VOL. 13, NO. 11, JUNE 2018 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

PERFORMANCE ANALYSIS OF COMPUTING TECHNIQUES FOR IMAGE DISPARITY IN STEREO IMAGE

Rakesh Y.
Department of Electronics and Communication Engineering, SRKIT, Vijayawada, India
E-Mail: rakesh.yemineni@gmail.com

ABSTRACT

Several computational techniques are proved to be efficient in computing the disparity of the stereo images. These are responsible for the change detection in the several applications. On the line, it is observed that a thorough analysis in terms of its performance is necessary for such stereo images. In this paper, conventional computing techniques like Normalised Cross Correlation (NCC), Sum of Absolute Differences (SAD) and Sum of Squared Differences (SSD) are considered for analysis. The standard image of Teddy is used for the analysis and the computation is carried out in MATLAB. Implementation and computation of the image disparity using these three computing techniques are performed and analyzed.

Keywords: image disparity, saliency, NCC, SAD, SSD.

1. INTRODUCTION

The manifold nature of visual attention leads psychologists, neurobiologists and computer scientists to understand simulate and implement its functionalities. The computer scientists, including Koch & Ullman [1], Itti *et al* [2], Frintrop *et al* [3] have been investigating and attempting to build biologically plausible visual attention model over the last three decades. As discussed in [4], the basic difference between computational models concerns the underlying structure which is either based on neural networks (Connectionist models) or on a collection of gray-scale, Saliency Maps (SM) (Filter models). The connectionist models designed on the assumption that each unit/pixel/neuron has different behavior whereas the filter models treat each pixel equally. The connectionist models claim to be more biologically plausible than the filter models but in practice, treating each unit differently is usually too costly. The filter model is well suited for the applications to real-world images and they can profit from approved techniques in computer vision. These models construct an SM, which is a two-dimensional gray scale image. The saliency map is used to represent the visual saliency of the corresponding location, independently of the particular feature that makes the location salient. The filter models compute saliency map using either bottom-up approaches or top-down approaches. The bottom-up models (also called as exogenous) are guided by low-level image properties which are solely derived from the images. However, the top-down models (also known as endogenous) are driven by cognitive factors such as knowledge, expectations, and current goals.

Several potential applications involving visual attention mechanism can be envisaged. According to Filipe & Alexandre [5] and Borj & Itti [6], the applications of CVA have been organized into three categories, namely vision and graphics, robotics and those in other areas.

It is noticeable that the significance of the content varies with length and breadth of the image and including time in the video. An information is contained in an image. From visual perception the information is perceived unevenly. The light rays which are incident on the centre of the retina have generally very high resolution and decreases down low areas the circumference. In several multimedia application, automatic detection of certain vital regions in the image or video. During the past two decades, there are several works which have presented a variety of techniques for saliency detection addresses several new problems. It is demonstrated in [7] on how to intelligently utilizing insignificant region of an image or a web page to include or display an advertisement, thereby avoiding interference with potentially useful data. Saliency mapping is one useful technique which can be used for this purpose. Same technique for image retrieval [8], in which the images with significant similarity are retrieved with respect to query or represent the image. In some cases, the saliency maps are used for thumb nail generation [9]. With saliency detection, it is reported that the object recognition, tracking, segmenting and classification is possible [10]. Even the image compression can also be benefitted with saliency [11]. In brief, the saliency detection schemes are broadly classified as the model based and task based. Similarly, the popular approaches are bottom up and top down approaches. The bottom up approaches often use the low-level features of the image in terms of color, resolution, brightness etc. however, the top down approach employs high-level features like objects. It is also possible to fuse both the approaches to make the approach more efficient. Saliency mapping (SM) is a technique through which visually attracting regions of an image or video are modeled. This computationally efficient technique elevates the significance of a particular region when compared with the entire image. This is constituted by relevant pixels in the mapping. The typical procedure involves in identifying significant pixels isolated within the image. This refers to the distinctness of the statistical features of the attracting region. The existing techniques purely rely on the natural human observation strategy. This is possible with the spot light or dense SM. The statistical details extracted from the region of the image are sufficient to exactly predict the most salient part of the image. However, in the current scenario, a high degree of digital vision application is

required. An application involving in separating the background, image segmentation and recognition. These applications often require very high and consistent resolution of images. The SM in high-resolution images often tries to concentrate the region which could be of the high degree of attractiveness to the observer. This can be reused as a property of SM which lead to excellent automated vision capture system.

In practice, there are some characteristics for sustained SM in the robotic application. It is essential to ensure the prominent location information of objects concerned on salient such that they are capable of defining specific boundaries. Similarly, it should have a uniform distribution of highlighting context on the entire SM object irrespective of their boundary. In addition, it should be capable enough to separate the unnecessary background and extract significant foreground information.

2. DISPARITY BASED SALIENCY DETECTION AND COMPUTING TECHNIQUES

The saliency map is a crucial outcome of many computational visual attention models. Each pixel in this map represents the relevance or importance of it with respect to the whole image. This map allows identifying the most distinct pixels in an image independently of the particular features that they possess. The existing visual attention models generate either spotlight or dense saliency map. The details captured in the spotlight saliency maps are sufficient to predict where human observers would fixate. However, more demanding higher-level computer vision applications such as figure background separation, object segmentation, object recognition and image understanding will require a high-resolution map. The dense and full resolution saliency maps capture image regions which the observer thinks is most interesting in the image. Because of this property, the dense saliency maps can be fed to high-level computer vision tasks. The essential characteristics of a saliency map which can be used in higher-level computer vision and robotic applications are to

- provide precise location of salient objects or parts of salient objects in the input image produce well-defined boundaries of important and relevant regions in an image
- uniformly highlight entire parts of salient objects rather than their boundaries
- efficiently suppress the irrelevant background and separate out important foreground regions
- produce dense and full resolution (same as that of the input image) saliency maps.

The saliency detection mechanism ideally uses the minimal set of tunable parameters and uses no prior knowledge and training corpus. Also, it uses the minimal set of feature maps which is computationally efficient. The aforementioned properties are set as goals of the proposed saliency detection mechanism.

Three computing techniques are used in this work. They are explained as follows in this section.

2.1 Sum of Squared Differences (SSD)

The computation of SSD is more complexed than SAD. It involves several additional steps. Initially, both the images are taken and transformed into one-dimensional vectors fN and gN. This is given as

$$SSD(i,j)=\left(\sum_{k=i}^{m+i-1} f(k,j)-\sum_{i=1}^{m} g(i,1)\right)^2+\left(\sum_{k=i}^{m+i-1} f(k,j+1)-\sum_{i=1}^{m} g(i,2)\right)^2 \mathrm{L} +\left(\sum_{k=i}^{m+i-1} f(k,n+j-1)-\sum_{i=1}^{m} g(i,n)\right)^2 \quad (1)$$

2.2 Sum of Absolute Differences (SAD)

The computation of NCC passes through several operations depending upon the size of ‘f' and ‘g'. As an alternative, the SAD method considers the sum of absolute differences between the image ‘g' and ‘f'. this is given as

$$SAD(i,j)=\sum_{x=1}^{n}\sum_{y=1}^{n}\left|f(i+x,j+y)-g(x,y)\right| \quad (2)$$

2.3 Normalized Cross Correlation (NCC)

Computing NCC [12] is a useful tool for comparing two images, searching for a pattern in an image, searching for a pattern in an image, registering an image, matching images etc.

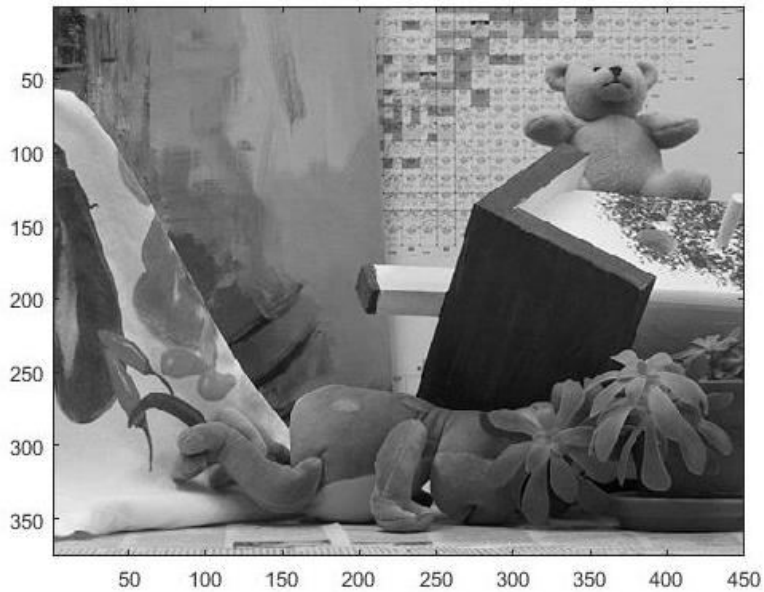
$$\lambda_{NCC}=\frac{\sum_{x=0,y=0}^{(n-1),(m-1)}\left(f(i+x,j+y)-\bar{f}(i,j)\right)\left(g(x,y-\bar{g}\right)}{\sqrt{\sum_{x=0,y=0}^{(n-1),(m-1)}\left(f(i+x,j+y)-\bar{f}(i,j)\right)^2\cdot\sum_{x=0,y=0}^{(n-1),(m-1)}\left(g(x,y-\bar{g}\right)}} \quad (3)$$

Where f is the first image and g is the second image. $\bar{f}$ and $\bar{g}$ are the mean values of images 'f' and 'g'.

3. RESULTS AND DISCUSSIONS

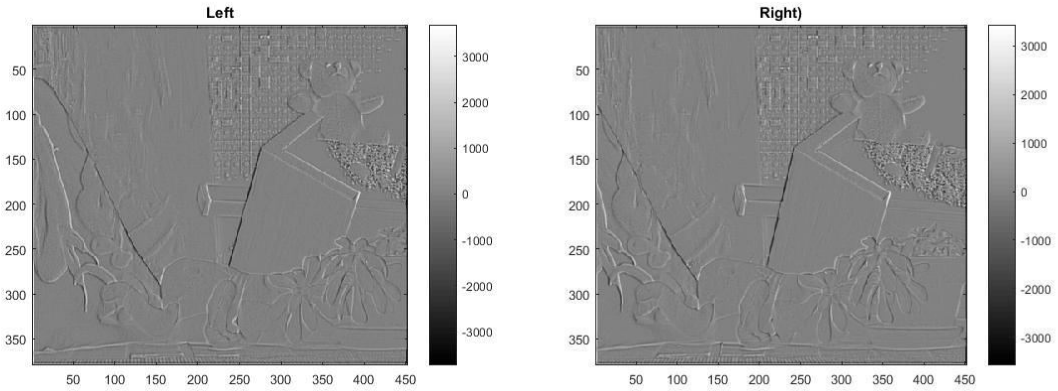
Results pertaining to the three techniques for investigation on the disparity between stereotype images are presented in this section. The two input images namely Teddy Left and Right are considered for the simulation based experimentation. These two colour images are as shown in Figure-1 (a) and 1(b). Initially, these images are converted into gray scale. The respective gray scale image is as shown in Figure-1(c). The simulation based experimentation is carried out using Matlab 7 on an i3 Processor with 4 GB RAM. Initially the SAD is applied on the Stereo images and the respective. Left to right disparity and the right to left disparity are as shown in Figure-2 (a) and (b) respectively. The corresponding distance Map and Disparity map amplitude are given in Figure-3.

(a) (b)



(c)

Figure-1. Image (a) Left (b) Right and (c) Grey scale.



(a) (b)

Figure-2. (a) Left to Right SAD and (b) Right to Left SAD.

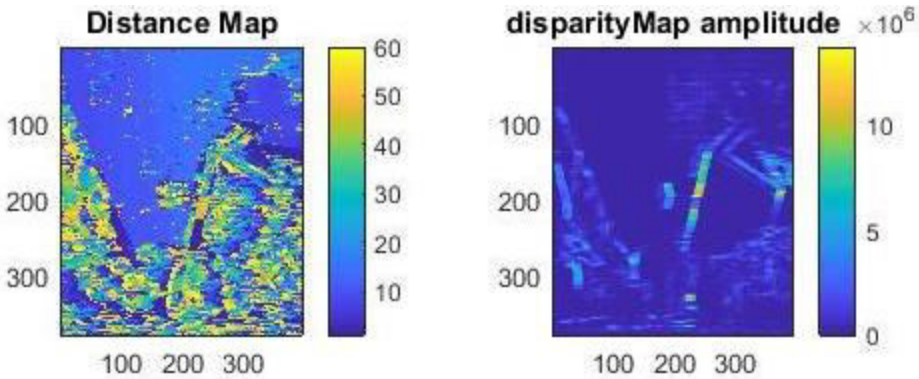


Figure-3. Disparity map distribution.

Similarly, the SSD based disparity study is performed using the procedure described in the flow chart in Figure-4. The corresponding SSD based Output image is as shown in Figure-5. Similarly, the NCC based Disparity is computed as discussed earlier and the corresponding disparity is evident from the Figure-6.

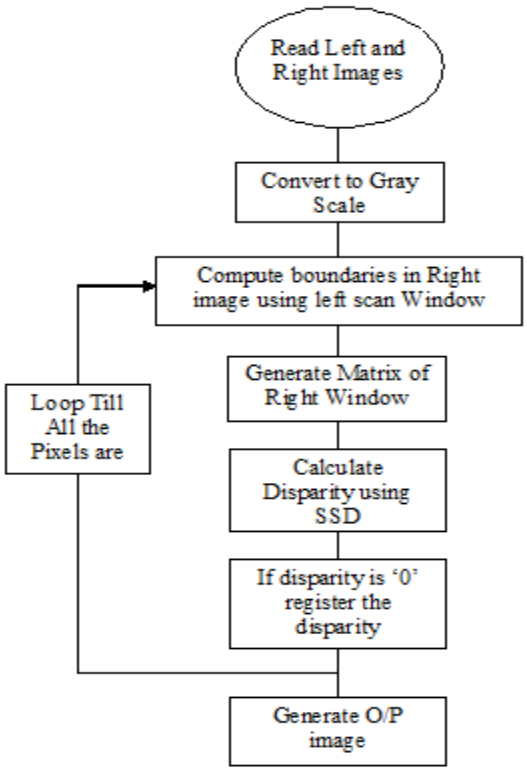


Figure-4. Flow chart describing SSD.

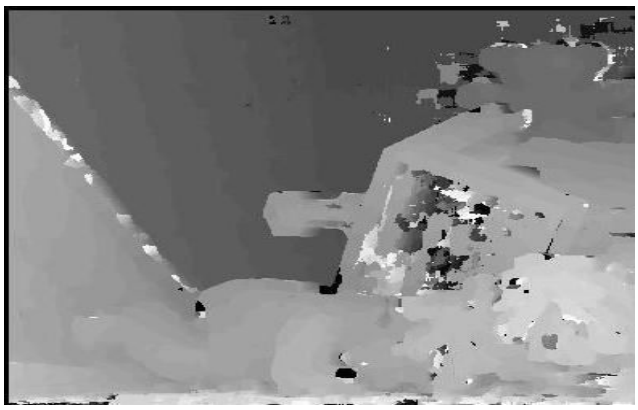


Figure-5. SSD based disparity.

Following the SSD based disparity analysis of the input images; the corresponding NCC based disparity analysis is performed as mentioned in the computing Section earlier. The respective output images depicting the NCC based disparity are obtained accordingly and are given as shown in Figure-6.

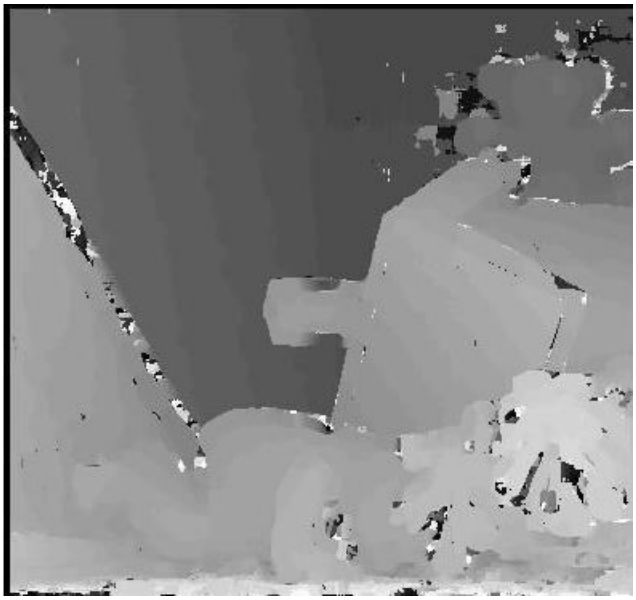


Figure-6. Normalized cross correlation.

4. CONCLUSIONS

Computation of disparity in the stereo images is successfully performed using three numerical techniques known as NCC, SSD and SAD. Though the NCC is simple, the computational process goes through conventional complexed calculations consuming computation time. Following NCC the SAD and SSD are performed. The visual representation of the disparity images clearly mentions the fact that the SSD has the obvious advantage of the squaring the differences rather than taking the absolute differences. This is evident from the output results of SSD and SAD. Extending this to 3D and video based applications would be a good scope of future work.

REFERENCES

[1] C. Koch and S. Ullman. 1985. Shifts in selective visual attention: towards the underlying neural circuitry. Hum Neurobiol. 4(4): 219-227.

[2] L. Itti, C. Koch and E. Niebur. 1998. A model of saliency-based visual attention for rapid scene analysis. Pattern Analysis and Machine Intelligence, IEEE Transactions on. 20(11): 1254-1259.

[3] S. Frintrop, E. Rome, A. Nchter and H. Surmann. 2005. A bimodal laser based attention system. Computer Vision and Image Understanding. 100: 124151.

[4] C. Guo and L. Zhang. 2010. A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression. Image Processing, IEEE Transactions on. 19(1): 185-198.

[5] Filipe S. and Alexandre L. A. 2013. A Comparative Evaluation of 3D Keypoint Detectors. In 9th Conference on Telecommunications, Conftele 2013, pp. 145-148, Castelo Branco, Portugal.

[6] A. Borji and L. Itti. 2013. State-of-the-Art in Visual Attention Modeling. in IEEE Transactions on Pattern Analysis and Machine Intelligence. 35(1): 185-207. doi: 10.1109/TPAMI.2012.89.

[7] H. Liu, X. Qiu, Q. Huang, S. Jiang and C. Xu. 2009. Advertise gently-in image advertising with low intrusiveness. in Image Processing (ICIP), 2009 16th IEEE International Conference on. IEEE. pp. 3105-3108.

[8] K. Gao, S. Lin, Y. Zhang, S. Tang and H. Ren. 2008. Attention model based sift keypoints _ltration for image retrieval. in Computer and Information Science, 2008. ICIS 08. Seventh IEEE/ACIS International Conference on. IEEE. pp. 191-196.

[9] L. Marchesotti, C. Cifarelli, and G. Csurka. 2009. A framework for visual saliency detection with applications to image thumbnailing. in Com puter Vision, 2009 IEEE 12th International Conference on. IEEE. pp. 2232-2239.

[10] U. Rutishauser, D. Walther, C. Koch, and P. Perona. 2004. Is bottom up attention useful for object recognition? in Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the

2004 IEEE Computer Society Conference on, vol. 2. IEEE. pp. II-37.

[11] N. Dhavale and L. Itti. 2003. Saliency-based multifoveated mpeg compression. in Signal processing and its applications, 2003. Proceedings. Seventh international symposium on. Vol. 1. IEEE, 2003, pp. 229-232.

[12] E. Z. Psarakis, G. D. Evangelidis. 2005. An enhanced correlation-based method for stereo correspondence with subpixel accuracy. Int. Conf. on Computer Vision (ICCV).