching that computes the coordinate of P, (x, y, z), from two pictures called left image and right image taken by two cameras separately (Figure 1). Based on

the triangulation from two views the axis of the camera lenses must be paralleled to axis z, and the two cameras mustbe kept the same height, that is, the same y.

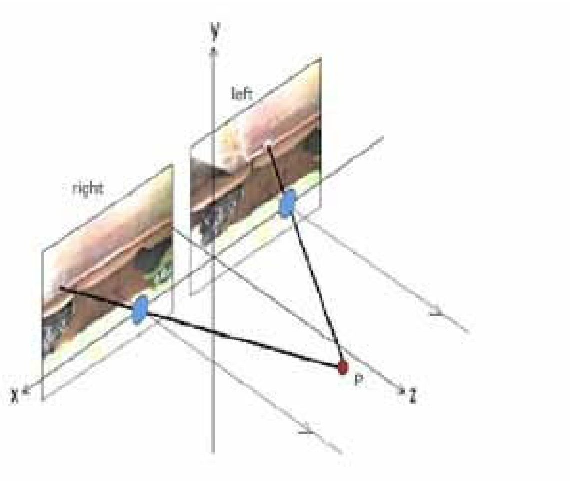



Figure:1 Left and Right images with projection point P

In Figure 1 and 2 the point P is our target we want to know P(x, y, z) exactly. For convenience we only show the x-z plane in Figure 2. Assume that the two cameras have the same focus distance f, the distance between the two cameras is d, pl(xl,yl) denotes the coordinates of P in the left image and pr(xr,yr) denotes the coordinate of P in right image, we can get P(x,y,z) by,

$$x = \frac{d}{2} . \frac{x_l + x_r}{x_l - x_r} \quad (1)$$

$$y = \frac{y_l . d}{x_l - x_r} \quad (2)$$

$$z = \frac{f . d}{m(x_l - x_r)} \quad (3)$$

$$x = \frac{d}{2} \cdot \frac{\frac{m_l . x_l}{f_l} + \frac{m_r . x_r}{f_r}}{\frac{m_l . x_l}{f_l} - \frac{m_r . x_r}{f_r}} \quad (4)$$

$$y = \frac{m_l}{f_l} \cdot \frac{y_l . d}{\frac{m_l . x_l}{f_l} - \frac{m_r . x_r}{f_r}} \quad (5)$$

$$z = \frac{d}{\frac{m_l . x_l}{f_l} - \frac{m_r . x_r}{f_r}} \quad (6)$$

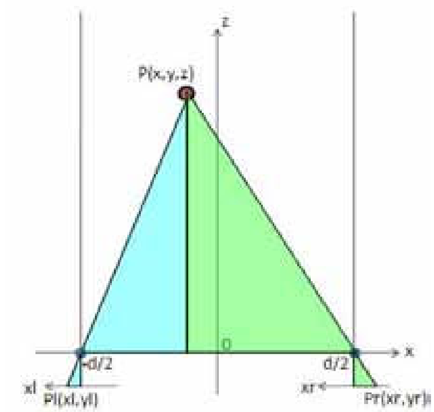
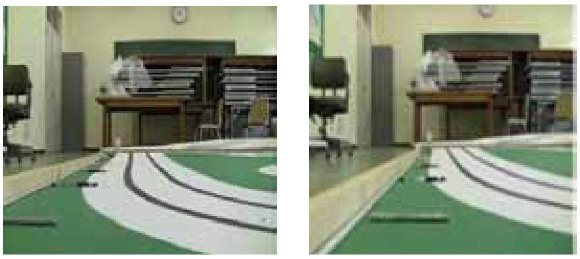



Figure:2 Coordinates in xy-plane with point

Figure: 3(a) and (b) show the left image and the right image taken by two different cameras.



(a) The image from left camera (b) The image from right camera

Figure: 3. Stereo images for computation (targets are the pens on the table)

**Proposed methods**

***Region Based Stereo Algorithms:***

***(a) Global Error Energy Minimization by Smoothing Functions***

In this method, we used block-matching technique in order to construct an Error Energy matrix for every disparity. Lets denote left image in RGB format by L(i, j,c) , denote right image in RGB format by R(i, j, c) and error energy by e(i, j, d) . For n*m window size of block matching, error energy e(i, j, d) can be expressed by,

$$e(i,j,d) = \frac{1}{3*n*m} \sum_{x=i}^{i+n} \sum_{y=j}^{j+m} \sum_{k=1}^{3} L(x, y+d, k) - R(x, y, k)))^2 \quad (7)$$

where, c represents RGB components of images and takes value of (1,2,3) corresponding to red, blue and green. d is the disparity. For a predetermined disparity search range (w), every e(i, j, d) matrix respect to disparity is smoothed by applying averaging filter many times. Averaging filter (linear filter) removes very sharp change in energy which possibly belong to incorrect matching. An other important properties of repeating application of averaging filter is that it makes apparent global trends in energy. (Local filtering in iterations could solve a global total variational optimization problem) Considering global trend in error energy naturally makes this algorithm a region based algorithm. For n*m window size, averaging filtering of e(i, j, d) can be expressed by following equation,

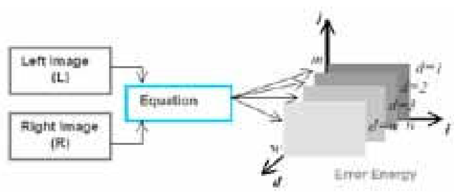
$$\tilde{e}(i,j,d) = \frac{1}{n*m} \sum_{x=i}^{i+n} \sum_{y=j}^{j+m} e(x,y,d) \quad (8)$$

After iterative application of averaging filtering to error energy for each disparity, we selected the disparity ( d ), which has minimum error energy e~(i, j,d) as the most reliable disparity estimation for pixel (i, j) of disparity map. Lets write basic steps of algorithm more properly,

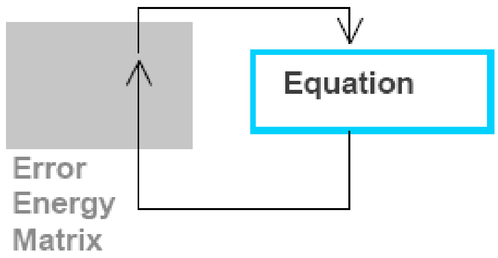
Step 1: For every disparity d in disparity search range, calculate error energy matrix.

Step 2: Apply average filtering iteratively to every error matrix calculated for a disparity value in the range of disparity search range.

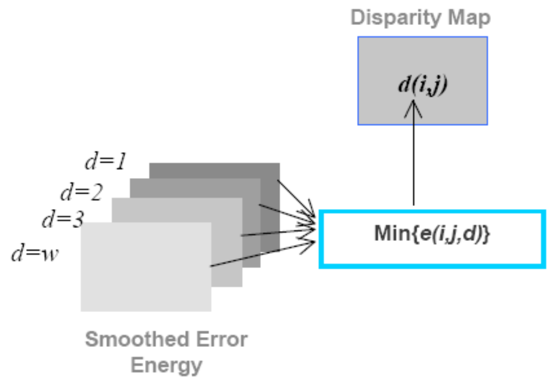
Step 3: For every (i, j) pixel, find the minimum error energy e~(i, j,d) , assign its disparity index ( d ) to d(i, j) which is called disparity map.



(a) Construction error enery matrix




(b) Smoothing energy matrix for every disparity values




(c) Disparity map generation by minimum every points

Figure :4 . Method using global error energy minimization by smoothing functions

(b) Line Growing Based Stereo Matching
Also proposed an algorithm based on region growing. In this manner, we consider region-growing mechanism in two phases operation. First phase, finding root point to grow region (Root Selection process) and the second phase, growing region for a root point corresponding to predefined rule.(Region Growing process) Our rule for associating a point to root point in the growing process is to have lower error energy than a predetermined threshold of error energy (Line Growing Threshold). In our application, being associated to a root points means to have the same disparity by root point. Thereby, the region emerged from all associated points has a disparity value. Actually, we should call the algorithm as disparity growing. Lets generally express steps of the algorithm in a list,

Step 1: (Root Selection process) Select a point, which isn't belonging to any grown region and find its disparity using energy function equation (1). Set it root point and set its disparity to region disparity then go to step 2. If you didn't find any disparity with lower enough error energy, repeat this step for the next point.

Step 2: (Region Growing process) Calculates error energy of neighbor points just for root point disparity, which was called region disparity. If it is lower than the predetermined error energy threshold, associate this point to region. Otherwise, left it free.

Step 3: Proceed the Step 2 until region growing any more. In the case that region growing is completed, turn back to step 1 to find out new root point to repeat these steps. When all points in image processed, stop the algorithm. Grown disparity regions composes disparity map d(i, j) .

In order to reduce complexity of the algorithm, we allow the region growing in the direction of rows since disparity of stereo image is only in row directions. So, only one neighbor, which is the point after searched point is inspected for region growing.

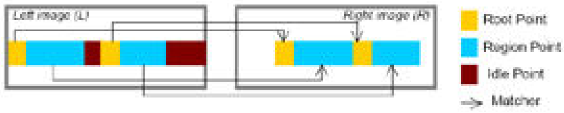
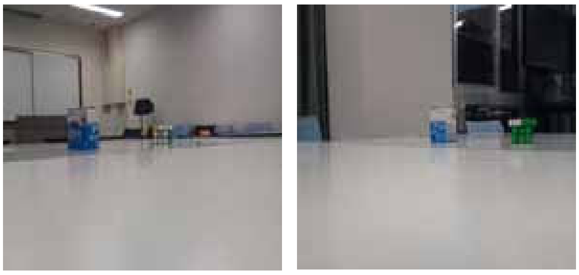



Figure: 5 Method using line growing

**Problem Statement**
In our work a number of cameras are located around of target and the two pictures taken by each pair are used for stereo matching calculation. It is impossible to compute the coordinate of target because the axes of camera lens are probably not paralleled each other. Assume these cameras having the same height we proposed a stereo axes correction algorithm and new equations to solve the unparalleled lens axes problem.



Picture from the left camera Picture from the right camera

.

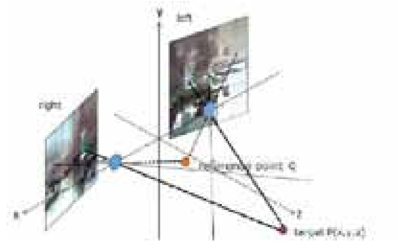


Figure :6 Stereo axes correction

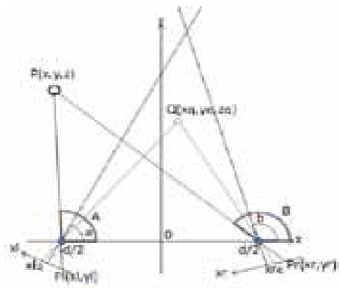


Figure:7 The principle of stereo axes correction

The Rubik's cube both in the left and right images of Figure 6 is our measured target. We do know the two axes of camera lens are not paralleled but we are not sure exactly how much they incline each other. In order to find out the lens axes inclined angle a and b of Figure 7 we pick up a known point Q(xq,yq,zq) as the reference point whose depth value zq and the distance between camera and Q can be measured by any measure tool. Figure 8 and (9) gives the method to measure the point Q.

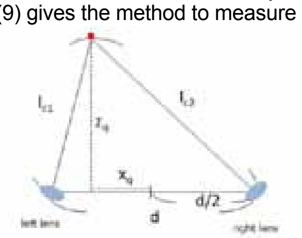


Figure:8 The Measure of the reference point

$$x_q = \frac{l^2.c_1 - l^2.c_2}{2.d} \tag{9}$$

$$z_q = \sqrt{l^2.c_1 -}\,(x_q + \frac{d}{2})^2 \tag{10}$$

Now the known Q($x_q$, $y_q$, $z_q$) into equations, where $x_{lq}$ is coordinate x of Q in the left image and $x_{rq}$ is coordinate x of Q in the right image,

$$\mathrm{a} = \tan^{-1}\left(\frac{z_q}{x_q + \frac{d}{2}}\right) + \tan^{-1}\left(\frac{x_{l_q}.\,m}{f}\right) \tag{11}$$

$$\mathrm{b} = \tan^{-1}\left(\frac{z_q}{x_q - \frac{d}{2}}\right) + \tan^{-1}\left(\frac{x_{r_q}.\,m}{f}\right) \tag{12}$$

A and B that are inclined angles of target P in the left image and the right image can be expressed by

$$\mathrm{A} = \mathrm{a} - \tan^{-1}\left(\frac{x_l.m}{f}\right) \tag{13}$$

$$\mathrm{B} = \mathrm{a} - \tan^{-1}\left(\frac{x_r.m}{f}\right) \tag{14}$$

Therefore the coordinate of target P can be obtained by

$$X = -\frac{d}{2} * \frac{\tan(A) + \tan(B)}{\tan(A) - \tan(B)} \tag{15}$$

$$y = \frac{y_l\,.m.\sqrt{\left(x + \frac{d}{2}\right)^2 + z^2}}{\sqrt{f^2 + (m.x_l)^2}} \tag{16}$$

$$\mathrm{z} = \tan(A).\left(x + \frac{d}{2}\right) \tag{17}$$

If the properties of the left camera and the right camera are different, let fl,ml and fr,mr stand for the focus distance and coefficients m of the left camera and the right camera separately, a, b, A and B are rewritten as follows,

$$\mathrm{a} = \tan^{-1}\left(\frac{z_q}{x_q + \frac{d}{2}}\right) + \tan^{-1}\left(\frac{x_{l_q}.m_l}{f_l}\right) \tag{18}$$

$$\mathrm{b} = \tan^{-1}\left(\frac{z_q}{x_q - \frac{d}{2}}\right) + \tan^{-1}\left(\frac{x_{r_q}.m_r}{f_r}\right) \tag{19}$$

$$A = a - \tan^{-1}\left(\frac{x_l\,.m_l}{f_l}\right) \tag{20}$$

$$B = b - \tan^{-1}\left(\frac{x_r\,.m_r}{f_r}\right) \tag{21}$$

Having the new equations not worry about whether the lens axes of cameras are paralleled or not anymore. And the shadow area of figure become more widely than before, if necessary we can incline leans axes intentionally to get better broad views. Meanwhile in the new method the reference point plays an important role to correct the unparalleled lens axes.

A critical issue in stereo matching is to measure the similarity (dissimilarity) between correspondences, which is calculated as a matching cost. Common matching costs defined based on the brightness constancy assumption, i.e., scene points have similar intensities in different views, are Absolute Difference and Squared Difference. Using the matching cost, many local and global stereo methods have been proposed to improve the matching accuracy, substantially.

Other matching cost functions obtain robustness to radiometric differences by removing or relaxing the brightness constancy assumption. Hirschmuller evaluated many of them such as Normalized Cross-Correlation (NCC) ,rank and census transforms , LoG and mean filters. As a more complicated measure, mutual information method can handle more complex radiometric transformations. When globally reasoning the image radiometric transformation, mutual information method is comparably sensitive to local variations such as vignetting.

Figure:9 Matching the Tsukuba (a) left image and (b) its right image with a global intensity bias

**Work**

The major innovative point is to combine color aggregation with local disparity estimation and adaptive window matching. It is able to accomplish a better matching accuracy while effectively reducing the time complexity thus improves the performance of the algorithm.

Compared with land image, there are more significant differences in underwater image quality. Medium's strong absorption of light and scattering properties, the underwater image has the characteristics of low contrast, high ambiguity and low image pixel resolution.

In underwater image based on stereo matching algorithm is already present and here we can see that which output will come if underwater image is based on window-based adaptive correspondence search algorithm but we can use horizontal line based method.

Using window and Horizontal line we use the following steps:

Take the disparity map estimated from the above step as the initial value, now we use window-based correspondence method to optimize the result.

For each pixel p and its neighborhood Np in reference image, the corresponding pixel pd and its neighborhood Np in target image, define the dissimilarity E (p, p4) between the two windows:

$$E(p, p_d) = \frac{\sum_{q \in N_p, q_d \in N_{p_d}} w(p,q) w(p_d, q_d) e(q, q_d)}{\sum_{q \in N, q_d \in N_{p_d}} w(p,q) w(p_d, q_d)} \tag{22}$$

Where e (q ,qd) is absolute difference and w( p, q) is the adaptiveweight:

$$\mathrm{e}(\mathrm{q}, q_d) = min\left\{\sum_{e \in (c,g,b)} |I_c(q) - I_c(q_c)|\,, T\right\} \tag{23}$$

$$w(p,q) = f\left(\Delta c_{pq}, \Delta g_{pq}\right)$$

$\Delta c_{pq}$ and $\Delta g_{pq}$ are color similarity and geometric proximity.

To better understand depth and disparity relation, let see stereo projection representation illustrated in the Figure 14. By considering the figure, one can derive relation between dept (Z) and disparity (d) by using basic geometrical calculations as following.

$$Z(i,j) = f.\frac{T}{d(i,j)} \tag{24}$$

If real location of object surface projected at pixel (i, j) is willing to calculate, following formulas can be used in calculation of (X ,Y) points after calculation of the Z .

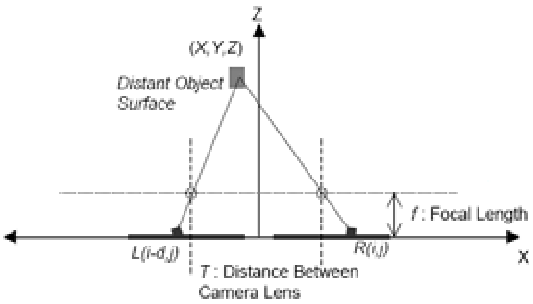



Figure:10 Representation of the stereo projection

In order to obtain smoother depth map to be used in applications such as robot navigation and recent trend for vision in various engineering application.

For that we take a various left and right angle with putting different distance of the camera/webcam and determine the disparity map for image, if time permits then try for real time interface with camera/webcam to computer system and prepare the stereo matching of the images and determine using smoothing with and without consider the reliability of non estimating pixel of images, also determine RMS error, time to be taken for execution in our system.

Our work progress chart show in figure 11.

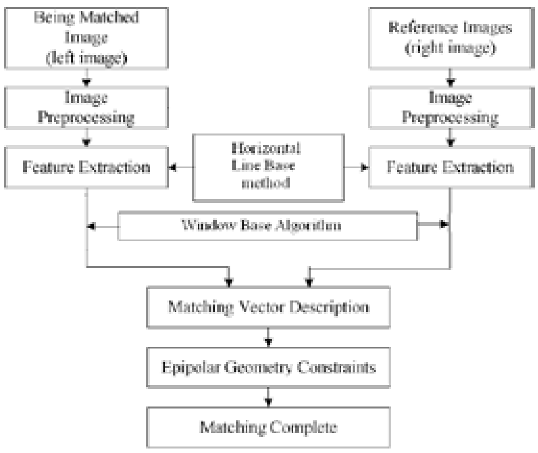
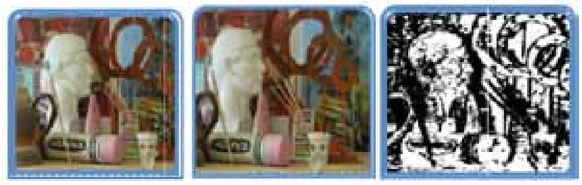



Figure:11 Flow chart of matching method of horizontal line and window based.

Left Camera Image Right Camera Image Disparity Map



Left Camera Image Right Camera Image Disparity Map

Figure:12 Horizontal Line Based Method image and Disparity Map

**Conclusion**

According to the stability demand of the underwater binocular image processing, this paper introduces an underwater image matching method based horizontal line base method. Because there is special corresponding relation between binocular images, we apply epipolar geometry constrain to eliminate the error matched points, enhancing the matching efficiency. In the study, we find that, in this special underwater environment, the preprocessing quality of the image and the scale sensitivity of the algorithm has a great influence on the image matching accuracy. So how to further improve the image preprocessing effect and reduce the algorithm scale sensitivity is our future research.

| Time (s) | Tsukuba | Venus | Teddy | Cones |
|---|---|---|---|---|
| Yoon's | 1800 | 2700 | 3600 | 3600 |
| Horizontal line based | 700 | 1000 | 2400 | 2500 |

| HL | high light and low noise | Output |
|---|---|---|
| 0.7 | 75 | 98.4% |

**REFERENCES**

[1] Huibin Wang, Hongye Sun, Jie Shen, Zhe Chen, "A Research on Stereo Matching Algorithm for Underwater Image", 2011 4th International Congress on Image and Signal Processing, 978-1-4244-9306-7/11/$26.00 ©2011 IEEE,pp.850-854. | [2] Shujun Zhang, Jianbo Zhang, Yun Liu, "A Window-Based Adaptive Correspondence Search Algorithm Using Mean Shift and Disparity Estimation", 2011 International Conference on Virtual Reality and Visualization, DOI 10.1109/ICVRV.2011.47, 978-0-7695-4602-5/11 $26.00 © 2011 IEEE,pp.319-322. | [3] Ningping Sun, Hiroshi Nagaoka, Takuya Shigemoto, "A Correction Algorithm for Stereo Matching with Web Cameras", 2011 International Conference on Broadband and Wireless Computing, Communication and Applications, DOI 10.1109/BWCCA.2011.88, 978-0-7695-4532-5/11 $26.00 © 2011 IEEE,pp.543-548. | [4] Jian Wei, Shigang Wang, Liwei Chen, Tianxiao Guan, "Adaptive Stereo Video Object Segmentation Based on Depth and Spatio-Temporal Information", 2009 World Congress on Computer Science and Information Engineering, DOI 10.1109/CSIE.2009.422,978-0-7695-3507-4/08 $25.00 © 2008 IEEE,pp. 140-144 | [5] M. Kaaniche, W. Miled, B. Pesquet-Popescu, A. Benazza-Benyahia and J.- C. Pesquet, "Dense Disparity Map Representations For Stereo Image Coding", ICIP 2009,978-1-4244-5654-3/09/$26.00 ©2009 IEEE,pp.725-728. | [6] Barıs Baykant ALAGÖZ, "Obtaining Depth Maps From Color Images By Region Based Stereo Matching Algorithms", OncuBilim Algorithm And Systems Labs. Vol.08, Art.No:04,(2008),pp.1-13. | [7] Andrea Fusiello, Emanuele Trucco, Alessandro Verri, "A compact algorithm for rectification of stereo pairs", Machine Vision and Applications (2000) Springer-Verlag 2000, pp.16–22. | [8] Ke Zhang, Jiangbo Lu,Gauthier Lafruit1, Rudy Lauwereins,Luc Van Gool, "Robust Stereo Matching With Fast Normalized Cross-Correlation Over Shape-Adaptive Regions", ICIP 2009, 978-1-4244-5654-3/09/$26.00 ©2009 IEEE,pp.2357-2360. | [9] N. Grammalidis and M.G. Strintzis, "Disparity and Occlusion Estimation in Multiocular Systems and their Coding for the Communication of Multiview Image Sequences," IEEE Trans. Circuits and Systems for Video Technology, vol. 8, no. 3, pp. 328–344, June 1998. | [10] Grammalidis and M.G. Strintzis, "Disparity and Occlusion Estimation in Multiocular Systems and their Coding for the Communication of Multiview Image Sequences," IEEE Trans. Circuits and Systems for Video Technology, vol. 8, no. 3, pp. 328–344, June 1998. | [11] Geoffrey McLachlan and Thriyambakam Krishnan. The EM Algorithm and Extensions. John Wiley & Sons, New York, 1996. | [12] L. Hong and G. Chen. Segment-based stereo matching using graph cuts. In CVPR, pages I: 74–81, 2004. | [13] M. Bleyer and M. Gelautz. A layered stereo matching algorithm using image segmentation and global visibility constraints. ISPRS Journal of Photogrammetry and Remote Sensing, 59(3):128–150, May 2005 | [14] J. Sun, N. Zheng, and H. Shum. Stereo matching using belief propagation. In ECCV, page II: 510 ff., 2002. | [15] Q. Y´ang, L. Wang, R. Yang, H. Stew´enius, and D. Nist´er. Stereo matching with color-weighted correlation, hierarchical belief propagation and occlusion handling. Accepted to CVPR 2006.