ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 2, Issue 8, February 2013

Comparison of Various Stereo Vision Cost Aggregation Methods

Suyog Patil[1], Joseph Simon Nadar[2], Jimit Gada[3], Siddhartha Motghare[4], Sujath S Nair[5]

***Abstract*— This paper explains various cost aggregation methods used in stereo vision systems and does comparison between these methods to find out optimum method for a given cost and for a given set of resources. The paper also throws light on optimum method for implementation of stereo vision system on FPGA's and SOC's. The paper concludes with the fact that though methods like NCC and ZNCC offer higher accuracy but they are computationally heavy for embedding into real time systems. SAD and SSD are suitable choice for embedded systems. However applications requiring higher accuracy should use correlation based method.**

***Index Terms*— Stereo Vision, 3D Reconstruction, Disparity, Cost Aggregation.**

I. INTRODUCTION

Stereo Vision systems are recently being used in consumer electronics which has resulted in much research in this field and introduction of new algorithms and cost aggregation methods. Most of the research is now being concentrated to make the algorithms more efficient and less time consuming so as to embed it in real time systems like mobiles and surveillance systems.

Stereo vision systems basically consist of three processes. The first is of the cost computation which finds out the required features between two images which may be intensity based features or colour based or other structural features like the edges and gradients across the image. The first step may vary according to various algorithms as distinct algorithms may use different approach to find correspondence between two images. However most of the algorithms use the same second step of cost matching function which may be pixel based or window based. Usually the cost is aggregated over a window. This aggregation of cost is the most important part of any stereo vision algorithm. The final step is of proper disparity selection based on the cost aggregated over the entire image. The output of this step is the disparity map which gives us the depth information of the two input images.

Fig 1: Steps in a stereo vision algorithm

This paper deals with the second step which is cost aggregation. The paper is an analysis of each method and finds out advantages and disadvantages of the methods over one another. All the methods have been implemented on matlab and the images used are the standard set of images provided by the Middlebury website [1].

II. COST AGGREGATION METHODS

There are many cost aggregation methods listed in literature but this paper deals with more widely used and easy to embed algorithms. [2], [6], [8]

Some of the notations used are

$I_L(x, y)$ –Left Image Intensity at point (x, y).

$I_R(x, y)$ –Right Image Intensity at point (x, y).

M_L – Mean Left Image Intensity over window W.

M_R –Mean Right Image Intensity over window W.

A. Sum of Absolute Difference(SAD)

In SAD method the correspondence is achieved by selecting a window of required dimension within the two images or the cost matrix and adding the difference between the elements over the entire window. This method is less time consuming but generally is susceptible to outliers.

$$SAD = \sum_W \{I_L(x, y) - I_R(x, y - d)\}$$

B. Zero Mean Sum of Absolute Difference(ZSAD)

It is similar to SAD but the mean of the window is subtracted from the aggregated cost. Due to this the cost aggregation is immune to linear intensity variations between the two images.

$$ZSAD = \sum_W |I_L(x, y) - M_L - I_R(x, y - d) + M_R|$$

C. Sum of Squared Difference(SSD)

In Sum of Squared Differences (SSD), the differences are squared and aggregated within a square window. This measure has a higher computational complexity compared to SAD algorithm as it involves numerous multiplication operations.

$$SSD = \sum_W \{I_L(x, y) - I_R(x, y - d)\}^2$$

D. Zero Mean Sum of Squared Difference(ZSSD)

In ZSSD method the cost is determined as in SSD also by subtracting the mean of the match area from each intensity value. However, subtracting the mean from the squared

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 2, Issue 8, February 2013

intensities adds to the computational complexity of the aggregation method.

$$ZSSD = \sum_{W} (I_L(x,y) - M_L - I_R(x,y-d) + M_R)^2$$

E. Normalized Cross Correlation(NCC)

In NCC method of cost aggregation a window of suitable size is determined and moved over the entire image or the cost matrix. The correspondence is thus obtained by dividing the normalized summation of the product of intensities over the entire window by the standard deviation of the intensities of the images over the entire window [3], [5]

$$NCC = \frac{\sum_w \{I_L(x,y).I_R(x,y-d)\}}{\sqrt{\sum_w I_L^2(x,y).\sum_w I_R^2(x,y-d)}}$$

F. Zero Mean Normalized Cross Correlation(ZNCC)

ZNCC is similar to NCC with the only difference of subtracting the local mean value of intensities. Thus the computation technique remains the same and correspondence can be obtained in the same way as NCC [3], [5]

$$ZNCC = \frac{\sum_w\{[I_L(x,y) - M_L].[I_R(x,y-d) - M_R]\}}{\sqrt{\sum_w[I_L(x,y) - M_L]^2.\sum_w[I_R(x,y-d) - M_R]^2}}$$

III. IMPLEMENTATION

All the aggregation methods have been implemented on MATLAB and the image used is the standard image Tsukuba which is provided by [1]. The codes for the methods were run on a machine with Intel Core i5 processor with inbuilt graphics. The codes for the listed methods were optimised to run efficiently in MATLAB environment. We have extensively used operations on matrices avoiding loops. This resulted in faster and efficient execution of methods.

IV. COMPARISON

SAD is the simplest algorithm and can be easily embedded into FPGA or SOC. SAD is an ideal choice for real time systems but it is more prone to intensity variation among the two images. SSD method has a higher computational complexity compared to SAD algorithm as it involves numerous multiplication operations. Normally, two areas which consist of exactly the same pixel values would yield a score of zero. However, these measures will no longer yield the correct results in the case of radiometric distortion, i.e., where the pixel values in one image differ from those in the other image by a constant offset and/or gain factor. The ZSAD and ZSSD have been devised to deal with this problem. However, the improved performance of the ZSAD over SAD and the ZSSD over SSD is offset by substantially increased computational complexity.

NCC algorithm is robust to the linear variation in the brightness due to different illumination conditions to the cameras. But it can be seen that due to more complex calculations of division, multiplication and square root its computational time is more than SAD, SSD. Hence it could only be used for real time application only if we are able to develop more efficient algorithm to speed up matching process. ZNCC algorithm has the advantages of NCC embedded into it. Like NCC, ZNCC gives good results as compared to other methods. ZNCC is immune to intensity distortions which is an added advantage over NCC. But it is computationally more expensive than NCC.

NCC and ZNCC give very good results with smaller windows whereas all other methods give proper results but with salt and pepper noise added to it. Therefore, NCC and ZNCC are proper for smaller windows and SAD, SSD, ZSAD and ZSSD usually give good results with larger windows which add to computational time.

A. Image and its Ground Truth

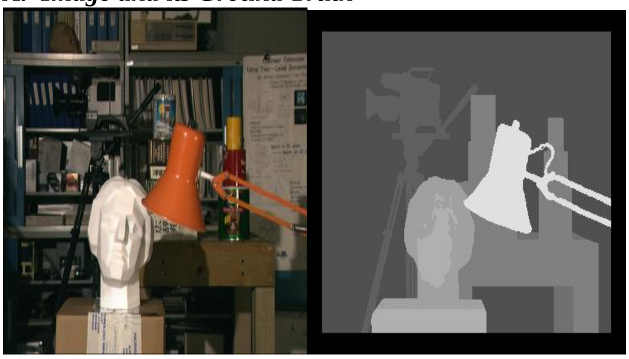
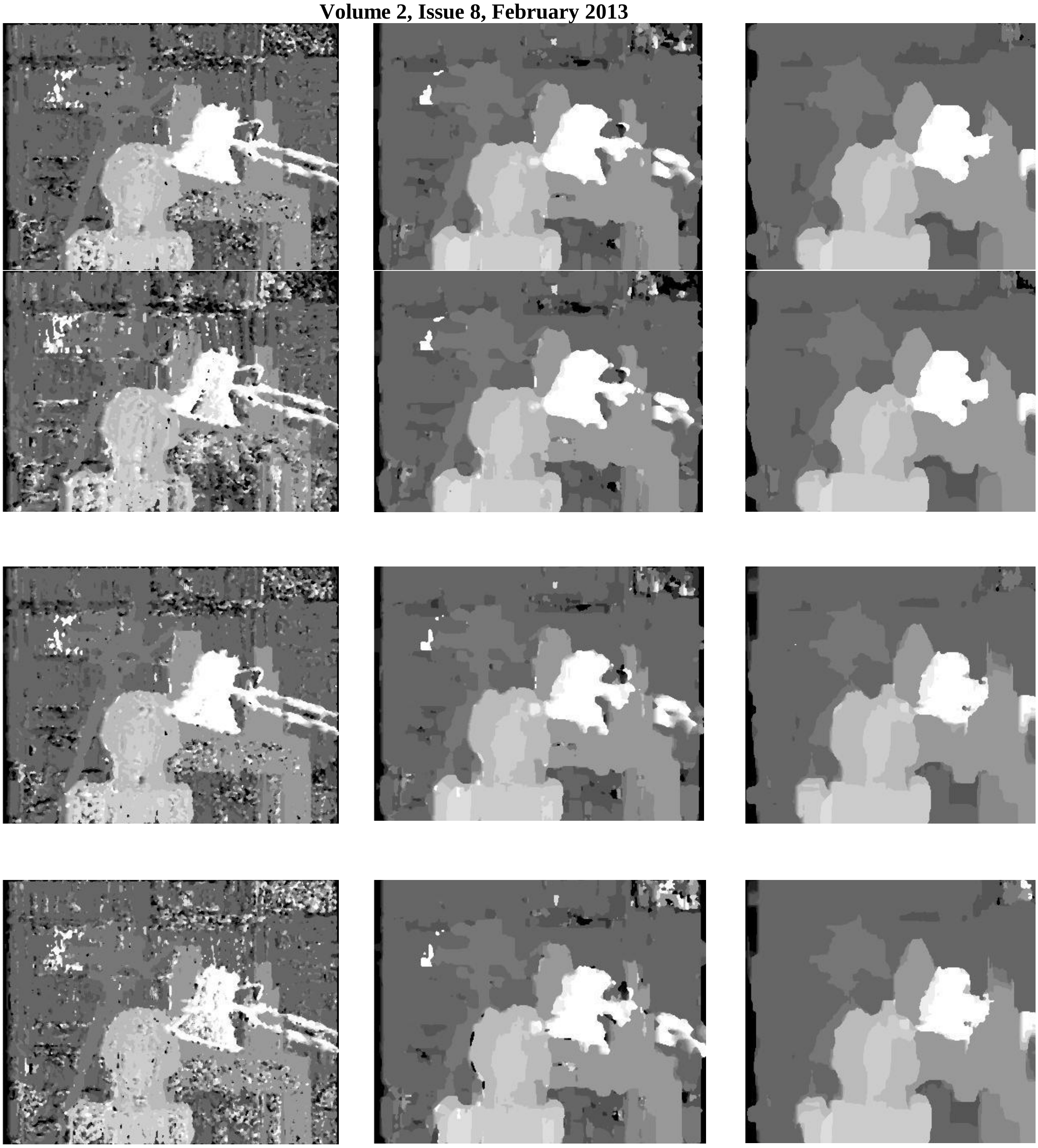


Fig 2: Tsukuba Image and Ground Truth [1]

B. Disparity Maps using different window sizes

Window size = 3 **Window size = 11** **Window size = 25**

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 2, Issue 8, February 2013



SAD

ZSA
D

SSD

NCC

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 2, Issue 8, February 2013

	Window size		
Algorithm	3	11	25
SAD	0.13817	0.14264	0.16032
SSD	0.13654	0.14596	0.17474
ZSAD	0.20162	0.27751	0.31131
NCC	0.21545	0.27938	0.33927
ZNCC	0.35073	0.48174	0.63432

Table I: Computational time in seconds

ZNCC

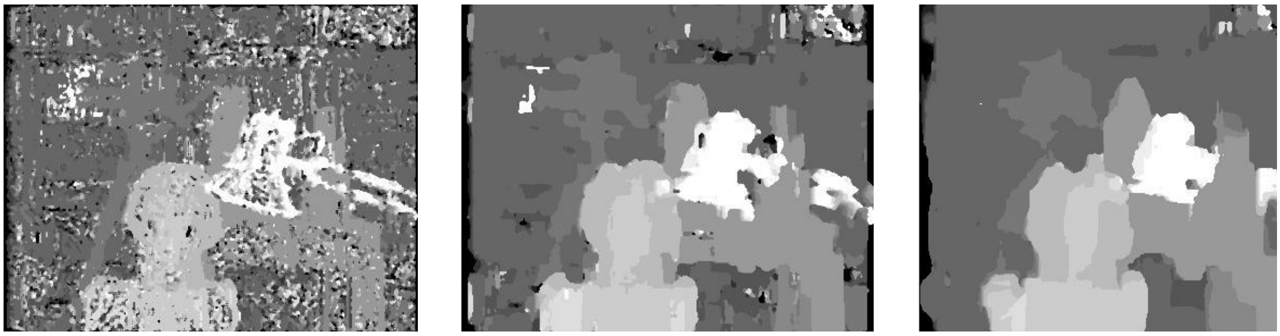


Fig 3: Comparison of Disparity Maps

C. Computational Time of the results

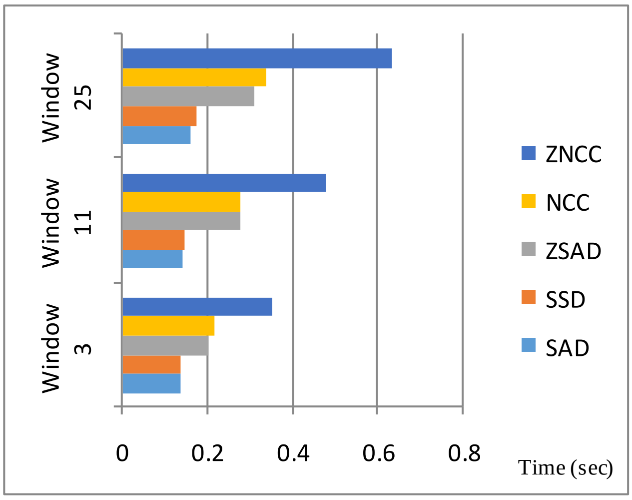


Fig 4: Time Comparison

V. CONCLUSION

It is clear from the result that NCC and ZNCC perform better with lower window size whereas the other method (SAD, SSD, ZSAD, and ZSSD) perform well with larger window size. NCC and ZNCC have higher accuracy as correlation is used but has limitation of high demand on computation time. As shown in the table1 time required for NCC and ZNCC is compensated by its improved performance over other methods. SSD and SSD would be the ideal choice for implementation on FPGA.

VI. FUTURE SCOPE

This paper only deals with the methods of aggregation which forms the backbone of any stereo vision algorithm. But along with method of aggregations the steps of the cost computation and disparity selection are equally important. These steps can also be improvised to make the stereo vision system more efficient. Also, occlusion handling techniques can be developed to improve the results. By improvising the stereovision system we can implement it on FPGA or SOC for real time applications such Face Recognition [7], 3D Face Reconstruction and 3D Terrain Mapping.

ACKNOWLEDGMENT

We wish to acknowledge Dr. A. N. Cheeran for her throughout support and guidance in every step from conceptualization to implementation of algorithms aiding in successful completion of this review paper.

REFERENCES

[1] Middlebury standard dataset for evaluation, tsukuba image Available: http://vision.middlebury.edu.

[2] Boguslaw Cyganec, J. Paul Siebert, "An Introduction to 3D Computer Vision Techniques and Algorithms " , John Wiley and sons, 2009.

[3] Yehu Shen, "Efficient normalized cross correlation calculation method for stereo vision based robot navigation",Frontiers of Computer Science in China, June 2011, Volume 5, Issue 2, pp 227-235.

[4] Luigi Di Stefano, Stefano Mattoccia, Federico Tombari, Article on "ZNCC-based template matching using bounded partial correlation".

[5] Banks, Jasmine, Porter, Reid, Bennamoun, Mohammed, & Corke, Peter (1997) "A generic implementation framework for stereo matching algorithms", In Chaplin, Bob I. & Page, Wyatt H. (Eds.) DICTA'97 and IVCNZ'97 Conference Proceedings, The Department of Production Technology, Massey University, Auckland, New Zealand, pp. 29-34.

[6] Nalpantidis Lazaros, Georgios Christou Sirakoulis and Antonios Gasteratos, "REVIEW OF STEREO VISION ALGORITHMS: FROM SOFTWARE TO HARDWARE", International Journal of Optomechatronics, 2: 435–462, 2008.

[7] Daniel Bardsley, "Stereo Vision for 3D Face Recognition", University of Nottingham, August 2005.

[8] Steffano Mattoccia, "Stereo-Vision: Algorithms and Applications", University of Bologna, May 22, 2012. Available : www.vision.deis.unibo.it/smatt.