PhD Forum: Towards an Embedded Stereo Matching Algorithm Based on Multiple Correlation Windows

M. A. Palacios-Ramos, H. D. Vázquez-Delgado, Abiel Aguilar-González, M. Pérez-Patricio
Instituto Tecnológico de Tuxtla Gutiérrez (ITTG), Carretera Panamericana Km. 1080, C.P. 29050, Apartado Postal: 599, Tuxtla Gutiérrez, Chiapas, Mexico
Contact:mperez@ittg.edu.mx

ABSTRACT

Stereo matching consists in extracting 3D information from digital images, such as those obtained by a CCD camera. It is an important issue under several real world applications, such as positioning systems for mobile robots, augmented reality systems, etc. In previous works one of the most popular trend to address the stereo matching challenge is that compares scene information from two viewpoints (left-right) with an eppipolar geometry via correlation metrics. In regard to the correlation metrics, most previous works compute the similarity between pixels in the left image and pixels in the right image using a correlation index computed on neighborhoods of these pixels called correlation windows. Unfortunately, in order to preserve edges, small correlation windows need to be used, while, for homogeneous areas, large correlation windows are required. To address this problem, we lay down on the hypothesis that small correlation windows combined with large correlation windows should deliver accurate results under homogeneous areas while at the same time edges are preserved. To validate our hypothesis, in this paper a similarity criterion based on the grayscale homogeneity of the correlation window being processed is presented. Preliminary results are encourageous, validates our hypothesis and demonstrated the viability performance and scope of the proposed approach.

CCS CONCEPTS

• **Computer systems organization** → **Real-time system specification**; • **Software and its engineering** → **Real-time systems software**.

KEYWORDS

Stereo Matching, Disparity map, Local correlation

ACM Reference Format:

M. A. Palacios-Ramos, H. D. Vázquez-Delgado, Abiel Aguilar-González, M. Pérez-Patricio. 2019. PhD Forum: Towards an Embedded Stereo Matching Algorithm Based on Multiple Correlation Windows. In *13th International Conference on Distributed Smart Cameras (ICDSC 2019), September 9–11, 2019, Trento, Italy.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3349801.3357128

ICDSC 2019, September 9–11, 2019, Trento, Italy

ACM ISBN 978-1-4503-7189-6/19/09…$15.00
https://doi.org/10.1145/3349801.3357128

1 INTRODUCTION

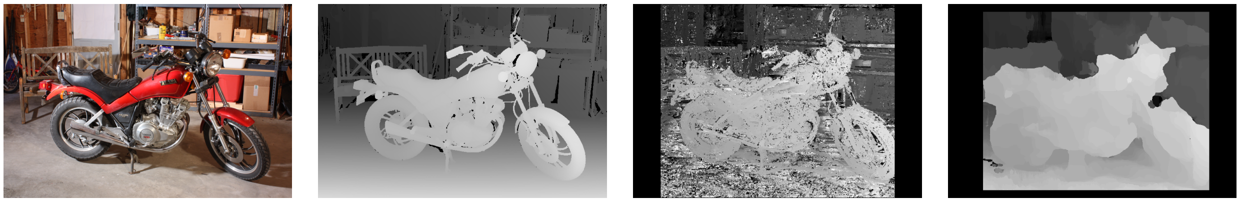
In traditional stereo vision, to compute pixel correspondences (disparity), most previous works compare a rectangular neighborhood (correlation window) centered on a reference pixel in one of the images from the stereo pair with similar neighborhoods for some pixels in the same raster line of the other image [4]. In order to compare correlation windows, previous works [1–3, 5] often use local correlation metrics such as the Sum of Absolute Differences (SAD)**Eq.**1: where $I_l(x+i, y+j)$ and $I_r(x+i+s, y+j)$ are the grey scale values of pixels in the left/right images from the stereo pair, respectively. $2 \times w + 1$ is the correlation window size, s is the shift of the correlation window and the maximal shift of the correlation window in the search image is s_m. A correlation coefficient is estimated for each shift and the shift that minimizes the correlation coefficient is retained as the disparity value at the pixel being processed.

$$C(x,y,s) = \sum_{i=-w}^{i=w} \sum_{j=-w}^{j=w} |I_l(x+i, y+j) - I_r(x+s+i, y+j)|, \tag{1}$$

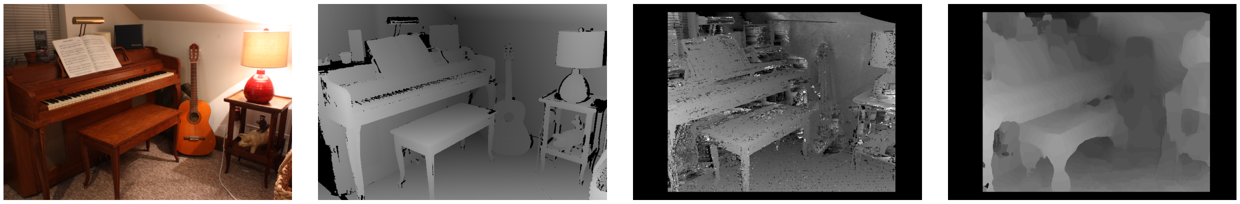
2 THE PROPOSED APPROACH

In **Fig.** 1 disparity maps by applying the SAD correlation metric are shown. Although there are several problems with the original SAD formulation, one of the most important issue is to select the correlation window size. High window size values allow to determine the correct correlation values in areas with uniform texture. However, these window sizes imply a high computational demand and erroneous values at certain points due to the blurring edges and that small features are eliminated, as illustrated in **Fig.** 1.(d),(h). On the other hand, small window sizes imply low computational demand but the correlation coefficient measurement is sensitive to noise, therefore, erroneous values at uniform texture regions are generated as seen in **Fig.** 1.(c),(g). To address this issue, we lay down on the hypothesis that small correlation windows combined with large correlation windows should deliver accurate results under homogeneous areas while at the same time edges are preserved. As early formulation, we propose a similarity criterion based on the grayscale homogeneity of the correlation window being processed. Let $k_1(x,y)$, $k_2(x,y)$ be the grayscale homogeneity of a small, large correlation windows, respectively. $k_n(x,y)$ is computed as shown in **Eq.** 2; where d_n is the disparity map of the stereo pair and $2 \times w + 1$ is the correlation window size.

$$k_n(x,y) = \frac{1}{((2 \times w)+1)^2} \sum_{u=-w, v=-w}^{u=w, v=w} (d_n(x+u, y+v) - d_n(x,y)) \tag{2}$$



(a) Scene: Motorcycle (b) Ground truth (c) SAD $w = 1$ (d) SAD $w = 20$



(e) Scene: Piano (f) Ground truth (g) SAD $w = 1$ (h)SAD $w = 20$

Figure 1: The new Middlebury Stereo Evaluation dataset [6]. Disparity maps by applying the original SAD formulation. In order to preserve edges, small correlation windows need to be used, see (c) and (g). For homogeneous areas, only large correlation windows deliver accurate results, as illustrated in (d) and (h).

Let $d_1(x, y), d_2(x, y)$ be the computed disparity values for a small or large correlation windows, respectively. $d_n(x, y)$ is computed as: $d_n(x, y) = \min_s(C_n(x, y, s))$; where $C_n(x, y, s)$ is the correlation function for a $n = 1$ (small), $n = 2$ (large) correlation windows, respectively (computed using the original SAD formulation, please see **Section** 1). Then, the final disparity value D is computed by comparing the homogeneity values, as illustrated in **Eq.** 3 and 4; where $d_1(x, y), d_2(x, y)$ are disparity values for a $n = 1$ (small), $n = 2$ (large) correlation windows, respectively; and β, Δ are threshold values defined by the user.

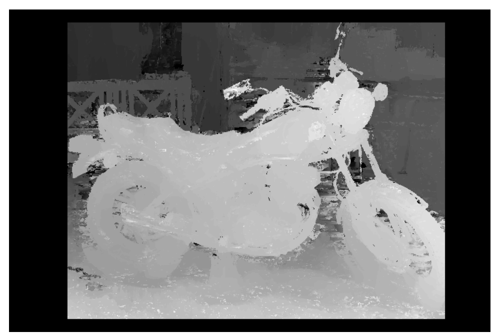
$$K(x,y) = \begin{cases} d_2(x,y), & k_2(x,y) \le \Delta \\ d_1(x,y), & otherwise \end{cases} \tag{3}$$

$$D(x,y) = \begin{cases} d_1(x,y), & k_1(x,y) \le \beta \\ K(x,y), & k_1(x,y) > \beta \end{cases} \tag{4}$$

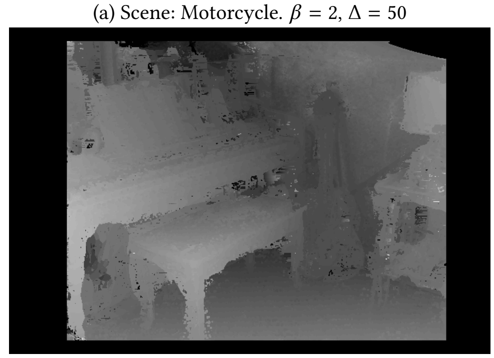
3 PRELIMINARY RESULTS

In order to validate our hypothesis, we have implemented our algorithmic formulation in a MatLab script. Then, we have estimated dense disparity maps for the Motorcycle and Piano scenes of the new Middlebury Stereo Evaluation dataset [6], see **Fig.** 2. In all cases $w_1 = 1$, $w_2 = 20$, and $\beta = 2$, $\Delta = 50$ for $n = 1$ (small), $\beta = 4$, $\Delta = 80$ for $n = 2$ (large) were used. Experimental results validate our hypothesis, demonstrating that it is possible to improve the original SAD formulation. Compared with the original formulation which reaches around 62% of accuracy for the Motorcycle and Piano scenes respectively; our algorithmic formulation improves it, reaching 67%, 70% of accuracy for the same scenes. In all cases, the accuracy was measured by applying **Eq.** 5; where I_1 is the ground truth. I_2 is the estimated disparity map. x is the horizontal resolution of the input image. y is the vertical resolution of the input image and N is set as $x \times y$.

$$\sigma = \frac{100\%}{N} \sum_{i=1,j=1}^{i=x,j=y} \begin{cases} 1 & if(I_1(i,j) - I_2(i,j)) >= 1 \\ 0 & otherwise \end{cases} \tag{5}$$



(a) Scene: Motorcycle. $\beta = 2$, $\Delta = 50$



(b) Scene: Piano. $\beta = 4$, $\Delta = 80$

Figure 2: Disparity maps by applying the proposed approach. Experimental results demonstrate the performance and scope of the proposed approach (qualitatively and quantitatively more accurate than the original SAD formulation).

4 CONCLUSIONS AND FUTURE WORK

In this paper, we have introduced a new stereo matching approach in which multiple correlation windows were used in order to improve the SAD correlation metric. Experimental results validated our hypothesis, outperforming in terms of accuracy the original SAD formulation. As work in progress, we are investigating about more robust similarity criteria, based on fuzzy logic and learning techniques. In other issue, to guarantee embedded capabilities we will implement our algorithmic formulation in an FPGA architecture. These new criteria, more detailed results, a detailled comparison with previous works and the FPGA architecture will be presented in a future forum.

REFERENCES

[1] Chai, Y., and Cao, X. Stereo matching algorithm based on joint matching cost and adaptive window. In *2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)* (2018), IEEE, pp. 442–446.

[2] Chang, Y.-J., and Ho, Y.-S. Fast cost computation using binary information for illumination invariant stereo matching. *IEEE Seoul Section Student Paper Contest, Paper C 4* (2016).

[3] Hallek, M., Boukamcha, H., Smach, F., and Atri, M. Real time stereo matching using two step zero-mean sad and dynamic programing. In *2018 15th International Multi-Conference on Systems, Signals & Devices (SSD)* (2018), IEEE, pp. 1234–1240.

[4] Hartley, R., and Zisserman, A. *Multiple view geometry in computer vision.* Cambridge university press, 2003.

[5] Pérez-Patricio, M., Aguilar-González, A., Arias-Estrada, M., Hernández-de León, H.-R., Camas-Anzueto, J.-L., and de Jesús Osuna-Coutiño, J. An fpga stereo matching unit based on fuzzy logic. *Microprocessors and Microsystems 42* (2016), 87–99.

[6] Scharstein, D., Hirschmüller, H., Kitajima, Y., Krathwohl, G., Nešić, N., Wang, X., and Westling, P. High-resolution stereo datasets with subpixel-accurate ground truth. In *German conference on pattern recognition* (2014), Springer, pp. 31–42.