DEPTH ESTIMATION AND 3-D RECONSTRUCTION ON HARDWARE SOFTWARE PLATFORM

APURVA NAIK, , GOURANG MULAY , ARTI KHAPARDE ,YOGESH PATKI, SAGAR POKALE AND SWAPNIL NALAWADE

Department of Electronics & Telecommunication, Maharashtra Institute Of Technology, Pune. India

Email: apurva.naik, gourang.mulay, arti.khaparde@mitpune.edu.in

ABSTRACT: Two or more images of same scene or object can compute representation of its shape using well known technique called stereo matching. Result of stereo matching can be utilized to extract depth information, generation of volumetric models, 3-D surface models,. We propose generation of 3D image using a stereo image. The innovativeness of the proposed method is 3-D generation from stereo using hardware platform and segmented stereo image matching. The 3D image generated on SAM9M10-G45 ARM -9 controller based board, making it portable low cost system for 3-D generation useful for industrial application. Prior to 3-D generation on ARM-9 board images are segmented using Particle swarm Optimization algorithm to reduce number of computations required for stereo matching and 3-D view is generated on PC as well as TFT display on the SAM-9 board.

KEYWORDS: Particle Swarm Optimization, SAM9M10-G45 ARM -9

INTRODUCTION

Stereo matching is the process of taking two or more images of same scene and estimating 3D model of the scene by finding matching pixels in the images and converting their 2D positions into 3D depths.

In order to recover depth information the stereo images should be brought into point-point correspondence. Correspondence points are the projections of a single point into the three dimensional scene. The difference between the locations of these two correspondence points is known as parallax or disparity, which is a function of position of the point in the scene, orientation and physical characteristics of the camera. Under simple imaging configurations amount of horizontal motion or disparity is inversely proportional to distance from observer. Measuring this disparity by establishing dense and accurate inter image correspondences is a challenging task. In computer vision the topic of stereo matching has been one of the most widely studied and fundamental problem and one of the most active research areas. Disparity Map is a 3D contour (three dimensions being X, Y and color) which illustrates the distances of various pixels from the viewer, as well as from each other. Till date many algorithms have been carried out to construct 3D images from a stereo image pair and later find out the disparity related to the same pair. These algorithms have been verified virtually using various compute units including CPUs, GPUs, and DSPs. So the research related to the stereo image processing has still been largely dependent on processing power. Creating computer vision hardware accelerators is difficult since applications must be portable. The standard C++ language path offers flexibility to change and portability across platforms.

Proposed system implements similar algorithms, scaled down to suit lower processing power, using lower resolution images for better output performance and using a more general programming platform like embedded C so as to suit any real time system.

Using a pair of stereo images, system is able to provide a 3D image in real time, keeping the details of produced image acceptable to human eye. System also provides a Disparity map which is a spatial representation of depths of various objects in the image.

In this paper, we are proposing an area-based segmentation method, as it generates a dense disparity map for 3D reconstruction. For some applications, such as image recognition or stereo vision, whole images cannot be processed, as it not only increases the computational complexity, but it also requires more memory. In literature, number of image segmentation algorithms like K-means [1], mean-shift, Particle Swarm Optimization (PSO) etc., have been proposed and extensively applied to stereo vision. The algorithm assumes that disparity values vary smoothly in those regions and that depth discontinuities only occur on region boundaries. Purely pixel-based methods are insufficient to express information of the image. The human identifies the objects by analyzing features of the objects such as color, texture and shape. Thus, segmentation-based stereo matching algorithm should be used. Segment-based methods have attracted attention due to their good performance on handling boundaries and texture less regions. They are based on the assumption that the scene structure can be approximated by a set of non-overlapping planes in the disparity space and that each plane of target image is coincident with at least one homogeneous color segment in the reference image. Segment based methods perform well in reducing the ambiguity associated with texture less regions and enhancing noise tolerance. The computational complexity is reduced due to much larger segments. Particle Swarm Optimization based segmentation algorithm is used prior to matching of stereo using MATLAB.

LITERATURE STUDIES

Stereo vision algorithms' theoretical matching cores are quite well established leading the researchers towards innovations resulting in more efficient hardware implementations such as FPGA /ASIC and embedded processors. Several biologically inspired algorithms have been explored in image segmentation [2]. Bio-inspired algorithms have been used in situations where conventional optimization techniques cannot find satisfactory solutions. For e.g., when the function to be optimized is discontinuous, non-differentiable and / or represents too many non-linearly related parameters [3]. Particle Swarm Optimization (PSO) is a one such machine learning algorithm. [4-5]

All stereo vision algorithms have a way of measuring the similarity of images. Typically matching cost is computed at each pixel. Common pixel based matching costs include sum of absolute or squared differences (SAD/SSD), normalized cross-correlation (NCC) and rank and census transform [6]. Matching costs that explicitly or implicitly handle radiometric differences has been evaluated.[7] Reliability of depth maps and computational cost of algorithm is key issue for implementing real time robust applications. Many studies suggesting reliable algorithms were made and some studies comparing performance of the algorithm can be found in the literature. [8]. Stereo algorithms based on local correspondences [9]-[10] are typically fast, but require an adequate choice of window size. This leads to trade-off between low matching ratios for small window sizes and border bleeding artefacts, for larger ones. As a consequence, poorly textured and ambiguous surfaces cannot be matched consistently.

Algorithms based on global correspondences overcome some of the aforementioned problems by imposing smoothness constraints on the disparities in the form of regularized energy functions. Optimizing such Markov random Field (MRF) based energy functions is in general Non-parametric e.g. Graph cuts [11] or belief propagation [12]. However, even on low resolution imagery they generally require large computational efforts and high memory capacities. Local methods typically aggregate image statistics in a small window. Thus imposes smoothness implicitly. Optimization is usually performed using winner takes all strategy which selects for each pixel the disparity with smallest value under some distance metric Real time performance is achieved using census transform and GPU implementation.
Dense and accurate matching can be obtained by global methods which enforce smoothness explicitly by minimizing MRF based energy function which can be decomposed as the sum of data fitting term and regularization term. Latest trends in these fields mainly demands real time execution speed with decent accuracy.

MOTIVATION

System can be utilized for augmented reality application for equipment maintenance using wearable display device such as glasses or goggles and wearable cameras.[13] Processing can be distributed between wearable device incorporating controller based hardware (in our case SAM-9) and server (CPU with MATLAB) as shown in figure 1.

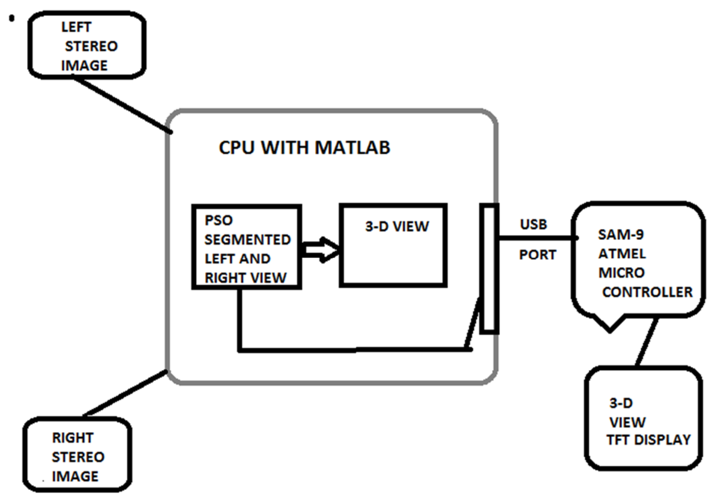


Figure 1. Block Diagram of System used for 3-D Generation and Depth Calculation

The server system can display on demand various 3-D views of any machine part from any view point. The wearable display allows service technician to look at the machine and view augmented reality overlays on the display illustrating how to service the machine. The 3-D depth surfaces can be obtained on PC server processing using MATLAB as well as on SAM -9.In future PC and SAM -9 can be connected with wireless module like ZIGBEE. Depth level calculations can also be used for wireless control of robot.[14].Though depth generation and 3-D cameras are available they are very costly. Our system will provide low cost solution for the same.

PROBLEM DOMAIN

The computer vision system of compressed 3-D generation and depth estimation providing reduction in stereo matching computational time on central server PC as well as generation of 3-D on hardware like ATMEL SAM-9 controller programmed using embedded c. Stereo matching problem is optimized at algorithm level and hardware level.

PROBLEM DEFINITION

The problem is to implement stereo correspondence algorithm on hardware processor for 3D image formation as well as real depth estimation making portable system for computer vision application.
Prior to stereo correspondence images are compressed using segmentation algorithm more specifically clustering based algorithm.

INNOVATIVENESS

The present invention provides a system and a method for segmentation-based depth estimation of stereo images. The method involves segmentation of real time 2D images for generation of stereo 3D images and estimation of depth levels from the 3D images. The stereo images and the depth levels estimated there from require minimum memory for storage and can be transmitted with high speed for use in various computer vision applications.

Another advantage of invention is implementation platform of embedded C and ATMEL processor making it suitable for use in portable equipment.

PROBLEM DESIGN
After application of segmentation of stereo image using Particle swarm optimization[1] and 3-D generation on PC server as shown in figure 2 an arm9 MPU has been used for implementation of algorithms A and B.

Algorithm 1 : 3-D Generation from stereo image with out segmentation

1. Align the two cameras taking into account the baseline distance (T) between the two cameras. ($T \approx 6.5$cm) [2]
2. Take the images from both the camera at the same instance of time. Store the left and right image separately.
3. Extract red channel information from the left image.
4. Extract cyan channel information from the right image.
5. Overlap this information to form a new image.
6. Display the image.

Algorithm 2: 3-D Generation from stereo image with PSO segmentation

1. Align the two cameras taking into account the baseline distance (T) between the two cameras. ($T \approx 6.5$cm)
2. Take the images from both the camera at the same instance of time. Store the left and right image separately.

3. Using the particle swarm optimization (PSO) algorithm calculate the segmented images of the given real time images. and store both the images
4. Take the segmented images as input and apply the Winners take all algorithm to calculate the disparity of each point in in the image.
5. Using disparities of the all points in the image calculate the disparity map, and display it.

Standard stereoscopic image pairs available on web are used as source images while creating 3D image and plotting the disparity map. The real time images will require camera calibration as additional step to give accurate stereo image pair. These standard images were loaded onto the SAM-9 kit using PC interface- Sam-Ba. Then using the embedded C program a generated 3D image was displayed on the LCD. Figure 2 shows **SAM9M10 G45 – EK** board used for 3-D image generation.

Figure 2. The development board

SAM9M10 G45 – EK

This is a media rich kit from Atmel. This kit features a powerful ARM9 processor, has a in built video decoder, a TFT display (touch screen supported), large storage capacity and fast serial communication.

Specifications
The SAM9M10-G45-EK is for evaluation of the SAM9M10 and the SAM9G45, two high-performance embedded MPUs optimized for industrial applications. The board includes the following features:

- Processing clock: 400 MHz Memory Clock 133 MHz
- Two high-speed USB hosts and one high-speed device port
- One Ethernet 10/100 interface

- Two high-speed multimedia card interfaces
- One LCD TFT display (480*RGB*272) with touch panel
- One composite video output
- One camera interface
- Several communication peripherals such as:
    - Universal Synchronous/Asynchronous Receiver Transmitter (USART)
    - Serial Synchronous Controller (SSC)
    - Two-Wire Interface (TWI)
- The external memory block is made of 3 memory types:
    - DDR2-SDRAM
    - NAND Flash
    - NOR Flash

Camera interface and quality processor make this kit ideal for image processing. The programming of this kit can be done through Keil microVision; it requires embedded c programming syntax. Figure 3 shows Major components of SAM9M10-G45 Board

Sam-Ba has been used as the programmer for the kit. This software is available from Atmel to burn programs on SAM9M IC on this kit, and then execute.

The board features Windows CE and Linux CE versions, embedded into it. Jumpers J9, J10 and J12 are closed along with all other jumpers when the kit is to be booted in one of the embedded OS. However, we have kept those jumpers open as we wanted to run the kit in programming mode. These connections are suggested in the kit user guide provided by Atmel. More details can be obtained from [15].

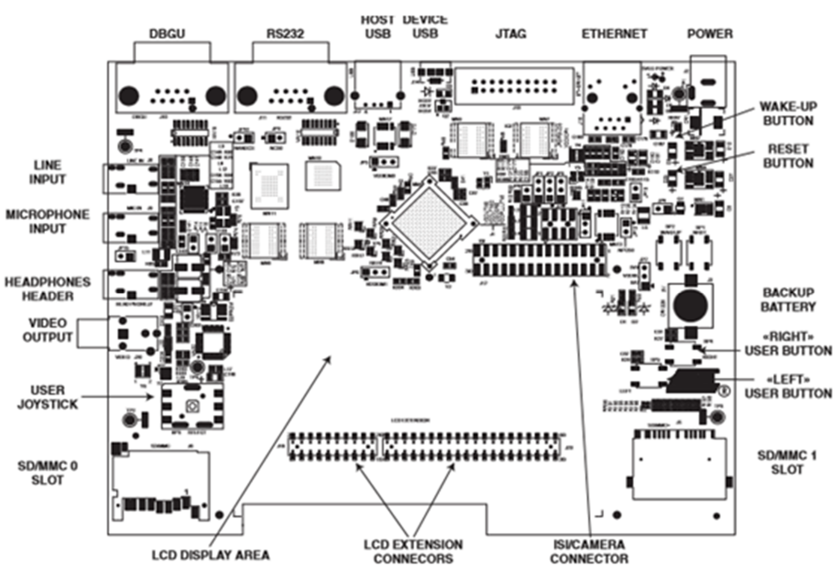


Figure 3 Major components of SAM9M10-G45 Board

Solution Methodologies

Stereo images are converted in .raw format using Photoshop CS6 and stored in DDRAM of the kit. The SAM9M10 features a DDR2/LPDDR memory interface and an External Bus Interface (EBI) to permit interfacing to a wide range of external memories and to almost any kind of parallel peripheral.

The SAM9M10-G45-EK board is equipped with DDR2/LPDDR devices featuring 128 MB of DDR2-SDRAM memory (16Meg*8*4).

The SAM9M10 features LCD controller. A 4.3" 480x272 Portrait Mode LCD provides the SAM9M10- G45-EK with a low power LCD display, back light unit and a touch panel, similar to that used on commercial PDAs.

Graphics and text can be displayed on the dot matrix panel with up to 16 million colors by supplying 24-bit data signals (8bitxRGB by default) or 16-bit data signals (5+6+5bitxRGB in option). This allows the user to develop graphical user interfaces for a wide variety of end applications.

Displaying image on LCD of the kit:

Two images were stored in the DDRAM of the kit, and were alternately displayed on the LCD.

1. Create a project for SAM9M10 using keil u-Vision 4
2. Build the project to obtain .bin file.
3. Use the Sam-Ba interface to as shown in Figure 4 connect the kit to computer.

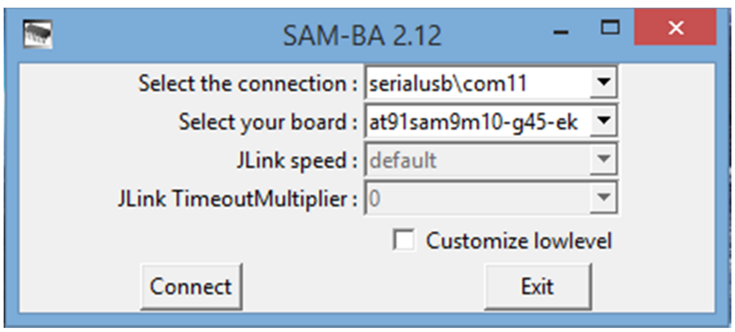


Figure 4 Sam-Ba interface connections with computer

4. Send the .bin file to the DDRAM of sam9.
5. Send images to be displayed in .raw format to locations specified in program.
6. Execute the .bin file using command window of Sam-Ba.

Algorithm implemented for 3-D image display is winner take all shown in figure 5.

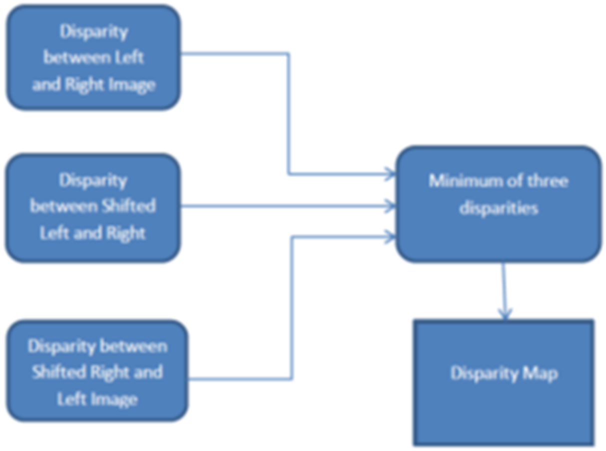


Figure 5 Basic steps in Winner Take all Algorithm for depth estimation Results

3D image obtained using MATLAB are shown in figure 6.These 3D images are obtained using segmented versions stereo images. The function used in MATLAB was 'cat'. Same function is implemented in optimized way using embedded C to obtain similar results on the SAM-9 kit.
3D image on the SAM9M10 kit is shown in figure 7

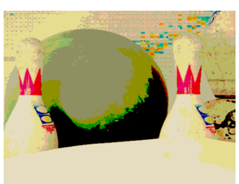
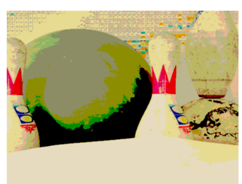
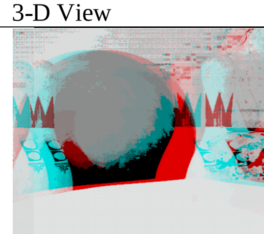
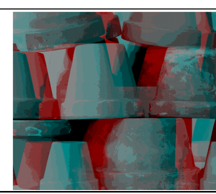
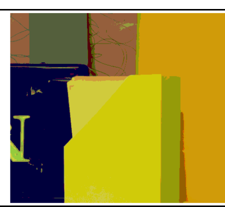
Segmented left	Segmented Right	3-D View
		
		
		

Figure 6 3D image using MATLAB: PC: Intel i5, 1.8GHz): 0.001608 seconds

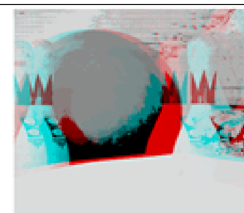
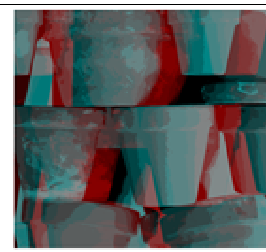
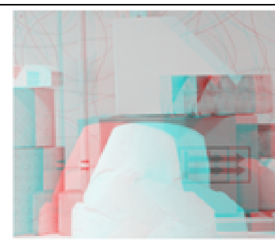
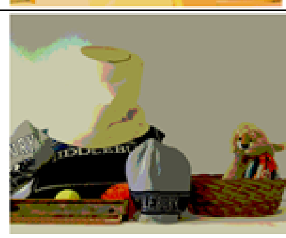
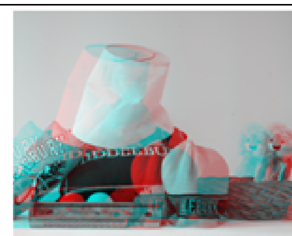
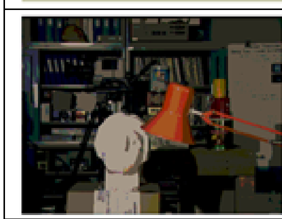
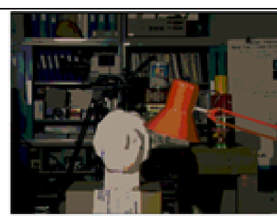
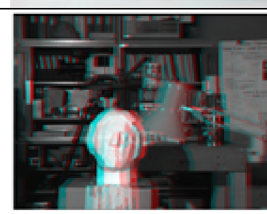
Segmented left	Segmented Right	3-D View
		
		
		
		
		

Figure 7 3D image on the SAM9M10 kit: Execution Time: (@ 400 MHz): 0.004896 seconds

CONCLUSION

3D image was obtained from a stereo image pair. The quality of image generated was acceptable to human eye, and the speed at which the image was obtained was also sufficiently fast. The Disparity Map was also obtained but was not processed with empirical formula yet.

Though the execution time for SAM9M10 kit was almost 4 times that of MATLAB, comparing the clock speeds, it was certainly faster than MATLAB. So we can imply that on a MPU even faster than SAM9M10, the algorithm developed will prove more efficient. The quality of the image and perception of depth in the image has been reflected in resultant image satisfactorily.

So in this era of real time system development, this system can deliver great results. 3D images of field view can be obtained and stored to help develop models, multi angle capturing of the same field can help in motion analysis. 3D imaging can further be extended to satellite image capturing, and depending on accuracies of cameras, space modelling is also possible.

FUTURE WORK

Segmentation of stereo images will be done on hardware using various algorithms like PSO, DPSO, FODPSO and then stereo matching will be carried out. Real time images will be taken through camera and after camera calibration 3-D view will be generated.

REFERENCES

Arti Khaparde, Apurva Naik, Manini Deshpande, Sakshi Khar, Kshitija Pandhari, and Mayura Shewale, "Performance Analysis of Stereo Matching Using Segmentation Based Disparity Map", ICDT 2013:The Eighth International Conference on Digital Telecommunications, 21-26, April 2013, Venice, Italy, pp. 38-43.

Kulkarni R.V., and Venayagamoorthy G.K., "Bio-inspired Algorithms for autonomous deployment and localization of Sensor", IEEE Transactions on Systems, 2010, 40(6), 663-675.

Floreano D., and Mattiussi C, "Bio-Inspired Artificial Intelligence: Theories, Methods and Technologies, Cambridge, MA:MIT Press, 2008.

A. Klaus, M. Sormann, and Karner, “K-Segment-Based Stereo Matching Using Belief Propagation “Pattern Recognition, 2006. ICPR2006. 18th International Conference, Page(s):15-18

D. Comaniciu and P. Meer, “Mean shift:A Robust Approach Toward Feature Space Analysis”, Pattern Analysis and Machine Intelligence, IEEE Transactions 2002, Page(s):603-619 D. Comaniciu and P. Meer, “Mean shift:A Robust Approach Toward Feature Space Analysis”, Pattern Analysis and Machine Intelligence, IEEE Transactions 2002, Page(s):603-619

Martin Humenberger, Christian Zinner, Michael Weber, Wilfried Kubinger,Markus Vincze, “A Fast Stereo Matching Algorithm suitable for embedded real time system”, Journal of computer vision and image understanding, 114(11), 2010, 1180-1202.

Baris Baykant Alagoz, "Obtaining Depth Maps From Colour Images By Region Based Stereo Matching Algorithms.", OncuBlim Algorithm and Systems Labs, 2008, Vol.08, Art. No: 04.

Daniel Scharstein, and Richard Szeliski, “A Taxonomy and evaluation of dense two frame stereo correspondence algorithms”, IJCV, volume 47, April 2002, Page(s) 7-42.

Heiko Hirschmuller and Daniel Scharstein, “Evaluation of costs functions for stereo matching”, CVPR IEEE Computer Society (2007).
Nalpantidis Lazaros, Georgious Christou Sirakoulis, and Antonious Gasteratos, “Review of stereo vision algorithms: From software to hardware”, International journal of Optomechatronics, 2008, Vol. 2, issue 4, 435-462.
Kolmogorov and V.Zabih, "Computing Visual Correspondence With Occlusion Using Graph Cuts", International Conference on Computer Vision, 2001, 508-515.
Felzenszwalb P., and Huttenlocher D., "Efficient Belief Propagation for Early Vision", International Journal of Computer Vision, 70(2006), 41-54.
Computer Vision Metrics, Survey, Taxonomy, and Analysis, Scott Krig, Apress Open, 2014
Apurva Naik, Keshav Velhal, Kunal Shah, Pratik Raut, Arti Khaparde “Wireless Transmission of Stereo Images and its Disparity Levels” IMMM 2014 : The Fourth International Conference on Advances in Information Mining and Management