Depth Detection Based Obstacle Avoidance

CS 684 : Embedded Systems Application Project

Group No : 19

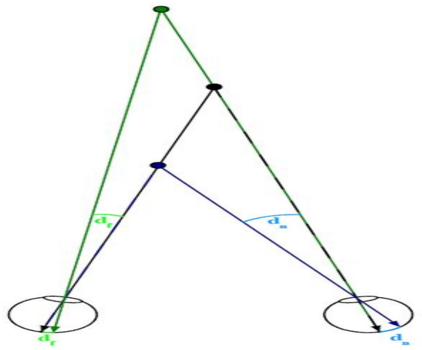
Hussain Manasawala 09307923 husainmanasa@ee.iitb.ac.in

Instructed by : Prof. Kavi Arya

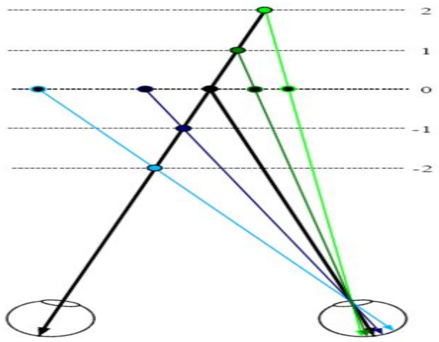
Submitted on – 15th November 2010

1. Introduction

Stereopsis is the process of visual perception leading to the sensation of depth from the two slightly different projections of the world onto the retinas of the two eyes. The difference in the two retinal images is called *Binocular Disparity* which refers to the difference in image location of an object seen by the left and right eyes, resulting from the eyes' horizontal separation. In computer vision, binocular disparity refers to the same difference seen by the two different cameras instead of eyes. The adjacent
stereo vision of eyes.



(a) Detection of Distance : Far and Near



(b) Detection of Depth in the Plane

The stereo vision of human eye motivates the computer vision to compute depth of an image through disparity estimation between two stereo images. It follows the same mechanism that human eye does.

Depth detection of an input image is done by the disparity estimation of stereo vision. Stereo vision refers to the ability to infer information about the 3D structure of a scene using two images captured from different viewpoints. The disparity map gives the relative depths between the stereo images.

1.1 Problem statement

The project involves the design and implementation of an engineering solution to the problem of autonomous geological obstacle detection and avoidance.

Given an Arena with objects placed in a random manner the Robot scans the area to create a Depth Map. Using this determines the position of the objects creating a position map and takes precise action to navigate around the obstacles.

The implementation has three main components :

(i) a stereo vision algorithm for depth detection,

(ii) construction of a position map of the area with the objects present and
(iii) Motion control for obstacle avoidance.

1.2 Requirements

Hardware:

i. One FB5 with ATmega2560 microcontroller.
ii. Two USB Cameras and their mounting mechanism.
iii. Two 2.4GHz XBee modules : One mounted on FB5 and other connected to PC through USB interface.

Software:

i. AVR Studio 4 - for coding n compiling C file to hex file and loading hex file on robot.
ii. MATLAB 7.9.0.529 (R2009b) – for image processing and controlling the robot through PC.

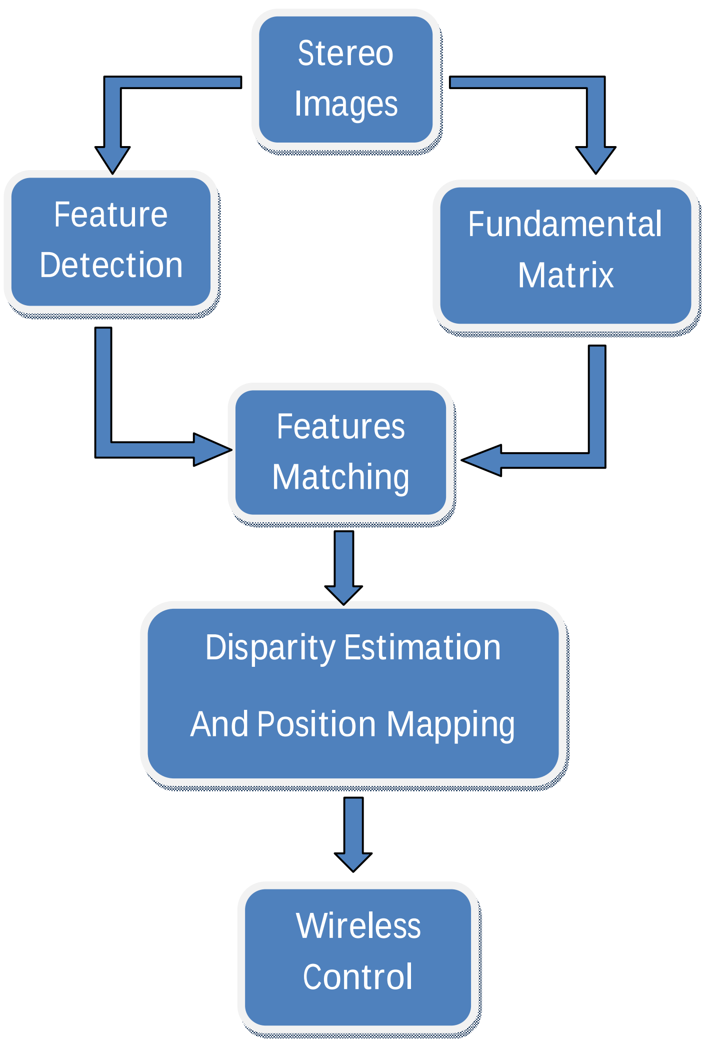
1.3 Assumptions and limitations

i. The Area is divided into virtual sections (grid) so as the robot moves only in steps.
ii. The surface of the ground and the objects are covered with non-reflective material.
iii. The robot is able to detect and navigate through objects of defined size and shapes but few in number.

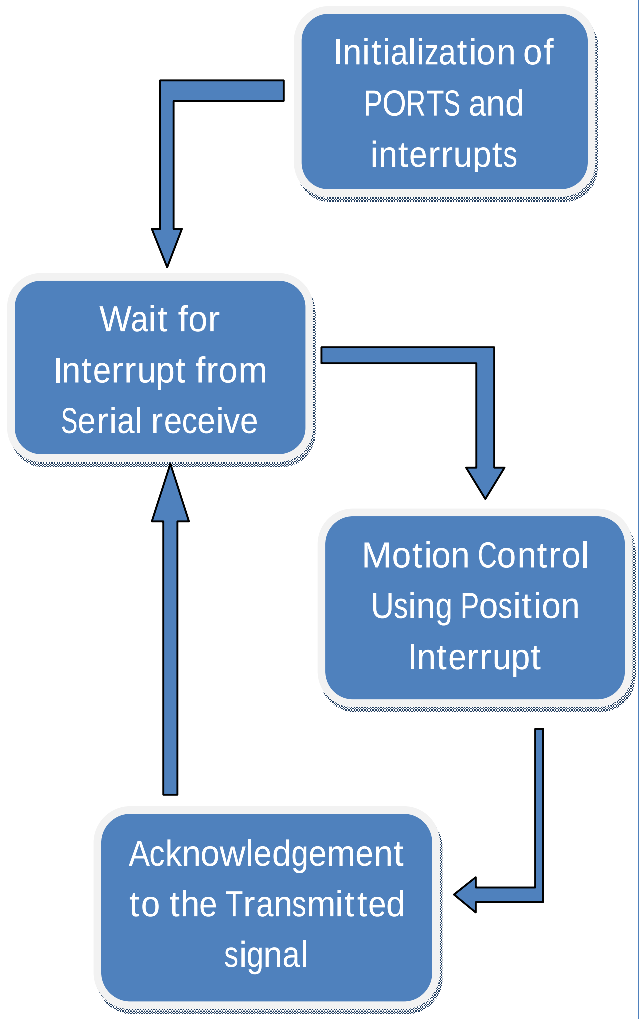
2. Overall Description

2.1 System Design

Block Diagram of the Embedded System in MATLAB:



Block Diagram of the Embedded System on FB5:



2.1.1 Embedded System in MATLAB:

- The FB5 robot has 2 cameras mounted on it with a specified distance between them. At a given point of time it captures a pair of ***stereo images*** of the scene in front of it. These images are transmitted to a computing station where image processing tool is installed.

- ***Feature detection*** involves the detection of the feature points which are the edges and corners in the stereo images. Harris' algorithm helps in finding the corners.

- ***Fundamental matrix*** describes the epipolar geometry of stereo vision system. It correlates the corresponding points of the stereo images using epipolar geometry.

- ***Feature matching*** is the procedure to match the corresponding points in the stereo images. This method calls for the computation of normalized cross correlation between the corresponding points in the stereo image pair.

- ***Disparity Estimation*** maps the disparity between similar points in the two images to a gray scale intensity. The pixel intensities describe the relative depth of points in the scene.

- Using this pair a depth/Position map is created through disparity estimation between the images. Using this depth map which is a gray image a 3D model of the space consisting of the objects present is created.

- From this the position and size information of the objects is transmitted back to the FB5 model wirelessly.

2.1.2 Embedded System on FB5 robot :

- The Ports and Interrupts are initialized to receive signal and move the robot in defined manner.

- The robot is waiting for the commands from the PC.

- On receiving the commands the robot decodes the information and also sends acknowledgement.

- Using this information the computer on the robot generates appropriate commands for the robot to navigate through the space avoiding the obstacles.

2. 2 Literature Review

2. 2. 1 Combined Edge and Corner Detection using Harris' Algorithm [1]

The mathematical basis of the edge and corner detection is explained here in brief.

Denoting the image intensities by I, the change E produced by a shift (x, y) is given by:

$$E_{x,y} = \sum_{u,v} w_{u,v}[I_{x+u,y+v} - I_{u,v}]^2 = \sum_{u,v} w_{u,v}[xX + yY + O(x^2, y^2)]^2$$

Where the first gradients are approximated by:

$$X = I \otimes (-1,0,1) = {\partial I}/{\partial x}$$
$$Y = I \otimes (-1,0,1)^T = {\partial I}/{\partial y}$$

Hence for small shifts E can be written as:

$$E(x, y) = Ax^2 + 2Cxy + By^2$$

where $A = X^2 \otimes w;\quad B = Y^2 \otimes w;\quad C = (XY) \otimes w$

The response is noisy because the window is binary and rectangular. Hence a smooth circular window, for example a Gaussian is used

$$w_{\boldsymbol{u},\boldsymbol{v}} = exp\{-{(u^2 + v^2)}/{2\sigma^2}\}$$

The change E, for the small shift (x, y) can be concisely written as:

$$E(x, y) = (x, y)M(x, y)^T$$

where the 2×2 symmetric matrix M is given by:

$$M = \begin{bmatrix} A & C \\ C & B \end{bmatrix}$$

(a) Left Image

(b) Right Image

so necessary to have the measure of corner and edge quality or response. The size of the response will be used to select isolated corner pixels and find the edge
It is also necessary to have the measure of corner and edge quality or response. The size of the response will be used to select isolated corner pixels and find the edge pixels.

Co nsider the trace and determinant of the matrix M as Tr and Det, then the corner response is given by:

$$R = Det - k(Tr)^2$$

R has high positive value in the corner region, high negative value in the edge region and small value in the flat region.

2. 2. 2 Fundamental Matrix

In computer vision, the fundamental matrix F is a 3×3 matrix which relates corresponding points in stereo images. In epipolar geometry, with the image coordinates U_r and U_l of corresponding points of stereo image pair, FU_r describes an epipolar line on which the corresponding point U_l on the other image must lie. That means for all pairs of corresponding points, the following holds:

$${U_l}^T F U_r = 0$$

2. 2. 3 Eight Point Algorithm

This is a linear method for extracting the Fundamental matrix from a list of matched pairs of points. The fundamental matrix has rank 2 and 7 degrees of freedom and hence can be recovered from only seven correspondences. However, the algorithm is much simpler if we have eight correspondences. For any given matched pair of points $U_l = [c_l\ r_l]$ and $U_r = [c_r\ r_r]$, the epipolar constraint as mentioned above is given by:

$$(c_l, r_l, 1)\begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix}\begin{pmatrix} c_r \\ r_r \\ 1 \end{pmatrix} = 0$$

Given 8 such correspondences , we can recover the matix F from the homogeneous system given below:

$$A^{8\times 9} f^{9\times 1} = 0$$

where $$f = (F_{11}\ F_{12}\ F_{13}\ \dots\ F_{33})^T$$

The least square solution for F is the eigenvector corresponding to the smallest eigenvalue of A, i.e. the last column of the matrix V in the Singular Value Decomposition (SVD), $A = UDV^T$. This gives the intermediate matrix F'. To obtain the correct rank 2 fundamental matrix, SVD of F' is done which is given by:

$$F' = USV^T$$

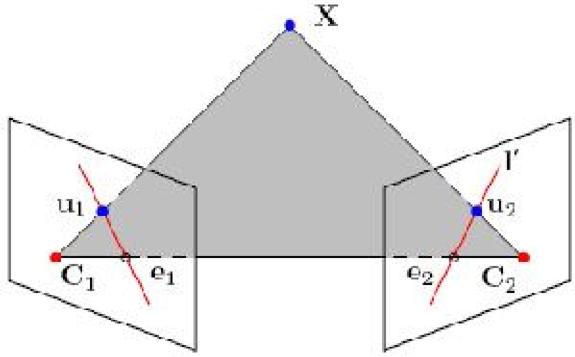
Let the diagonal matrix obtained from the SVD is $S = diag(r, s, t)$, then the correct rank 2 fundamental matrix F is given by:

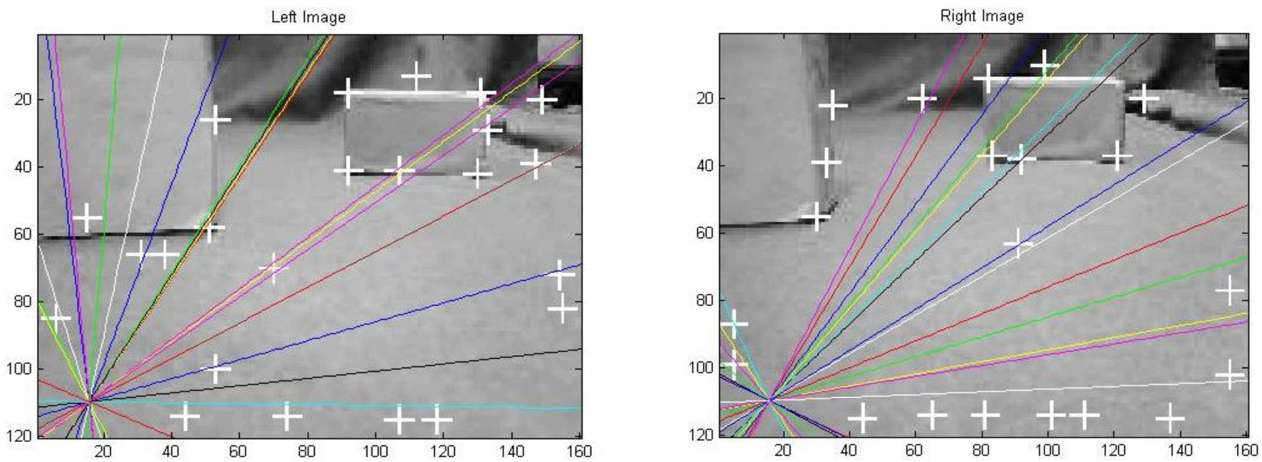
$$F = U \times diag(r, s, 0) \times V^T$$

To summarize, the steps for generic 8 - *Point Algorithm* are listed below.

- For each matched point, add a new row to matrix A mentioned above
- Find F' , an intermediate value for F by taking the singular value decomposition (SVD) of A
- Refine F' by taking SVD , setting the smallest diagonal value equal to zero and recombining

2. 2. 4 Epipolar Geometry The epipolar geometry is motivated by stereo matching i.e. by searching for the projections of a scene point X in two different views of the same rigid scene. Let's suppose that $X \in \mathbb{R}^3$ projects onto $u_1 \in \mathbb{P}^2$ in the first view and onto $u_2 \in \mathbb{P}^2$ in the second. From the fundamental properties of central projection follows that the centers of the cameras C_1 , C_2 and the points u_1 , u_2 and X are coplanar. Scene points together with baseline $\overline{C_1C_2}$ create a group of planes called *Epipolar Planes*. Each of such planes intersects the projective planes of the two views in straight line called *Epipolar Line*. The following figure illustrates the epipolar geometry.





(a) Left Image (b) Right Image

2. 2. 5 Feature Matching

The matching value of pixel p in one stereo image and pixel q in the other image is computed using the modified normalized cross-correlation with a window of suitable size defined by:

$$Matching\ Value\ (p, q) = \frac{\sum_{p_i \in A_p, q_j \in A_q} [(I_{p_i} - \bar{I}_p) \times (I_{q_j} - \bar{I}_q)]}{\left[\sum_{p_i \in A_p} (I_{p_i} - \bar{I}_p)^2 \times \sum_{q_j \in A_q} (I_{q_j} - \bar{I}_q)^2\right]^{1/2}}$$

where A_p and A_q are windows of same size with centers p and q respectively, I_{p_i} and I_{q_j} are the corresponding pixel intensities of pixels p_i and q_j and $\bar{I}_p$ and $\bar{I}_q$ are the average pixel intensities in the two windows A_p and A_q respectively. The SSD algorithm calculates the sum of the squared distances between corresponding points in the stereo images which when minimized gives the matched points (check). The location of the best match can further be refined to subpixel accuracy by fitting a quadratic polynomial curve to the minimum SSD value and its two neighbouring values. The location of the maximum value of the polynomial curve is the subpixel location. This can be implemented using HOUGH TRANSFORM.

2. 2. 6 Disparity Estimation

The disparity values are then determined by computing the EUCLIDEAN DISTANCE between each pair of corresponding points. Finally assemble all the disparity values into a map, where pixel intensities describe the relative depth of points within a scene. . Disparity Filling may be performed to estimate the Disparity values of unreliable pixels using both binocular and monocular information. This makes the Disparity image smoother.

Depth Map

2.3 Additional hardware used

2.3.1 Communications Interfaces:

Communication between PC and FB5 robot is established by XBee (802.15) wireless communication module. The baud rate is 9600 for two XBee modules, the destination address in XBee modules are serial number of another XBee module. Synchronization between two modules is done by following steps:

1. XBee1 connected to PC will send motion signals to XBee2 which present on the FB5 robot.
2. XBee2 will respond immediately to Xbee1 by acknowledgement signal.

As such both the modules should be programmed to work at a baud rate of 9600 with 8 data bits, no parity bit, and one stop bit and without handshaking signals.

2.3.2 Image Acquisition:

A pair of stereo images is captured with the use of two identical USB cameras mounted on
the FB5 robot at a pre-calculated distance between them. The output of these cameras is a YUY2_160 x 120 image which are used to generate the Depth Map thus the Position Map of the scene in front.

2.4 Design challenges

- Detection of obstacle through image processing involving of finding accurate distance of the objects from the robot.
- Removing errors in motion control regarding accurate distance traversing and angle rotation.
- Minimizing the processing time.
- Decision making for obstacle avoidance.

3. Implementation

2.3 Design steps

- Problem definition
- Depth Detection Algorithm development
  - Feature Points Detection
  - Fundamental Matrix computation
  - Feature Matching
  - Disparity Estimation
  - Depth/Position Mapping
  - Wireless Control
- Robot Navigation Algorithm Development
  - Position interrupt based Motion control
  - Wireless Communication
- Integration
- Implementation, testing and debugging

2.5 Timeline

- **27_09_10 to 2_10_10:**
  i. Code in C for controlling FB5 through XBee.

- **4_10_10 to 9_10_10:**
  i. Functions for FB5 to move in predefined steps in the direction transmitted through XBee.
  ii. Coding in Matlab for Feature points detection and epipolar geometry.

- **11_10_10 to 16_10_10:**
  i. C code for FB5 to move in controlled manner along a path derived from a predefined matrix.
  ii. Testing of Matlab code on pre-captured stereo pair images.

- **18_10_10 to 23_10_10:**
  i. Testing of the C code to move FB5 in predefined steps in the direction transmitted through XBee.
  ii. Code in Matlab for capturing pair of stereo images of specific objects (ball, cube, similarobjects available in lab) at intervals and storing in a folder as stereonL.bmp and stereonR.bmp from left n right camera respectively.
  iii. Code in Matlab for generating Depth Map from two stereo pair images.
  iv. Testing of Depth detection code on captured stereo image pairs.

2.5 Problems Faced

- First it was planned to used wireless Cameras using TV tuner cards. But The two TV tuner receivers could not be detected simultaneously because of their similar IDs.

  Possible solution:
  USB cameras were used to capture stereo pair images.

- Lighting conditions in the lab created shining of surfaces which gave false depth.

  Possible solution:
  Setting up an Arena with objects of predefined size and shape covered with none reflective materials and also positioning of the light sources improved the results.

- Xbee interference from neighboring teams caused unpredictability. Proximity overruled xbee id difference and caused interference.

  Possible solution:
  Wireless communication testing was done using USB serial communication.

- Precise 90 degree rotation by robot, requires that the wheels move exact shaft counts. Determining this shaft counts has been a pain. It varies from robot to robot and also between left rotation and right rotation.

  Possible solution:
  After trial and errors, the values for 90 deg rotation were fixed.

- The acknowledgement transmitted back by the Robot was not received in MATLAB which result the wait sequence to time-out causing delay and abrupt movement of the Robot.

  Possible solution:
  The MATLAB function for serial communication was worked out.

2.7 Performance metrics

- Correctness – Each individual object should be detected and depth should be calculated accurately.
- Safety – With more than one objects, safety issues of avoiding collision with the objects.

4. Innovation, Creativity and Reusability Index

4.1 Innovations in project

- A new approach, different from the traditional methods of object finding/grabbing which are based on color and/or shape detection, is developed to detect objects based on their positions, thus making it independent of object properties.

- Since the algorithm is based on position finding, it facilitates the mapping of objects in the scene, thus creating map of the area.

4.2 Concepts/modelling tools learnt in CS 684

The knowledge of modeling the design in the form of state machines helped to implement the design in a generic manner

4.3 Future Scope

- The Depth detection algorithm can be worked out to run faster and give more satisfactory results and accurate object detection.

- Making the image acquisition system wireless will facilitates Robot to traverse a distant location.

4.4 Conclusion

In this project we are able to achieve a system that detects obstacles with a considerable efficiency. The task proposed was successfully implemented and tested on a number of test setups with satisfactory results.

This project lays a stepping stone in the field of Position Mapping and object detection using Image processing.

From here the idea can be developed into a complete Embedded system without any human input required to work in distant locations where human expedition is not possible.

The setup may even be developed into a module which can be used to achieve other tasks such as object acquisition, target follow, area mapping, etc.

5. Appendices

5.1 References

[1] Chris Harris and Mike Stephens, *'A Combined Corner and Edge Detector'*, Plessey Research Roke Manor, United Kingdom, The Plessey Company plc. 1988.

[2] Feng Zhao, Qingming Huang, Wen Gao, *'Image Matching By Normalized Cross-Correlation',*

[3] Klaus, Sormann, and Karner**,** '*Segment-Based Stereo Matching Using Belief Propogation and Self-Adapting Dissimilarity Measure*'.

5.2 Documents

i. AVRstudio_software_manual
ii. FIRE BIRD V Hardware Manual
iii. ATMEGA2560 Processor Datasheet
iv. MATLAB Product Help

5.3 Deliverables

i. Readme
ii. Program source code, with inline documentation
iii. Final Presentation
iv. Attach CD with everything in it, video included.