Autonomous Multi-Sensor Platform with Stereo Vision

Paul Mugur Svasta, and Iaroslav Andrei Hapenciuc

CETTI Blvd. Iuliu Maniu, nr. 1-3
Complex LEU, Corp B, cam. 301,
paul.svasta@cetti.ro, andrei.hapenciuc@cetti.ro

Abstract: *This paper is describing a different way to investigate the terrain topography: stereo images. This is the closest way of terrain investigation to the human sight. By gathering as much information from images the need for other sensors on the robot will be no more. This is an obvious fact since for human more than 90% of information about the surrounding environment is coming from sight.*

An equipment to capture and process stereo images had been build and the software was created to prove the accuracy of this method to detect and avoid obstacles and to investigate the terrain topography for uncharted areas.

To extract the information needed for the navigation from the stereo optic images, complex algorithms of shape recognisance, pattern mach and cross correlation will be involved together with image calibration and sharpening. Because the processing power requirements will be huge methods to decrease the computation need will be used to increase the processing speed

The sensor and control board are build on one PCB. High power digital signal had to been as far as possible from the high accuracy analog inputs. The magnetic compass proves a very high accuracy: 0.5 degrees. For fast movements, the gyroscopes can give better measurement,. The triaxial accelerometer is a poor way to measure movement. It is mainly used as an inclinometer to help the algorithm of the electronic compass. For the measurement of speed and distance, an odometer is used. It is having a very good resolution, few millimeters, but it is insensitive to car sliding and mechanical interconnection imperfection. For these the accelerometer can be used to improve the result. The GPS results are used for less accurate measurements since its accuracy of positioning is around 5-10m.

The information from all sensors is centralized into a powerful MCU. The MCU alone can be used for controlling the robot movement on a simple route. For more complex situations, this acquisition module is connected to a small PC able to conduct more complex data analysis.

1. INTRODUCTION

The goal of this system is to give to an autonomous system the power of sight. The system tries to replicate as closely it can the human visual system, and so the two cameras connected in parallel resample a lot with two eyes. Yet, some of the features of human visual apparatus were not replicated or where implemented in a different manner. One of this is the fact that the cameras are parallel while the human eyes tend to converge both to a single point.

In order to obtain the desired results complex image processing techniques had been used, and adapted to the purpose. The main two parameters that drive the project intro actual configuration had been: accuracy and speed.

Figure 1: Acquisition system

In the process of robot guidance on a certain trajectory, sight have an essential role because it is implying complex algorithms, large amount of computation power, characterization and interpretation process of the information from the image of the 3D space in which the robot must move. This complex process implies more operations which are distinct phases:

- **Acquisition** is the process to transform de optic image in a digital image in the computer memory;
- **Preprocessing** is the process of image parameters improving: noise contrast, detail enhancement, etc;
- **Segmentation** is the process that is braking the image in smaller images;
- **Description** is the process that is computing different parameters for object identification;
- **Recognition** is the process of identifying the objects;
- **Interpretation** is the process that is giving sense to the entire picture.

Using advanced techniques in image processing it is possible to make a topographic investigation of the terrain using the most common and cheap image capturing devices: web cams.

The solution used in this project is using the stereo optical images for distance measurement to the most important terrain details.

These images will be captured by two web cam rigidly connected laterally to a distance of almost 15cm, similar as in human sight. The image from one camera will be slightly shifted laterally relative to the other images because of the parallax error. By measuring this parallax error it is possible to compute the distance according to the figure 2

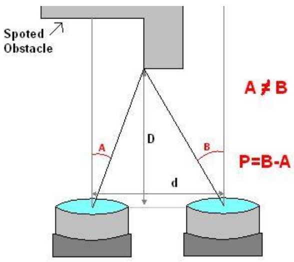

Figure 2: Distance measurement from parallax error

2. ACQUISITION

For this project, as acquisition devices had been selected the most common and cheap device: web cams. Them had been choose because of their advantages: low cost, electronic sweep of the image, colour identification. Them also have disadvantages: huge amount of the data that needs to be processed, high complexity of the algorithm to extract the terrain topography information. Similar to animals the systems is made from two “eyes” – the web cams, and a “brain” – the computer, as is visible in figure 1.

The GUI of the acquisition software in presented in figure 3:

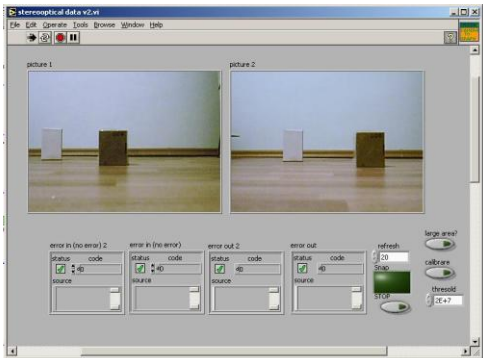

Figure 3: GUI of the acquisition software

The acquired images are 320x240 pixels in size ad the frame rate can reach 30fps.

3. PREPROCESSING

After acquisition, is necessary to bring parameters of the two images to very close values, so the image recognition techniques will be as accurate as possible.

Except this colour adjustment an image calibration will take place. The purpose of this calibration process is to make any item recoded with one of the camera on a certain line will be present on the image recorded by the other camera on the same line, but shifted laterally. This will increase the chances of correct object recognition, and decrease the risk to identify a similar but different object.

This calibration is in fact a position and rotation calibration, and is compensating the alignment errors if the web cams are not perfect parallel. The calibration is performed by a routine called Image Calibration.vi (see figure 4).

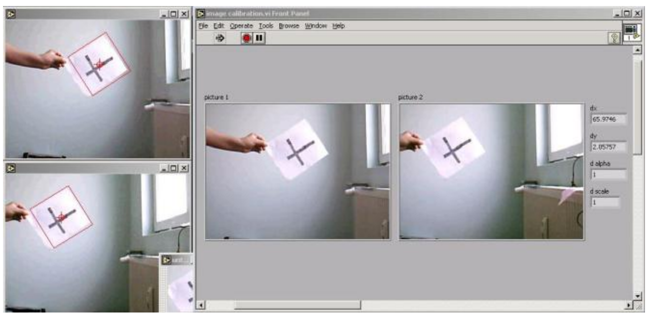

Figure 4: Image calibration.

For the detection of the calibration image it is used a pattern recognition technique invariable to rotation. Even more, because the detection of the position of the calibration image must be done with very high accuracy the parameter that is indicating the minimum score for a detected image is quite high 900 out of the 1000 which is the case of perfect identical objects.

The calibration image can be in any position and any orientation as long as it is entire seen by both web cams. Calibration coefficients will be very close, as is visible from figure 4.

4. Segmentation

Segmentation is the process thru which the big picture is divided into smaller high interest areas. Because in the case of navigation is impossible to predict the shape and the position of the obstacles, it will be used an algorithm that will detect interest area around of which the pattern recognition process will be more efficient.

At the base of this algorithm to detect good areas to be searched stays the fact that such an area to be find easier by the pattern recognition techniques needs to have a wide variation of pixel colour and intensity. The result of this detection algorithm is visible in the intensity graphs from figure 5.

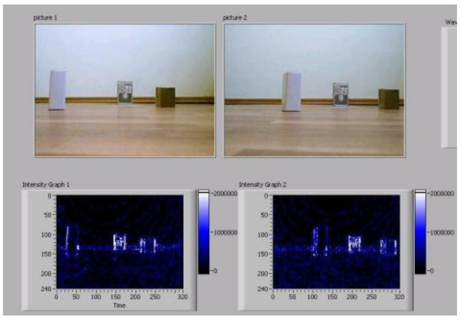

Figure 5: Detection of interest areas

The picture will be segmented vertically in 12 or 6 levels depending on the chosen size for the interest areas: 20x20 pixels or 40x40 pixels. For larger areas the recognition is done more precise but will be fewer such interest areas so the numbers of point to describe the 3D terrain will be smaller than for the case when are used 20x20 pixels area. In figure 8 are shown ho the image is segmented in 6 verticals levels, and the position of interest areas on the 3-th line.

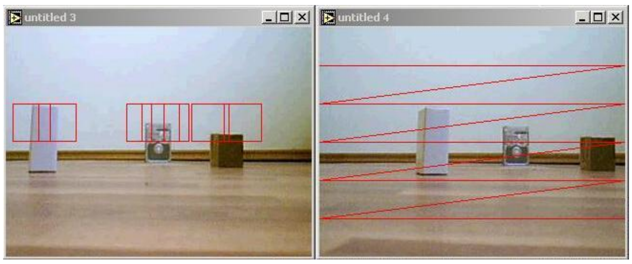

Figure 6: Segmentation for 40x40 pixels interest areas

As is visible in figure 6 the interest areas can overlap horizontally but not vertically. It would have been possible to overlaps them vertically also to obtain a higher number of levels that are describing the terrain but this it would have increase the number of interest areas a lot and because the computation power is limited it would have taken too much time. Good results had been obtained with interest areas of 20x20 pixels.

5. Recognition

5.1 Theory of shape recognition.

Shape recognition techniques are allowing the localization of areas from a picture which are similar with a template. The shape recognition techniques is finding images even if is low contrast, fog, noise, shifting or rotation of the image or template.

To use the shape recognition techniques first must be created an image of the object that will be searched this is the template. After this the shape recognition algorithm is searching the image for the template and is computing a score for each image found. This core is 1000 for perfect mach and 0 for no relation between the two images.

The following is the basic concept of correlation: Consider a sub-image w(x,y) of size K × L within an image f(x,y) of size M × N, where K ≤ M and L ≤ N. The correlation between w(x,y) and f(x,y) at a point (i,j) is given by where i = 0,1,…M – 1, j = 0,1… N – 1, and the summation is taken over the region in the image where w and f overlap.

$$C(i,j) = \sum_{x=0}^{L-1} \sum_{y=0}^{K-1} w(x,y)f(x+i,y+j) \qquad (1)$$

Figure 4 illustrates the correlation procedure. Assume that the origin of the image f is at the top left corner. Correlation is the process of moving the template or subimage w around the image area and computing the value C in that area. This involves multiplying each pixel in the template by the image pixel that it overlaps and then summing the results over all the pixels of the template. The maximum value of C indicates the position where w best matches f. Correlation values are not accurate at the borders of the image.

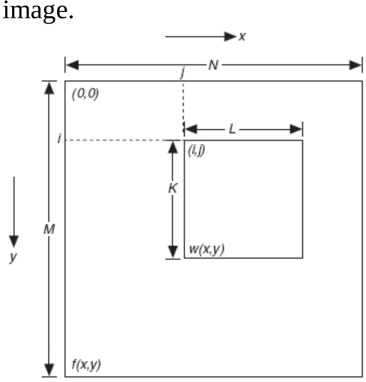

Figure 7: Correlation function computation

Basic correlation is very sensitive to amplitude changes in the image, such as intensity, and in the template. For example, if the intensity of the image f is doubled, so will the values of c. You can overcome sensitivity by computing the normalized correlation coefficient, which is defined as:

$$R(i,j) = \frac{\sum_{x=0}^{L-1}\sum_{y=0}^{K-1}(w(x,y)-\overline{w})(f(x+i,y+j)-\dot{f}(i,j))}{\left[\sum_{x=0}^{L-1}\sum_{y=0}^{K-1}(w(x,y)-\overline{w})^2\right]^{\frac{1}{2}}\left[\sum_{x=0}^{L-1}\sum_{y=0}^{K-1}(f(x+i,y+j)-f(i,j))^2\right]^{\frac{1}{2}}} \qquad (2)$$

Where w (calculated only once) is the average intensity value of the pixels in the template w. The variable f is the average value of f in the region coincident with the current location of w. The value of R lies in the range –1 to 1 and is independent of scale changes in the intensity values of f and w.

5.2 Algorithm implementation

Once image segmentation is completed, the interest area recognition process must occur. These interest areas are defined only in the left image and they are being searched on the right image but only on the same level as in the left image. This way, similar images from different levels will not interfere. The images from the right image will be slightly shifted to left because of the parallax error. By measuring this “left shift” it is possible to measure the distance up to that detail.

The shape recognition algorithm consider a successful find any image with a score higher than 600 – this is because the same object might look different in the two web cams. When a good score is computed and aim cross is drowned at the found image coordinate (see figure 8)

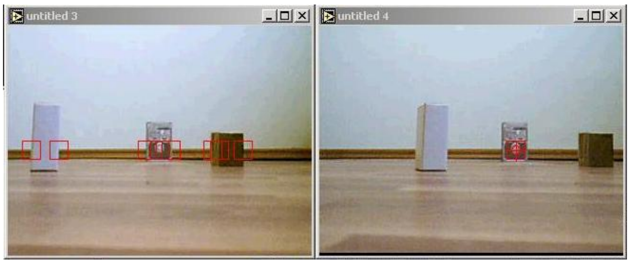

Figure 8: segmentation and recognition of 20x20 interest images

Because the image is segmented vertically on 12 levels, it is possible to draw a topographic map of the terrain with a resolution of 12 levels.

6. MEASUREMENTS

To test the entire system it was conducted an experiment in which a simple case of terrain topography was investigated. The 3D picture of the terrain is visible in figure 9 (at the left of the picture are visible the two electronic eyes).

Figure 9: Perspective image of the topography

The captured images by the stereoscopic system are presented in figure 10 – see the effect of calibration on the right image:

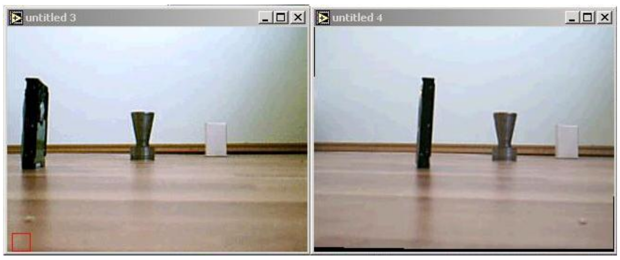

Figure 10: stereoscopic images

The result can be plotted in different forms. Each result is indicated by a circle with different colour according to the level on which the detail is present.

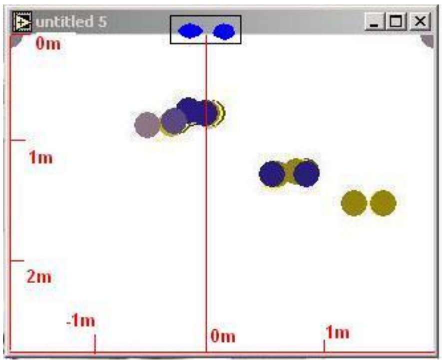

Figure 11: Result- top is 0m away from web cams, lower line is 2.4m from web cams

7. SYSTEM ARCHITECTURE

The architecture of the system in presented as a block diagram in figure 12. It is visible de interconnection of subsystems: at the left the PC –the brain of the robot-, in the middle the sensor and control unit that is giving senses to the machine and at the right the driving unit, equivalent with the muscles of a human been.

In figure 13 is a detailed block diagram of the acquisition and control unit. There is visible the complexity of the system, and the big number of peripheral devices used. In the following sections each peripheral will be described in detail. The picture of the acquisition and control block diagram is presented in figure 14.

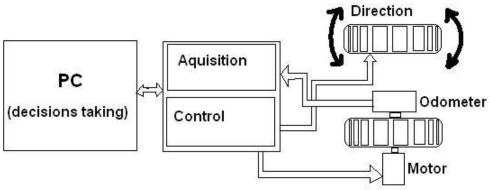

Figure 12 - System Block Diagram.

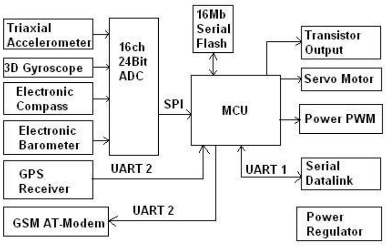

Figure 13 - Module Block Diagram.

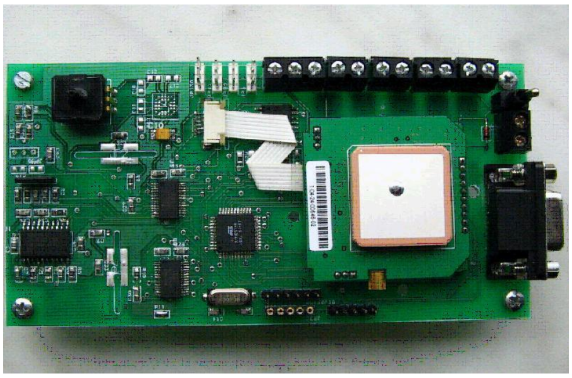

Figure 14:Acquisition and Control Module

7.1 Robot Mechanical Frame

For proving the feasibility of this project an mobile platform had been chosen to carry the equipment.

Because of the big weight and because the terrain where it will move won't be all the time plain it was chosen a 2WD OFF-ROAD auto model frame (figure 15). Each wheel is equipped with individual suspensions, making possible to run on difficult terrain. Each suspension can be preset for a certain lifting force so is possible to load the auto model with an asymmetric load, heavier at the back and easier in front.

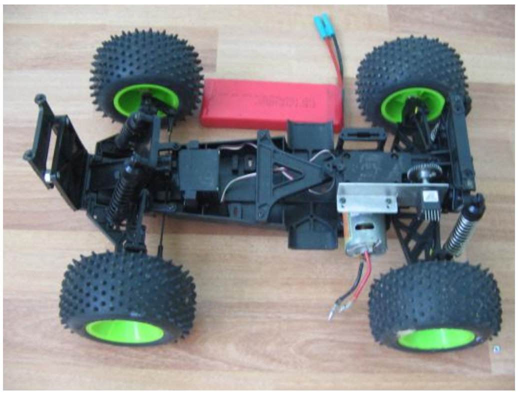

Figure 15. 2WD OFF-Road Frame

Driving wheels are at the back and are connected thru a differential. It is having only one gear and inside of it was mounted the odometer, which is in fact an optical quadrature encoder. To motor used is a 60W 12V brushed motor. The power source is a 3 cell Li-Po with 4000mAh capacity.

The front are the steering wheels connected to a 5kg*cm torque servomotor. The control of these servo motors are done thru short periodic pulses with a period of 20ms. The widths of these pulses are varying from 1ms – maximum left turn to 2ms – maximum right turn. For 1.5ms it is going straight.

7.2 Computer

For this project were tried more types of computers. The largest one was a made out of a mother board in mini-ATX format (25cmx24cm). Because the computation power of that processor was weak it was replaced by a SBC from AAEON model GENE-8310 (figure 16). This computer is in a 3.5' format and the processor – a ULV 1.6GHz Celeron – is using a passive heat sink. This is decreasing a lot the power requirements and noise.

Figure 16: GENE-8310 SBC

7.3 Electronic compass

The same way sailors are using the compass to indicate them the heading, the same way out autonomous navigation system needs a compass. But an ordinary one will be hard to integrate in the electronic system and fragile so an electronic compass had to be developed.

For accurate measurement of the earth magnetic field three magnetic sensors must be used. One for each Cartesian axe. This way cam be measured the direction and the amplitude of the magnetic field. It is important to know than not only the direction but also the intensity of the earth magnetic field is changing during a day, or depending to altitude or global position.

For having an accurate indication of the magnetic north pole is important to know the horizontal plane in the location. For this purpose the triaxial accelerometer will be very handy.

To get the information from all three axes will be used 2 ICs: HMC1001 and HMC1002 from Honeywell. The HMC1001 is sensitive on one axe, vertical, and HMC1002 is having two axes of sensitivity, both horizontal:

Figure 17 - Magnetic sensors

If the position of the earth magnetic vector is known in the 3D space, and the position of the local horizontal plane then by doing the projection of the magnetic vector on the horizontal plane is obtained the correct direction of the magnetic north pole. It is important to underline that if the correct position of the horizontal is not know, the direction of the magnetic north pole will be wrong as is visible from figure 18.

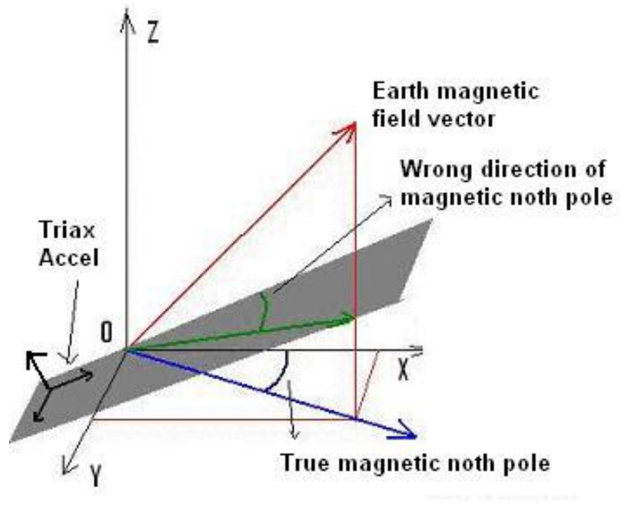

Figure 18 – Electronic compass method

The accuracy of determining the direction and magnitude of the earth magnetic field is one again limited by the sensors noise. The two types have a value of 29nV/Hz^-2, while de sensitivity is 12mV/Gauss, since them are powered from 3.3V.

The magnetic field intensity is around 0.5Gauss. Do determine the error in measuring it must be evaluated the accuracy of the sensors. At a bandwidth of 10Hz the noise is around 92nV, at a sensitivity of 12mV/Gauss the accuracy will be: 8uGauss.

At such low values of the signal the accuracy of the ADC come into question. The ADC used is a 24Bit model, but practically only 20bits are trusted. At a voltage reference of 3.3V the accuracy of the ADC is 3uV this value is higher than the accuracy of the sensors (92nV) and so it is not possible to take advantage of the full sensor accuracy.

With a 64x gain, the signal is amplified in the same time with the noise, and the new value will be: Sensitivity 768mV/gauss and noise: 6uV. Because now the sensor noise is higher than the ACD noise it will be possible to use the measurement system at the maximum accuracy: 8uGauss.

With such accuracy the angular errors will be under 0.5degrees. This value is a lot better than the one obtained from gyroscopes, then why are not used only these ones. The answer is that three magnetic sensors can’t determine alone any arbitrary position in space: will be 2 ore more positions in space that will give the same values at the output of the sensors.

As long as these positions are fairly different one from other them can be distinguished by using the info from gyroscopes, as long as the information from them is accurate enough to separate the positions with common output.

8. CONCLUSIONS

By using cheap hardware was possible to build a system able of terrain topography investigation. The resulted precision is influenced by a serried of factors like: shape recognition algorithm accuracy, image acquisition performance, and last but not least the computation power of the PC on which the entire system is implemented.

By optimising the image processing algorithms it was possible to get performance similar to the one based on LASER telemetry, the main difference is the mechanical construction and price.

Even more, there are situations when obstacles even visible are thousands of km away so a laser telemeter might become inefficient.

9. REFERENCES

[1] National instruments, “IMAQ Vision – concept manual”, www.ni.com

[2] National instruments, “IMAQ Vision for LabVIEW – User Manual”, www.ni.com

[3] Xilinx ISE WebPack www.xilinx.com

[4] PixART CMOS Sensor User Manual www.PixART.com.tw

[1] www.Atmel.com: "Atmega162 Datasheet"

[2] Integrated Power Module, www.irf.com

[3] Freescale Semiconductor, MEMS Triaxial accelerometers:

[4] Analog Device, iMEMS Gyroscopes.

[5] www.Memsic.com: “Inclination sensing of moving vehicles”

[6] Colonel Engineer Ion Musat, PhD. General Nicolae Irimie: “Considerations regarding the conception of a computing system of parameters for rocket guidance”. . 1996