# Stereo matching based on Census transformation of image gradients

C. Stentoumis[a*], L. Grammatikopoulos[b], I. Kalisperakis[b], G. Karras[a], E. Petsa[b]
[a]Lab. of Photogrammetry, Dpt. of Surveying, National Technical University of Athens, GR-15780, Greece; [b]Lab. of Photogrammetry, Dpt. of Surveying, Technological Educational Institute of Athens, GR-12210, Greece

## ABSTRACT

Although multiple-view matching provides certain significant advantages regarding accuracy, occlusion handling and radiometric fidelity, stereo-matching remains indispensable for a variety of applications; these involve cases when image acquisition requires fixed geometry and limited number of images or speed. Such instances include robotics, autonomous navigation, reconstruction from a limited number of aerial/satellite images, industrial inspection and augmented reality through smart-phones. As a consequence, stereo-matching is a continuously evolving research field with growing variety of applicable scenarios. In this work a novel multi-purpose cost for stereo-matching is proposed, based on census transformation on image gradients and evaluated within a local matching scheme. It is demonstrated that when the census transformation is applied on gradients the invariance of the cost function to changes in illumination (non-linear) is significantly strengthened. The calculated cost values are aggregated through adaptive support regions, based both on cross-skeletons and basic rectangular windows. The matching algorithm is tuned for the parameters in each case. The described matching cost has been evaluated on the Middlebury stereo-vision 2006 datasets, which include changes in illumination and exposure. The tests verify that the census transformation on image gradients indeed results in a more robust cost function, regardless of aggregation strategy.

**Keywords:** 3D reconstruction, matching, stereo, census, gradients, cost

## 1. INTRODUCTION

Generation of dense 3D information is a fundamental task in most applications in photogrammetry and computer vision (3D reconstruction, DSM production, novel view synthesis, automatic navigation, object detection). Methods for acquiring 3D information can be distinguished as *passive* and *active*. Image-based approaches (*passive*) are lately proven to be competitive to laser and optical scanners (*active*) in terms of accuracy, while exhibiting a clear advantage as regards cost and flexibility. Several theoretical alternatives exist for exploiting images in producing 3D information (e.g. shape from shading). A fundamental procedure is *image matching*, i.e. automatic determination of pixel correspondences among images. This approach may be seen as consisting of two steps: establishment of *sparse* correspondences for camera calibration/orientation; and *dense* matching for 3D surface reconstruction. *Stereo-matching* algorithms, as that presented here, usually exploit the *epipolar constraint*; hence they typically operate on rectified images to produce a *disparity map*.

Existing and emerging application fields have guided the research interest towards a significant number of proposed stereo-matching algorithms. The effectiveness of such algorithms has been extensively discussed in a number of surveys[1–4]. Scharstein and Szeliski [3] have categorized algorithms by splitting them into four fundamental components: *matching cost computation*, *support aggregation*, *disparity optimization* and *disparity refinement*. In [5,6] the question of support region formation is discussed, while in [7] the cost function itself is evaluated under different optimization algorithms. Enlightening comments may also be found in [1,4]. Publications [8,9] provide respective surveys focusing on criteria for hardware implementation and real-time performance. The assessment of stereo-methods poses difficulties due to diverging criteria set by different applications; real-time performance, manipulation of large amounts of data, accuracy and computationally manageable complexity are such possible requirements. Moreover, issues such as scene occlusions, surface non-continuity, differing convergence angles, illumination changes, camera geometry or type of imagery (aerial / close-range) inevitably differentiate the treatment of the matching problem. These aspects often depend on the recorded scenes and the underlying application. Two very commonly used online platforms for benchmarking stereo-matching algorithms are that of Middlebury College[*] and the Karlsruhe Institute of Technology (KITTI)[†] dataset for autonomous driving[10,11].

[*] http://vision.middlebury.edu/stereo/
[†] http://www.cvlibs.net/datasets/kitti/eval_stereo_flow.php?benchmark=stereo

*cstent@mail.ntua.gr; phone +302107722683; fax +302117708184

In *matching cost computation* a dissimilarity measure is attributed to each pixel for every value in the disparity range. A wide spectrum of such matching measures has been proposed. Most common measures for similarity between pixels are their absolute or squared difference, their normalized cross correlation or measures relying on input images transformed by filters such as the median, the mean, the LoG, or more sophisticated tools like bilateral filtering[12]. Non-parametric image transformations, such as *rank* and *census*[13], produce robust results based on binary relationships of pixels with their neighbourhood. *Birchfield-Tomasi* dissimilarity measure[12] copes with differences in image sampling, and, recently, the *mutual information* cost[14] has been proposed for effectively handling radiometric differences. On the other hand, pixel-wise descriptor measures, like DAISY[15] and SIFT variations[16], have yielded promising results in global formulations for wide-based stereo.

During the last decade, the census transformation (CT) has become increasingly popular as the core for dense matching functions; additionally, modifications have been proposed for optimizing its performance. The original matching cost has been evaluated on different colour spaces for a typical rectangular-window-based cost aggregation[17]. The authors found no considerable improvement of disparity maps for colour spaces other than RGB. The modified census transformation (MCT) approach[18] compares each pixel intensity against the average intensity in a neighbourhood instead of the central pixel intensity; following this, [19] suggested the application of MCT on bidirectional gradient images using Sobel filter for optical flow. The two latter works also propose the computation of CT on a sparse neighbourhood for speeding up the results with small loss in quality. Another modification of census is cross comparison census (CCC)[20], which takes into consideration not the binary relations among a pixel and its neighbours, but the relations among each pixel in the defined neighbourhood and its four adjacent pixels defined in a clockwise direction.

Matching cost on census transformation, as well as some of the above costs, has been combined with several other costs. In [21] census cost ($C_{\mathrm{CI}}$) was combined with absolute difference on colour ($C_{\mathrm{ADC}}$) and produced top-scoring results in the Middlebury benchmark. In [22] the original census transformation was extended through its application on image gradients (CG); the total cost was found by combining costs $C_{\mathrm{CG}}$, $C_{\mathrm{ADC}}$ with the absolute difference on gradients $C_{\mathrm{ADG}}$, and proved to be efficient for sub-pixel accuracy in Middlebury and a variety of different scenarios. More recently [23], the results of [22] have been improved in the vicinity of edges by the addition of a cost component based on Weber's law. Finally, a cost combination of $C_{\mathrm{CI}}$, or $C_{\mathrm{CCC}}$, with the differences of the mean pixel intensity absolute differences from its neighbours has been proposed in [20] for handling the radiometric differences of the KITTI stereo-dataset.

Cost computed per pixel is supported by a neighbourhood around pixels in the *cost aggregation* step. Three main approaches exist for addressing this issue: use of support weights, arbitrary window shapes, and variations of rectangular windows. For a review of aggregation methods for the purposes of real-time stereo matching see [5]. A more thorough evaluation of both performance and accuracy for the most important aggregation methods has been conducted in [6]. Disparity selection in local (region-based) methods is mostly performed in the winner-takes-all (WTA) mode, i.e. the disparity with the lowest aggregated cost is chosen. Global approaches, on the other hand, perform *disparity optimization* on an energy function defined over all image pixels by simultaneously imposing a smoothness constraint. Regarding this method, the most important techniques rely on partial differential equations[24] and graph-cuts[25]. Between local and global methods a class of algorithms have been derived from *semi-global* matching[14]. Finally, disparity maps from the selection/optimization step can be refined, and sub-pixel disparity values may be computed. A more thorough discussion on state-of-the-art research on stereo-matching is found in [22]. Moreover, a list of stereo-matching algorithms are cited in the aforementioned benchmarking sites (Middlebury and KITTI), where developments and new trends in the field are being reported.

This contribution presents a novel matching cost which shows robustness against illumination differences between the images of the stereo-pair. This cost is estimated on the separate census transformations of *x* and *y* image gradients and yields a census binary vector of doubled length. It is claimed here that when the census transformation is applied on gradients the invariance of the cost function to illumination changes is significantly strengthened. The estimated pixel-wise cost values are aggregated through adaptive support regions based on cross-skeletons, as well as through basic rectangular support windows. Both aggregation methods are thoroughly tuned in order to compare matching cost computed from census transformation on intensity against census on gradients for their most appropriate parameters. Hence, it is experimentally verified that the proposed census modification on image gradients indeed results in a more robust cost function, regardless of aggregation method. Although no discernible matching differences emerge for stereo pairs without radiometric differences, non-linear illumination changes lead to inferior disparity maps when obtained from the original census cost. The evaluation is performed on the complete Middlebury 2006 dataset[‡], since it includes images with different

[‡] http://vision.middlebury.edu/stereo/data/scenes2006/

illumination and exposure settings. Initial aspects of the present contribution have been presented in [22].

The rest of the publication is organized as follows: Section 2 describes the proposed matching cost; Section 3 refers to the two aggregation strategies used here to evaluate the proposed cost function based on the modified census; Section 4 presents the detailed results of the evaluation. Finally, the paper concludes with the remarks of Section 5.

## 2. CENSUS ON INTENSITY GRADIENTS

Census ($T_C$) is a well-known non-parametric image transformation[13]; each pixel **p** is mapped to a binary vector **I** of length m×n, which encodes the neighbouring pixels with intensities less than that of **p** in a neighbourhood $N_{m\times n}$. Thus, if pixel **q** is in the neighbourhood N of **p**, then:

$$c(p,q) = \begin{cases} 0, I(p) \leq I(q) \\ 1, I(p) > I(q) \end{cases} \quad (1)$$

$$T_c(p) = \bigotimes_{q \in N_p} c(p,q) \quad (2)$$

Census transformation $T_C$ depends on how a pixel relates to its surroundings within the image patch. It is hence robust against linear changes in brightness/contrast, namely radiometric distortions which do not modify the ordering of intensity values. In the present contribution the transformation is performed not on image intensity function $\boldsymbol{I}$ but on its gradients $\partial\boldsymbol{I}/\partial x$, $\partial\boldsymbol{I}/\partial y$. Image derivatives are related to characteristic structural image features (points, edges) and are, of course, widely used as a contributing source of information in various fields of image processing and computer vision. In image matching they are used e.g. in gradient-based methods in global optimization formulations, feature-based methods and local stereo [4,26,27]. The modified census transformation $T_{MC}$ provides the extended binary vector

$$c_i(p,q) = \begin{cases} 0, \partial\mathbf{I}/\partial i(p) \leq \partial\mathbf{I}/\partial i(q) \\ 1, \partial\mathbf{I}/\partial i(p) > \partial\mathbf{I}/\partial i(q) \end{cases} \quad (3)$$

$$T_{MC}(p) = \bigotimes_{i \in \{x,y\}} \bigotimes_{q \in N_P} c_i(p,q) \quad (4)$$

In Eq. (3) $i$ is the direction of the gradient; in Eq. (4) ⊗ denotes the act of concatenation, following the original definition of $T_C$. The direct introduction of image gradients in the two directions into the binary vector doubles the size of the produced vector $T_{MC}$, thus exploiting the representational potential of image gradients and their robustness against illumination changes.

In either transformation case, the matching cost $C'$ between a pixel **p** of the reference image and its corresponding pixel **p**' in the matching image is the Hamming distance, which represents the number of unequal elements in the two binary vectors:

$$C' = \sum_{x=1}^{n} \left( T_c^{ref}(p) \oplus T_c^{mat}(p') \right) \quad (5)$$

The final matching cost $C$ is smoothed via a robust exponential function, in order to diminish the influence of large, or outlying, pixel-wise costs:

$$C = 1 - \exp(-C'/\lambda) \quad (6)$$

The regularization parameter $\lambda$ was defined in all our experiments as $\lambda = l_T/3$, where $l_T$ is the length of the binary vector $T$, in order to softly constrain all cost values to a specific field of the exponential function.

## 3. AGGREGATION ALGORITHMS

Local algorithms of stereo-matching are based on the definition of a neighbourhood around each pixel. It is supposed that the pixels within this neighbourhood share the same disparity, thus fronto-parallel surfaces are favoured. In this work the matching cost derived from census transformation on image gradients is tested against the cost derived from the original census transformation. To this end, two aggregation strategies are applied: cross-based aggregation and typical rectangular windows.

Cross-based aggregation[28] is a highly adaptive method, which has been proven efficient regarding both accuracy and computational load. It relies on the assumption that pixels of a support region ought to have similar colours and are expected to decrease in coherence with their distance from the reference pixel. Here, a modification of the initial method is used[22], since it is proven to be more accurate through exploiting a linear threshold for the support region formation. The construction of such cross-based support regions is achieved by expanding around each pixel **p** a cross-shaped skeleton to create four segments, which define the corresponding sets of pixels H(**p**) and V(**p**) in the horizontal and vertical directions, as seen in Figure 1.

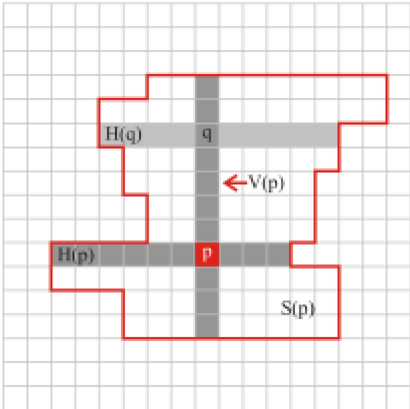



Figure 1. Expansion of the cross-based support region S(**p**) driven by the skeleton of each pixel. The skeleton pixels for V(**p**) and H(**p**) sets are calculated only once per pixel. Then, if a pixel **q** belongs to V(**p**), the corresponding horizontal arm H(**q**) is added to S(**p**). S(**p**) consists of the union of H(**q**) for all pixels **q** which participate in V(**p**). [After [28].]

The intersection of the two cross windows (S(**p**($x$,$y$))∩S(**p**($x$,$y$+$d$))) developed on the reference and the matching image is used, as it is more robust against local projective distortions and radiometric differences in the stereo-pair. As a consequence, the support region has different shape for each possible disparity value. Aggregation is applied on cost using the combined support region S. Aggregated pixel costs $C_{aggr}$ are normalized by the number of pixels in the support region:

$$C(p,d) = C_{aggr} / \|S(p,d)\| \tag{7}$$

The minimum length of all cross-segments is 1 pixel to ensure a minimum support region $S$ of 9 pixels, which is also the minimum window size for the tests under rectangular windows.

In this work, aggregation of pixel-wise costs $C_{pw}$ in typical rectangular windows of constant size $L_N$ in both dimensions has been tested, in order to further evaluate the performance of matching when applying census transformation on image gradients:

$$C(p,d) = \sum_{p \in N} C_{pw}(p,d) \tag{8}$$

The cost aggregation step of matching algorithms could become computationally expensive, especially when support regions vary for each pixel. Thus, in the case of cross-regions costs are aggregated via integral images, whose use can drastically reduce computational load since summations involving matching cost for a pixel need to be performed only once[28]. In the case of rectangular support regions, the aggregation of pixel-wise costs can be efficiently performed by convolving through filters.

The cross-based aggregation was evaluated for the combinations of parameters $\{L, t, k\}$ = {10…55, step 5; 5…30, step 5; 7…15, step 2} (Figure 2), where $L$ is the maximum length of a support skeleton, $t$ the maximum colour difference, and $k$ is the census window length in each dimension. The rectangular windows were tested for the combinations of window size $L_N$ and census window size $k$ $\{L_N, k\}$ = {3…27, 3…21} (Figure 3). All 21 stereo-pairs of the Middlebury 2006 data-set have been used for parameter tuning as well as for producing the results of Figures 2 and 3.

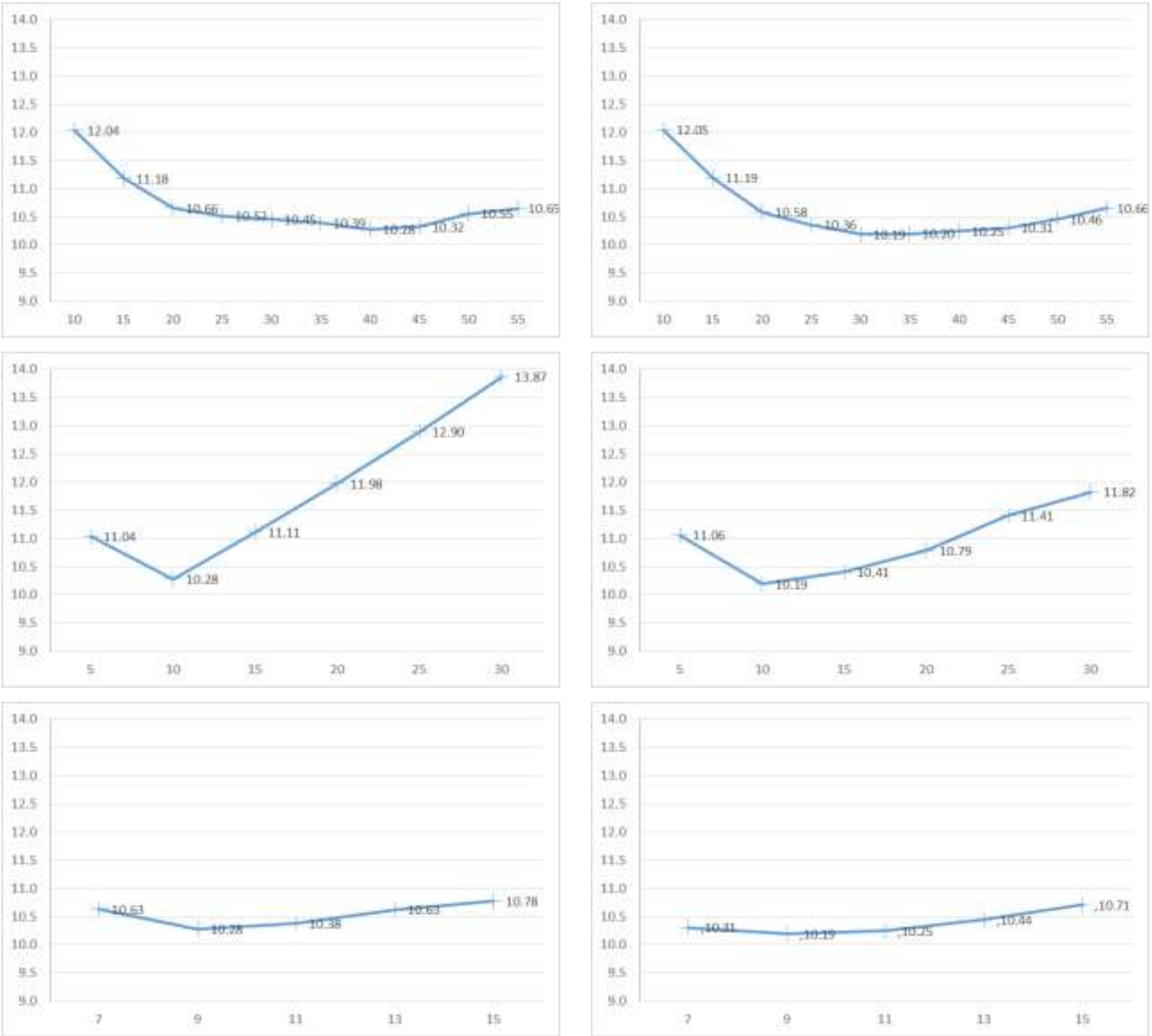



Figure 2. Diagrams from all pairs of the Middlebury 2006 dataset presenting the response of the cross-based algorithm to the tuning of individual parameters with the rest of the parameter set remaining constant at the values of the least average disparity error. The error percentages refer to non-occluded pixels. *Left*: matching on census-transformed image gradients; *right*: matching on census-transformed intensity images. *Top* to *bottom*: $L$, $t$, $k$.

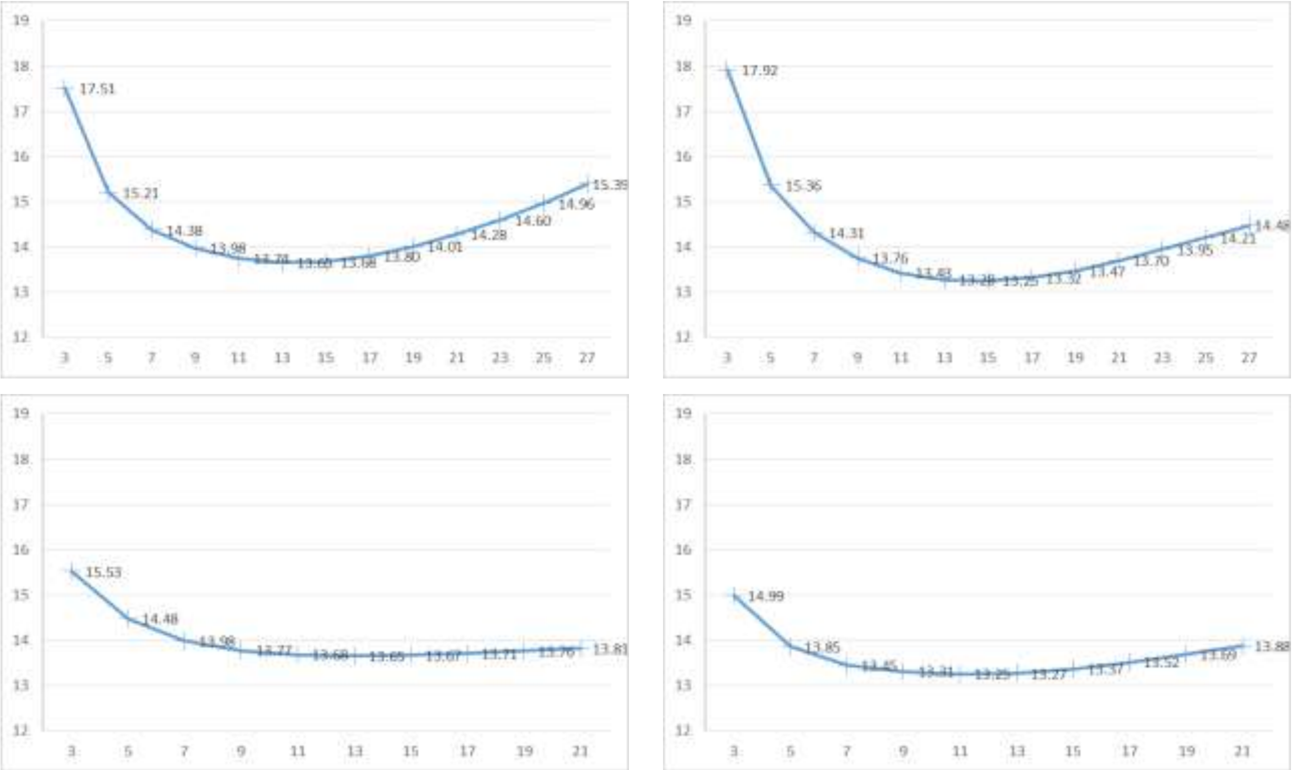



Figure 3. Diagrams from all pairs of the Middlebury 2006 dataset presenting the response of the rectangular windows algorithm to the tuning of individual parameters with the rest of the parameter set remaining constant at the values of the least average disparity error. The error percentages refer to non-occluded pixels. *Left*: matching on census-transformed image gradients; *right*: matching on census-transformed intensity images. *Top:* $L_N$, *bottom*: *k*.

## 4. EVALUATION

Original census on intensity (CI) and modified census transformation on gradients (CG) are evaluated on the complete Middlebury 2006 dataset, which includes different variants of each image of all pairs under three illumination conditions (1, 2, 3) and three exposure settings (0, +1, +2). The parameter tuning for each stereo-matching algorithm was performed for the default image configuration (illumination set 1, exposure setting +2) which corresponds to typical stereo pair conditions. Hence, the comparison of the two cost functions is done for the best performing parameter sets (Table 1) of each census transformation.

Table 1. Parameter values used for evaluating CG and CI after tuning

| Cross-based regions | | | Rectangular regions | | |
|---|---|---|---|---|---|
| | CG | CI | | CG | CI |
| *L* | 40 | 30 | $L_N$ | 13 | 15 |
| *t* | 10 | 10 | | | |
| *k* | 9 | 9 | *k* | 13 | 11 |

The average matching errors of the disparity maps estimated via the two cost functions are presented with respect to the combinations of exposure and illumination modes in Figure 4. The comparison regards the *"non-occluded"* and *"all"* pixels whose estimated disparity differs more than 1 pixel from the correct (ground truth) disparity. The errors obtained from the cross-based support region are presented in the upper diagrams, whereas the errors for the typical rectangular support regions are presented in the lower diagrams. Moreover, in Figure 5 a selection of four stereo-pairs and the estimated disparity maps are presented, in order to highlight the improvement in matching results when the cost is derived from census on gradients. The first two rows show the reference and matching image with default exposure +2, where the strong differences in illumination modes are illustrated. The forth and the fifth row illustrate the disparity maps estimated from CI and CG, respectively; areas in images where illumination changes create harsh changes in colour (e.g. different shadows, glowing) are estimated more accurately with the proposed method.

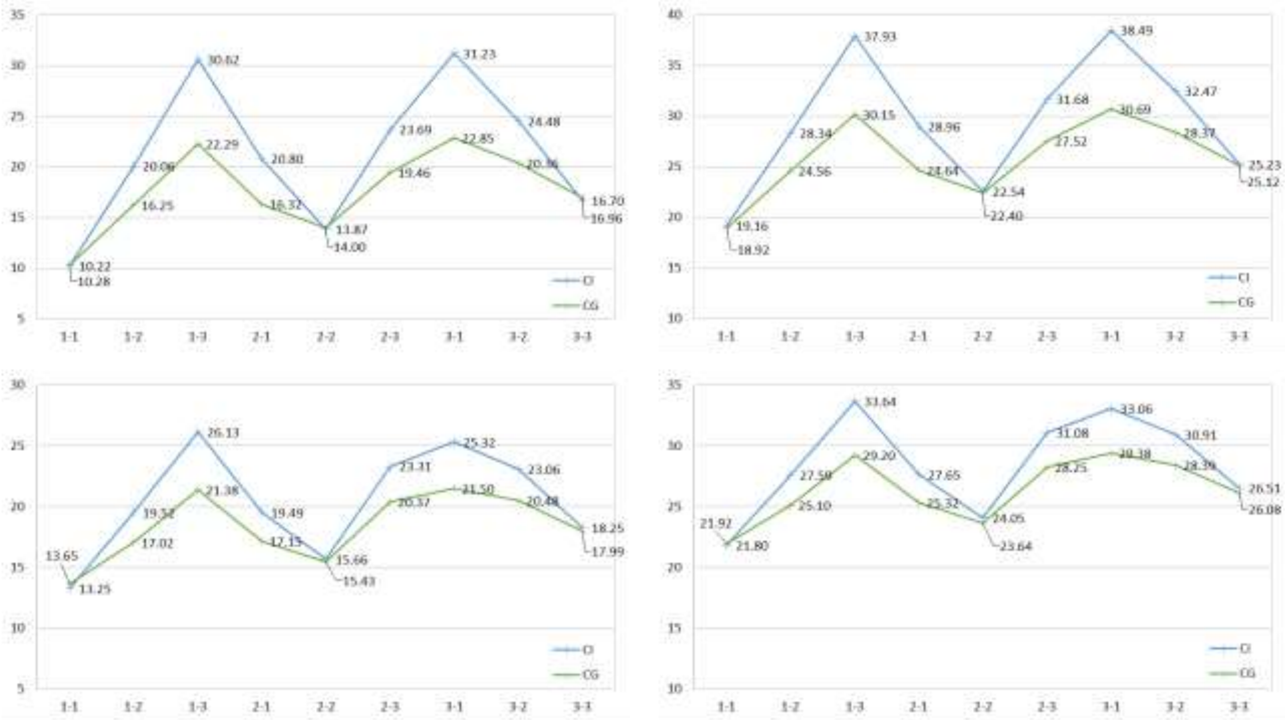
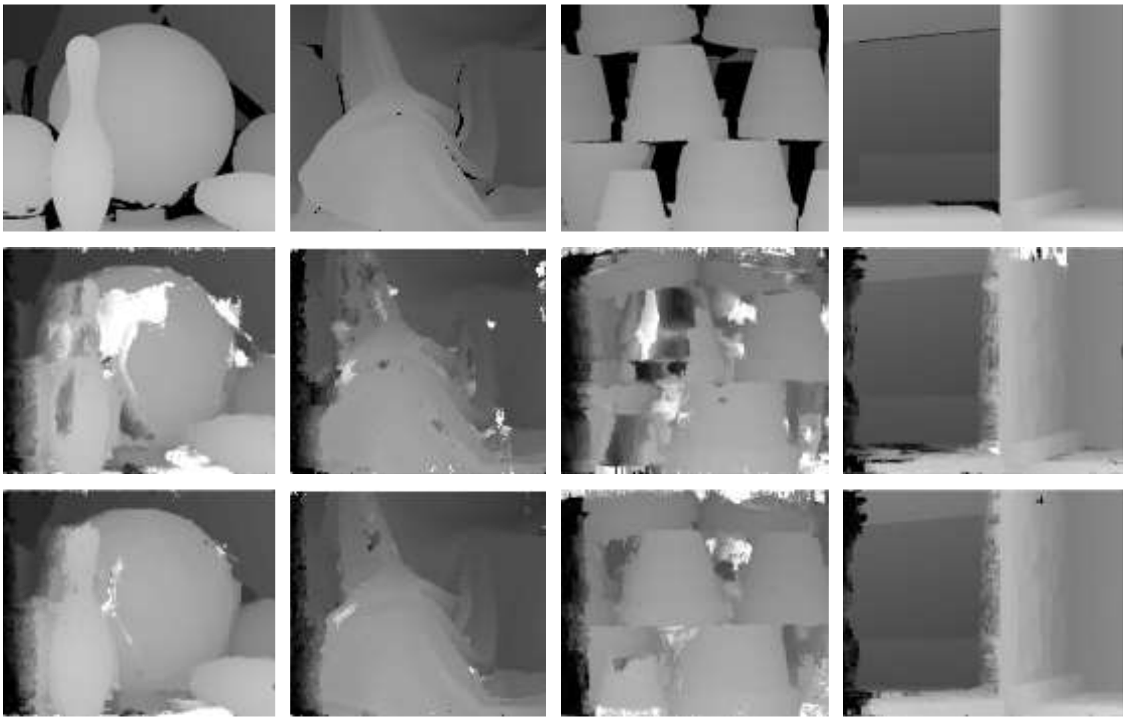



Figure 4. Comparison of the original census on intensity (CI) and the modified census on gradients (CG) under illumination changes for different aggregation methods evaluated on the Middlebury 2006 dataset. *Above*: cross-based support regions; *below*: typical rectangular support regions. *Left*: results for non-occluded pixels; *Right*: results for all pixels.



Bowling2 14.37 | Cloth3 8.02 | Flowerpots 22.24 | Wood1 1.66

Figure 5. Results of the original census on intensity (CI) and the modified census on gradients (CG) under illumination differences for cross-based aggregation. From top to bottom: base image (illumination mode 1); matching image (illumination mode 3); true disparity map; disparity map from CI; disparity map from CG; stereo-pair name and error improvement.

In Figure 6 the average matching percentage errors of the disparity maps of images with exposure differences calculated via CI and CG are presented for cross-skeleton based support regions. Results are also presented for the two different sets of *“non-occluded”* and *“all”* pixels. Changes in exposure affect the image intensities in a linear manner; hence CG does not yield better disparity maps than CI in the presence of exposure differences. This result is expected since census transformation is by definition invariant to linear changes.

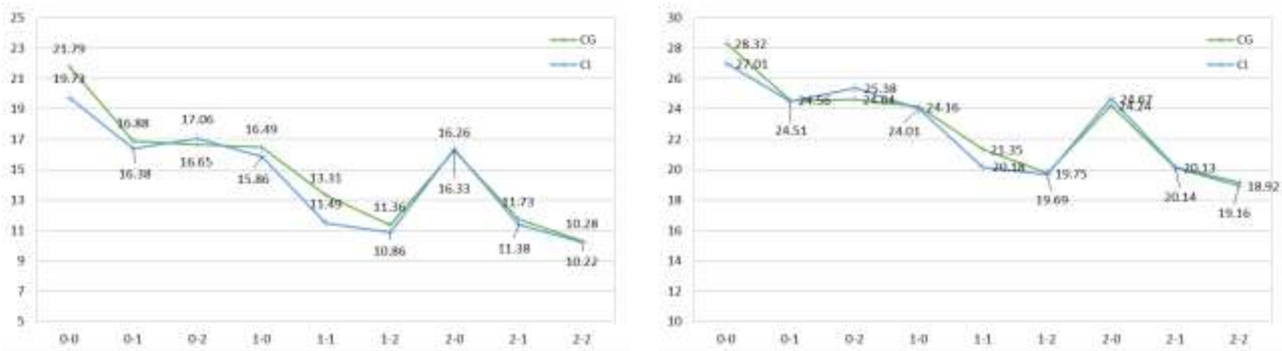



Figure 6. Comparison of the original census on intensity (CI) and the modified census on gradients (CG) under under different exposure settings for cross-based aggregation, also evaluated on the Middlebury 2006 dataset. *Left*: results for non-occluded pixels; *Right*: results for all pixels.

## 5. CONCLUDING RESULTS

In this paper a novel approach of census transformation was presented, and its efficiency on stereo matching of images with significant illumination differences is stated. The estimated disparity maps are essentially the same in cases with no illumination changes, but they are significantly better under illumination changes. In fact, the stronger the illumination changes the wider is the difference in efficiency between the original census transformation on intensity and its modification. On the other hand, exposure changes are linear changes and the census transformation is inherently resistant to this type of changes, thus no differences in disparity map accuracy worth mentioning emerged under exposure changes. As a future task, this cost function should be tested under a graph-cut optimization scheme and in conjunction with semi-global matching, in order to further verify its efficiency. Moreover, the theoretical justification of this matching cost needs to be elaborated. Finally, a similar evaluation should be done on the KITTI dataset, as this also presents radiometric differences within the stereo pairs.

## ACKNOWLEDGEMENT

This research has been co-funded by the European Union (European Social Fund) and Greek national resources under the framework of the Archimedes III: Funding of Research Groups in T.E.I. of Athens project of the Education & Lifelong Learning Operational Programme.

## REFERENCES

[1] Dhond, U. R.., Aggarwal, J. K., “Structure from stereo-a review,” IEEE Trans. Syst. Man Cybern. **19**(6), 1489–1510, IEEE (1989).

[2] Banks, J.., Corke, P., “Quantitative evaluation of matching methods and validity measures for stereo vision,” Int. J. Rob. Res. **20**(7), 512–532 (2001).

[3] Scharstein, D.., Szeliski, R., “A taxonomy and evaluation of dense two-frame stereo correspondence algorithms,” IEEE Int. J. Comput. Vis. **47**(1), 7–42 (2002).

[4] Brown, M. Z., Burschka, D.., Hager, G. D., “Advances in computational stereo,” IEEE Trans. Pattern Anal. Mach. Intell. **25**(8), 993–1008 (2003).

[5] Gong, M., Yang, R., Wang, L.., Gong, M., “A performance study on different cost aggregation approaches used in real-time stereo matching,” Int. J. Comput. Vis. **75**(2), 283–296 (2007).

[6] Tombari, F., Mattoccia, S., Di Stefano, L.., Addimanda, E., “Classification and evaluation of cost aggregation methods for stereo correspondence,” Proc. Comput. Vis. Pattern Recognit., 1–8, IEEE, Anchorage, AK (2008).

[7] Hirschmüller, H.., Scharstein, D., “Evaluation of stereo matching costs on images with radiometric differences,” IEEE Trans. Pattern Anal. Mach. Intell. **31**(9), 1582–1599 (2009).

[8] Wang, L., Gong, M., Gong, M.., Yang, R., “How far can we go with local optimization in real-time stereo matching - A performance study on different cost aggregation approaches,” Proc. Int. Symp. 3D Data Process. Vis. Transm. (3rd 3DPVT’06), 129–136, IEEE (2006).

[9] Nalpantidis, L., Sirakoulis, G. C.., Gasteratos, A., “Review of stereo vision algorithms: from software to hardware,” Int. J. Optomechatronics **2**(4), 435–462, Taylor & Francis (2008).

[10] Hirschmüller, H.., Scharstein, D., “Evaluation of cost functions for stereo matching,” Proc. Comput. Vis. Pattern Recognit., 1–8, IEEE (2007).

[11] Geiger, A., Roser, M.., Urtasun, R., “Efficient large-scale stereo matching,” Comput. Vision–ACCV 2010, 25–38, Springer (2011).

[12] Birchfield, S.., Tomasi, C., “A pixel dissimilarity measure that is insensitive to image sampling,” IEEE Trans. Pattern Anal. Mach. Intell. **20**(4), 401–406 (1998).

[13] Zabih, R.., Woodfill, J., “Non-parametric local transforms for computing visual correspondence,” Proc. Eur. Conf. Comput. Vis. **801**(May), 151–158, IEEE, Stockholm, Sweden (1994).

[14] Hirschmüller, H., “Stereo processing by semiglobal matching and mutual information,” IEEE Trans. Pattern Anal. Mach. Intell. **30**(2), 328–341 (2008).

[15] Tola, E., Lepetit, V.., Fua, P., “A fast local descriptor for dense matching,” Pr, 1–8, IEEE, Alaska, USA (2008).

[16] Strecha, C., Bronstein, A., Bronstein, M.., Fua, P., “LDAHash: Improved matching with smaller descriptors.,” IEEE Trans. Pattern Anal. Mach. Intell. **34**(1), 66–78 (2011).

[17] Miron, A., Bensrhair, A., Rogozan, A.., Ainouz, S., “Cross-comparison census for colour stereo matching applied to intelligent vehicle,” Electron. Lett. **48**(24), 1530–1532 (2012).

[18] Froba, B.., Ernst, A., “Face detection with the modified census transform,” Proc. IEEE Int. Conf. Autom. Face Gesture Recognit., 91–96, IEEE (2004).

[19] Puxbaum, P.., Ambrosch, K., “Gradient-based modified census transform for optical flow,” Proc. Int. Symp. Vis. Comput. **6453**, G. Bebis, R. Boyle, B. Parvin, D. Koracin, R. Chung, R. Hammoud, M. Hussain, T. Kar-Han, R. Crawfis, et al., Eds., 437–448, Springer Berlin Heidelberg, Berlin, Heidelberg (2010).

[20] Miron, A., Ainouz, S., Rogozan, A.., Bensrhair, A., “A robust cost function for stereo matching of road scenes,” Pattern Recognit. Lett.(38), 70–77, Elsevier B.V. (2014).

[21] Mei, X., Sun, X., Zhou, M., Jiao, S., Wang, H.., Zhang, X., “On building an accurate stereo matching system on graphics hardware,” ICCV Work. GPU Comput. Vis. Appl., 467–474, Barcelona (2011).

[22] Stentoumis, C., Grammatikopoulos, L., Kalisperakis, I.., Karras, G., “On accurate dense stereo-matching using a local adaptive multi-cost approach,” ISPRS J. Photogramm. Remote Sens. **91**, 29–49 (2014).

[23] Wu, Q., Zhao, X.., Li, X., “Stereo Matching Algorithm of Color Image Based on Weber’s Law and Census,” Proc. 10th Int. Conf. Intell. Inf. Hiding Multimed. Signal Process., 199–202, IEEE (2014).

[24] Faugeras, O.., Keriven, R., “Variational principles, surface evolution, PDE’s, level set methods, and the stereo problem,” IEEE Trans. Image Process. **7**(3), 336–344 (1998).

[25] Kolmogorov, V.., Zabih, R., “Computing visual correspondence with occlusions using graph cuts,” IEEE Int. Conf. Comput. Vis. **2**, 508–515, IEEE Comput. Soc (2001).

[26] Scharstein, D., “Matching images by comparing their gradient fields,” Int. Conf. Pattern Recognit. **1**, 572–575, IEEE Comput. Soc. Press (1994).

[27] Klaus, A., Sormann, M.., Karner, K., “Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure,” IEEE Int. Conf. Pattern Recognit., (3) 15–18 (2006).

[28] Zhang, K., Lu, J.., Lafruit, G., “Cross-based local stereo matching using orthogonal integral images,” IEEE Trans. Circuits Syst. Video Technol. **19**(7), 1073–1079 (2009).