

PAPER

An Efficient Local Stereo Matching Algorithm for Dense Disparity Map Estimation Based on More Effective Use of Intensity Information and Matching Constraints

Ali M. FOTOUHI^{†a)}, *Student Member* and Abolghasem A. RAIE[†], *Member*

SUMMARY In this paper, a new local matching algorithm, to estimate dense disparity map in stereo vision, consisting of two stages is presented. At the first stage, the reduction of search space is carried out with a high efficiency, i.e. remarkable decrease in the average number of candidates per pixel, with low computational cost and high assurance of retaining the correct answer. This outcome being due to the effective use of multiple radial windows, intensity information, and some usual and new constraints, in a reasonable manner, retains those candidates which satisfy more constraints and especially being more promising to satisfy the implied assumption in using support windows; i.e., the disparity consistency of the window pixels. Such an output from the first stage, while speeding up the final selection of disparity in the second stage due to search space reduction, is also promising a more accurate result due to having more reliable candidates. In the second stage, the weighted window, although not necessarily being the exclusive choice, is employed and examined. The experimental results on the standard stereo benchmarks for the developed algorithm are presented, confirming that the massive computations to obtain more precise matching costs in weighted window is reduced to about 1/11 and the final disparity map is also improved.

key words: *matching cost, disparity candidate, weighted window, radial window, validation-correction*

1. Introduction

Stereo vision or recognition of 3D structure from 2D images of the scene is one of the fundamental problems in machine vision and its wide modern applications have made it an active research area within the last few years. The algorithms proposed for correspondence search, which is the main problem in stereo vision, can be categorized into two groups: feature-based and area-based. In feature-based algorithms, the matching of feature elements such as corners, edges, and lines is considered and sparse disparity/depth map is produced. Area-based stereo algorithms accomplish the stereo matching for all pixels of the images and produce the dense disparity/depth map. The area-based algorithms could also be categorized into global and local approaches. Due to the fact that neighboring pixels in an image have common candidates in the other image to match, and selecting a match for each pixel effects selecting matches for other ones, global matching algorithms are proposed [3], [8]–[12]. In these algorithms an energy function is defined on the

basis of intensity consistency, disparity continuity and ordering constraints and is minimized on image line/surface. Because of using iterative optimization [8], [9] or dynamic programming [3], [10]–[12], as the minimization technique, the global matching approaches are computationally expensive. These algorithms usually have numerous parameters which are not simple to determine, and they often need an initial estimation of disparity values as well as accurate color segmentation of images. On the other hand, local matching area-based algorithms [1], [2], [4]–[7] considered in this paper, having less computational complexity, generally to compute reliable matching cost of pixels from different images, use some kind of statistical correlation of intensity or color of supporting windows around the pixels by the implied assumption of disparity consistency in the windows area. In these algorithms, generally, final selection of disparity of a pixel is performed simply by finding the disparity candidate with minimum matching cost among all disparity candidates for that pixel, known as WTA (Winner Takes All) rule.

In area-based algorithms, while the disparity continuity constraint may be used to determine matches in homogeneous or repetitive textured regions, it cannot be used to determine matches adjacent to depth discontinuities in a precise manner. Algorithms providing smooth disparity map may lose the details and maps with sufficient details are considered to be noisy. Though, the final aim is to find the matches, the occluded pixels in an image may not be matched with a pixel from another image. Obviously, thus, the dense stereo matching is an intrinsically complicated and ambiguous problem. In local algorithms, appropriate selection of the size and shape of the window to obtain a smooth disparity map with sufficient details is of great importance. In areas with low texture, the larger window may cover more intensity variations, causing more reliable matches. The larger window, on the other hand, may encompass more pixels with different disparity at depth discontinuity boundaries or on small objects of the image; hence, those parts of the windows corresponding to the disparity of the more textured regions dominate in the matching process independent of the disparity of the proposed pixel. This can cause an incorrect estimation of disparity in object boundaries and elimination of the details. The method of adaptive window is proposed in [1] and [2] to decrease this effect. In algorithm [1], the size and the shape of the window change iteratively on the

Manuscript received September 10, 2008.

Manuscript revised December 29, 2008.

[†]The authors are with Mobile Robot Research Lab, Electrical Engineering Department, Amirkabir University of Technology, Tehran, Iran.

a) E-mail: fotouhi@aut.ac.ir

DOI: 10.1587/transinf.E92.D.1159

basis of local variations of intensity as well as the present estimation of disparity of pixels within the window. The adaptive window methods are dependent upon the initial values of disparity and generally are computationally expensive. The multiple window technique, as a simplification of the adaptive window method, for reducing the computational complexity, limits the search space to windows with specific shapes and sizes [3], [4]. For instance, in [3], for each disparity candidate, the statistic correlation is computed for nine windows and the result of the window with minimum matching cost will be considered. In this approach, using minimum matching cost, while matching cost of different candidates may have been derived from different windows encompassing different regions of the image with different statistical characteristics, might need to be confirmed. In some other methods, assigning appropriate weights to the pixels of the window, is considered [5]–[7]. The main idea in these methods is to assign bigger weights to the pixels of the window being more probable to have disparities equal to the disparity of the proposed pixel at the center of the window. Contrary to the methods [5] and [6], determining the weights through initial disparity estimation, in [7] weights are determined according to their color similarity and geometrical proximity to the pixel at the center of the window. The weighted window of [7] shows better performance compared to other window-based methods; however, its efficiency is dependent upon computing weights through complex mathematical functions as well as using large windows.

Although all the window-based algorithms mentioned above have shown correspondence search improvement, what they have in common is that through massive and in some cases iterative computations for size and shape determination of windows and weight of pixels, a matching cost is computed and used to do the final selection. An ideal matching cost with high discriminating power which should reflect all necessary constraints and conditions and could select the match among large number of candidates with arbitrary statistical information, regardless of its computational complexity, seems unreachable. On the other hand it seems that proper search space reduction in a reasonable manner, retaining those candidates which satisfy more constraints and especially being more promising to satisfy the implied assumption in using support windows, i.e. the disparity consistency of the window pixels, could improve accuracy and computational performance. Based on this idea, in this paper an area-based local matching algorithm is proposed and developed. Section 2 of the paper describes the general functioning of the proposed algorithm. In Sect. 3 the first stage of the proposed algorithm, which is a new preprocessing algorithm to reduce the number of disparity candidates in each pixel, is considered. Section 4 investigates the weighted window, which is used in the second stage of the proposed algorithm for final decision making of disparity of each pixel. Section 5, presenting the experimental results, evaluates the effective performance of the proposed method to decrease the disparity candidates as well as the general

performance of the algorithm to estimate the dense disparity map. In Sect. 6 the concluding remarks are presented.

2. Logic and Block Diagram Overview of Proposed Algorithm

In this section, an overview of the general block diagram of the proposed algorithm and supporting ideas behind it, are presented. The two stages of the algorithm, shown in Fig. 1 are named “search space reduction” and “final disparity selection” abbreviated as SSR and FDS in the rest of the paper. At SSR the aims are: 1) to reduce the number of candidates per pixel as much as possible, 2) to retain those promising to fulfill necessary matching constraints and conditions, including the correct answer, 3) to maintain a computationally low cost procedure. Noting that all window-based approaches implicitly assume the disparity consistency of the window pixels (abbreviated as DCWP in this paper), i.e. all pixels in a support window are of the same depth, it seems reasonable to conclude that adaptive-window, multiple-window, weighted-window and similar techniques are trying to fulfill this assumption with massive computations. In other words, if a support window fulfilling this assumption could be determined; simpler cost functions could yield better results. Considering this important requirement, the sequence of steps and reasonable constraints and ideas employed in this stage, explained and justified in more details in the next section, are as follows.

- By usual epipolar, minimum and maximum disparity constraints, the initial set of candidates for each pixel p denoted by $cand_{p,int}$ is determined.
- Eight matching cost matrices of eight radial windows, named $cost_{\alpha}$ for a radial window in direction α , are computed. Using radial support windows around each pixel and in different directions makes it much more probable to have one of them more satisfying DCWP; hence, its corresponding window could be selected with a simpler matching cost function.
- Using “MinMin” process on each of the matching cost matrices $cost_{\alpha}$, for some eligible pixels, a candidate from $cand_{p,int}$ satisfying the left-right consistency con-

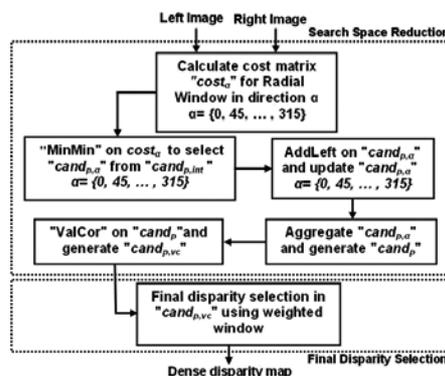


Fig. 1 Block diagram of proposed algorithm.

straint, named $cand_{p,\alpha}$, is determined.

- For each direction α , the pixels having not assigned $cand_{p,\alpha}$ from the MinMin process, are assigned one, using “AddLeft” process, based on disparity continuity constraint. So, after MinMin and AddLeft steps, for each direction, a $cand_{p,\alpha}$ from $cand_{p,int}$ is determined for each pixel p .
- By aggregating $cand_{p,\alpha}$ candidates from the previous steps, total candidates for each pixel p of the left image, known as $cand_p$, are obtained.
- Using validation-correction or “ValCor” process on the aggregated candidates, $cand_p$, candidates of $cand_{p,vc}$ are obtained. In this step, based on $cand_p$, the DCWP for the window of each retained candidate for each pixel p is double checked and if not satisfied is deleted or corrected.

In the FDS, the final disparity selection for each pixel p based on weighted window technique, is carried out on $cand_{p,vc}$ and dense disparity map is obtained. The weighted window computes more reliable matching costs through massive computations. However, in the proposed algorithm, due to low cost preprocessing in the first stage, the total computation cost is reduced by the search space reduction ratio. Furthermore, since non-reliable and noisy candidates are eliminated, more precise disparity map is expected. In the next two sections of the paper, the two stages of the algorithm are considered in more details.

3. Search Space Reduction

Considering the aims and steps of SSR, the precise definitions, relations, and formulas used in implementation and evaluation of algorithm, alongside necessary discussions and justifications are presented in this section. The initial set of candidates for each pixel p , $cand_{p,int}$, as usually used, is simply determined by epipolar and possible range of disparity constraints. With two rectified images, according to the epipolar constraint, if d_{min} is the minimum disparity and d_{max} is the maximum disparity of the matched pixels, the matching candidates of pixel p from the line i of the left image in location x , will be the pixels from the line i of the right image in locations $x - d_{max}$ to $x - d_{min}$. If the left border of the image is not considered, in pixel scale, we have $cand_{p,int} = \{d_{min}, \dots, d_{max}\}$. So, the number of disparity candidates for each pixel and also the average number of disparity candidates per pixel, abbreviated as $ANCP$ and used for performance evaluation in this paper, are in this case the same and equal to:

$$ANCP = d_{max} - d_{min} + 1 \tag{1}$$

In the local matching algorithms, it is necessary to compute $ANCP$ matching costs per pixel, and in the global matching algorithms, it is necessary for energy function to be evaluated, in each pixel, in this number of candidates and minimized by an optimization algorithm. Therefore, low cost reliable and appropriate reduction of $ANCP$ is valuable in both local and global techniques.

3.1 Radial Windows and Related Processes

Radial window definition: Eight radial windows with 45° of angle differences proposed in this paper, are illustrated in Fig. 2. In any direction shown, a ray of pixels with the length of n is considered and for any ray, two supporting regions in its left and right sides are also considered. L_α denotes a vector with the length of n , including the intensity value of the center pixel and $n - 1$ pixels in its continuance in the left image for a ray in direction α . $L_{1,\alpha}$ and $L_{2,\alpha}$ denote intensity vectors with the length of n for the two neighboring sides of the main ray for a ray in the direction of α in the left image. For instance, in Fig. 2 for $\alpha = 135$ and $n = 5$ the pixels taking part in vector L_α are demonstrated in black color, while the pixels taking part in $L_{1,\alpha}$ and $L_{2,\alpha}$ vectors are shown in grey color. In the same way, for the matching candidate pixel in the right image, vectors R_α , $R_{1,\alpha}$ and $R_{2,\alpha}$ are defined.

Matching cost function definition: The matching cost of the two pixels of p_l from the left image and p_r from the right one, in direction α , is defined as follows:

$$\begin{aligned} cost_\alpha(p_l, p_r) &= \sum_{i=1}^n \left| \left(L_\alpha(i) + \frac{L_{1,\alpha}(i)}{2} + \frac{L_{2,\alpha}(i)}{2} \right) \right. \\ &\quad \left. - \left(R_\alpha(i) + \frac{R_{1,\alpha}(i)}{2} + \frac{R_{2,\alpha}(i)}{2} \right) \right| \\ &\quad + \sum_{i=1}^n \left| (L_{1,\alpha}(i) - L_{2,\alpha}(i)) - (R_{1,\alpha}(i) - R_{2,\alpha}(i)) \right|; \\ &\quad \alpha = \{0, 45, \dots, 315\} \tag{2} \end{aligned}$$

The first part of the cost function is the sum of absolute differences of intensity value of the radial window pixels of the left and right images, with the ratios indicated for main and neighboring rays. Thus, in the computation of matching cost, the intensity information of a larger area take part, which in turn cause the reduction of noise effect, especially in homogenous or repetitive textured regions. Second part of the cost function is the sum of absolute gradient differences of the intensity value of pixels of the neighboring rays in the left and right images. The less the intensity gradient difference of the radial windows of the left and right images, the less the matching cost will be. If there is a great gradient difference, and especially if the gradient direction in the specified regions of the two images is different, the match-

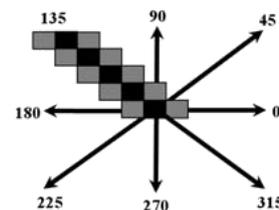


Fig. 2 Radial window pattern in proposed method.

ing cost will be increased. The matching cost function (1), while having low computational complexity, uses the intensity information in an effective manner.

Matching cost matrix calculations: Having calculated the matching costs of pixels for any pair of horizontal corresponding lines of the two images, using the matching cost function (1), eight matching cost matrices each belonging to one of the directions shown in Fig. 2, denoted by $cost_\alpha$; $\alpha = \{0, 45, \dots, 315\}$ in the block diagram, will be obtained.

MinMin process: In each matrix $cost_\alpha$, each point being minimum of both the row and the column, which means that is satisfying left-right consistency constraint, will be used to select a reliable candidate, named $cand_{p,\alpha}$. The existence of such a point on the i th row and j th column means that pixel p in location i from the left image and pixel q in location j from the right one, in two corresponding horizontal lines, are the best matching candidate for each other. The pixels that are not assigned a candidate in this process should mainly belong to homogenous and occluded regions; thus, being reasonable to avoid assigning them a candidate based on the matching cost.

AddLeft process: To determine the disparity candidate, $cand_{p,\alpha}$, those pixels remained unmatched through MinMin process, while they are supposed to be on the occluded or homogenous regions, on the basis of disparity continuity constraint, are assigned the exact $cand_{p',\alpha}$ of the pixel p' on the left side. Note that the pixels remained unmatched on the left side border of the image, are the exceptions of this rule and are assigned the exact $cand_{p',\alpha}$ of the pixel p' on the right. Note also that this rule holds if the dense disparity map of the left image is intended, otherwise the “left” and “right” words should be interchanged.

Aggregation process: In this process, for each pixel p , the outputs of MinMin and AddLeft processes on the eight related matching cost matrices, namely $cand_{p,\alpha}$; $\alpha = \{0, 45, \dots, 315\}$ are collected and named $cand_p$. Note that while cost matrix computations and MinMin process are valid for both left and right images, this process like AddLeft, is image dependent and is done for the pixels of the left image, as the disparity map of the left image is intended here. Although $cand_p$ is a set of eight candidates for each pixel p of the left image, each one corresponds to one direction, note that the disparity values of these eight candidates, are not necessarily distinct and in fact for many pixels they all may be the same. Note also that, since FDS is only carried out on distinct candidates, performance evaluation of SSR and so ANCP are determined in terms of the number of distinct candidates.

Comments: Two important points about using radial windows in this paper should be emphasized:

- The notion of multiple windows is used here differently from the way used in usual multiple-window approaches. Here matching cost matrix of the set of windows for each direction, $cost_\alpha$, is processed separately and comparing matching cost of different windows for a pixel, which is not justifiable, is not carried out.
- Because of using eight slim radial windows in different directions, while the more neighboring pixels take part in the matching cost computation, the probability of a continuum in which the overlapping with depth change does not occur will be highly increased and in case of desired functioning of the cost function, the existence of correct answer among the eight candidates obtained, will be certified with a high probability. Only in the cases where an image pixel is on a small cavity or a sharp projection in the scene, there will be depth change in every side around the pixel. However, the existence of such cases in the image is rare in practice and can be negligible. In addition, it seems that none of the other window-based methods can produce the reliable matching cost in these cases.

3.2 Validation-Correction Process

The new proposed process named validation-correction, or “ValCor” in the block diagram, is described in more details here. The process will be carried out on the obtained disparity candidates from MinMin and AddLeft processes for each pixel of left image, namely $cand_p$. The aim is to reduce the number of disparity candidates, by eliminating non-valid noisy ones, and to decrease the number of pixels without correct disparity candidate, by inserting new valid non-present disparity candidates.

The ValCor process is done on the basis of loosened DCWP, meaning that the cost aggregation in the support window area for a pixel may be valid if most of the pixels within the window have the same disparity as the disparity of the proposed pixel. Thus, the validity of a disparity candidate for each pixel, which is supported by one direction, is simply considered as the percent of pixels, within the correspondent radial window, having the same disparity value amongst their disparity candidates. In case a disparity candidate is supported by several directions, its validity is considered as the maximum validity computed for those supporting directions. If the calculated validity for a candidate is less than a threshold value, this candidate will be eliminated from the list of candidates for the nominated pixel. This stage of the ValCor process is known as “validation”.

In the most local matching algorithms as well as the algorithm of the present paper, the necessary computations to determine the disparity are performed in the pixel scale, whereas the disparity amounts especially on slanted surfaces can be on a sub-pixel scale. Thus, in the ValCor process, the validity of disparity values with the difference of one from any disparity candidate in the nominated pixel is also computed. Providing that these new disparity values have sufficient validity, they will be added to the list of disparity candidates for the nominated pixel. This part of the algorithm is known as the “correction”. Considering $cand_{p,vc}$ for each pixel p in the beginning of ValCor process as $cand_{p,vc} = cand_p$, the steps of ValCor process with more details are as follows:

1. The computation of validity for each disparity candidate: The validity of disparity candidate d_p for pixel p , shown by $v(d_p)$, is computed as:

$$v(d_p) = \text{MAX}_{\alpha \in S} \left\{ \frac{1}{N_R} \sum_{\bar{p} \in R_\alpha} B(\text{cand}_{\bar{p}}, d_p) \right\} \quad (3)$$

In Eq. (3), S is the set of directions which support disparity candidate d_p . R_α is the region from the left image occurring in the radial window located in pixel p and direction α . N_R is the total number of pixels present in this region, equal to $3n$ for the radial windows of the length n , and $\text{cand}_{\bar{p}}$ is the set of disparity candidates for pixel \bar{p} of this region after the aggregation process of SSR. Function B is defined as:

$$B(\text{cand}_{\bar{p}}, d_p) = \begin{cases} 1 & \exists d_{\bar{p}} \in \text{cand}_{\bar{p}}, d_{\bar{p}} = d_p \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

2. The elimination of disparity candidate with low validity: The disparity candidate d_p for pixel p will be eliminated from disparity candidates list of pixel p if its validity is less than a threshold value of T . That is to say:

$$v(d_p) < T \rightsquigarrow d_p \notin \text{cand}_{p,vc} \quad (5)$$

3. Correction: The validity of disparity candidates of $d_p - 1$ and $d_p + 1$, can be computed from Eq. (3). If the amount of validity of each of them is greater than the threshold T then, the same amount will be added to $\text{cand}_{p,vc}$, i.e. the set of disparity candidates of pixel p . In other words:

$$\begin{aligned} v(d_p + 1) > T &\rightsquigarrow (d_p + 1) \in \text{cand}_{p,vc}, \\ v(d_p - 1) > T &\rightsquigarrow (d_p - 1) \in \text{cand}_{p,vc} \end{aligned} \quad (6)$$

The validation phase of the ValCor process, has an effective role in the elimination of incorrect or noisy disparity candidates and consequently the reduction of $ANCP$, i.e. the average number of disparity candidates for each pixel. In this phase, the incorrect and scattered disparity candidates generally located in homogenous or low textured regions as well as occluded ones will be eliminated with a high probability, as they are not supported by the disparity candidates of the neighboring pixels. The correct disparity candidate in homogenous or low textured regions, as well as the disparity candidates equal to the disparity of one of the surfaces of the occlusion sides, will enjoy the support of disparity candidates of neighboring pixels and consequently would be kept with a high probability. To evaluate the validity of a disparity candidate, all supporting directions and for each direction all disparity candidates for the pixels present in the correspondent radial window will be considered. Thus, there is a very low possibility of the elimination of the correct disparity from the disparity candidates of a pixel, by the validation process. The correction phase of the ValCor process leads to the reduction of the number of pixels without correct disparity candidate, especially for the slanted surfaces in the image.

4. Final Disparity Selection

In the second stage of the proposed algorithm, it is necessary to decide the disparity of each pixel amongst the selected candidates of the first stage through using an efficient matching cost function. In this paper, the weighted window proposed in [7] which has shown more accurate results with respect to other window-based methods, is used. The main idea of the weighted window method is that those pixels of the window being more probable to have a disparity equal to that of the center pixel, have more significance in the computing of the matching cost. Relation (7) states that the weight of pixel q in a window with p as the center pixel, $w(p, q)$ should be proportional to the probability of the equality of pixel q disparity, d_q , and pixel p disparity, d_p .

$$\omega(p, q) \propto \text{Pr}(d_p = d_q) \quad (7)$$

In the proposed method, if ΔC_{pq} is the color difference and Δg_{pq} is the geometric distance of pixel q to the center pixel of the window, p , the weight allocated to the pixel q , $w(p, q)$ is computed as:

$$\omega(p, q) = \exp \left(- \left(\frac{\Delta C_{pq}}{\gamma_c} + \frac{\Delta g_{pq}}{\gamma_p} \right) \right) \quad (8)$$

In Eq. (8), γ_c and γ_p are constant coefficients. The more similar the color of the proposed pixel to that of the center pixel, and the more approximate the proposed pixel to the center pixel, it is more probable that these two pixels be of the same surface and identical disparity; thus, the proposed pixel must have more weight. A comparison of Eqs. (7) and (8) confirms this point. The matching cost of pixel p_l from the left image and its corresponding pixel $p_{r,d}$ from the right image for the disparity candidate d , is determined as:

$$\begin{aligned} \text{cost}(p_l, p_{r,d}) &= \frac{\sum_{q_l \in W_{p_l}, q_{r,d} \in W_{p_{r,d}}} \omega(p_l, q_l) \omega(p_{r,d}, q_{r,d}) e_0(q_l, q_{r,d})}{\sum_{q_l \in W_{p_l}, q_{r,d} \in W_{p_{r,d}}} \omega(p_l, q_l) \omega(p_{r,d}, q_{r,d})} \quad (9) \end{aligned}$$

W_{p_l} is the window surrounding pixel p_l in the left image and $W_{p_{r,d}}$ is its corresponding window in the right image. q_l and $q_{r,d}$ are the corresponding pixels of the left and right images, located in the windows W_{p_l} and $W_{p_{r,d}}$ respectively. $e_0(q_l, q_{r,d})$ represents the pixel-based raw matching cost and can be expressed as:

$$e_0(q_l, q_{r,d}) = \min \left\{ \sum_{c \in \{r,g,b\}} |I_c(q_l) - I_c(q_{r,d})|, T \right\} \quad (10)$$

I_c is the intensity value of color c and T is a truncation value determining the limit of matching cost. In [7], after the matching cost computation through Eq. (9) for the set of disparity candidates of pixel p_l ($\text{cand}_{p_l,int}$) the disparity of each pixel p_l is simply selected using the WTA rule as:

$$d_p = \arg \min_{d \in \text{cand}_{p_l,int}} \text{cost}(p_l, p_{r,d}) \quad (11)$$

The most important shortcoming of the weighted window method is its high computational cost, which as discussed in Sect. 5, is improved significantly because of using reduced disparity candidates of $cand_{p,vc}$, while its error performance is also improved.

5. Evaluation of the Proposed Algorithm

5.1 Bench Marks

The proposed algorithm is evaluated on standard stereo images “Tsukuba”, “Sawtooth”, “Venus” and “Map”, generally used to compare different area-based algorithms [13]. In Fig. 3 (a) left images and in Fig. 3 (b) ground truth disparity maps of these stereo images are presented.

5.2 Evaluation of Search Space Reduction Stage

In Sect. 2, three objectives were enumerated for SSR. Here, the performance of the SSR, regarding the first two objectives, namely reduction rate of candidates as well as the capability of retaining correct answers amongst remaining candidates, is evaluated. The third objective of SSR, regarding computational cost, is evaluated in related Sect. 5.4.

To implement the SSR, radial windows with the length of $n = 15$ are used. This length being about half of the side length of the square window in the FDS, as determined in the next subsection, seems sensible, although not necessarily being the only choice. Considering this ratio for the size of windows in the two stages, radial windows in SSR would become radial scans of the region covered by the square window in FDS. Furthermore, reasonable amount of $T = 0.75$, is used in the process of ValCor as the threshold value.

To evaluate the first objective, the average number of distinct disparity candidates for each pixel, $ANCP$, before applying SSR, as computed from Eq. (1), for each image is presented in the first row of Table 1. The value of $ANCP$, after applying SSR, is presented in the second row of Table 1. The last row of the table shows the reduction ratio of $ANCP$ for each image. Significant reduction ratio from 8.2 to 16.2 with the average of 11 on the standard bench marks represents remarkable achievement on the first objective of SSR.

In reduction of $ANCP$, each step in SSR has different behavior. For instance, for Venus image after the MinMin and AddLeft processes, the value of $ANCP$ reduces from 20 to 2.08 and using ValCor process reduces to 1.91. Although, reduction of candidates due to the latter process compared to the preceding ones is not noticeable, it should be noted

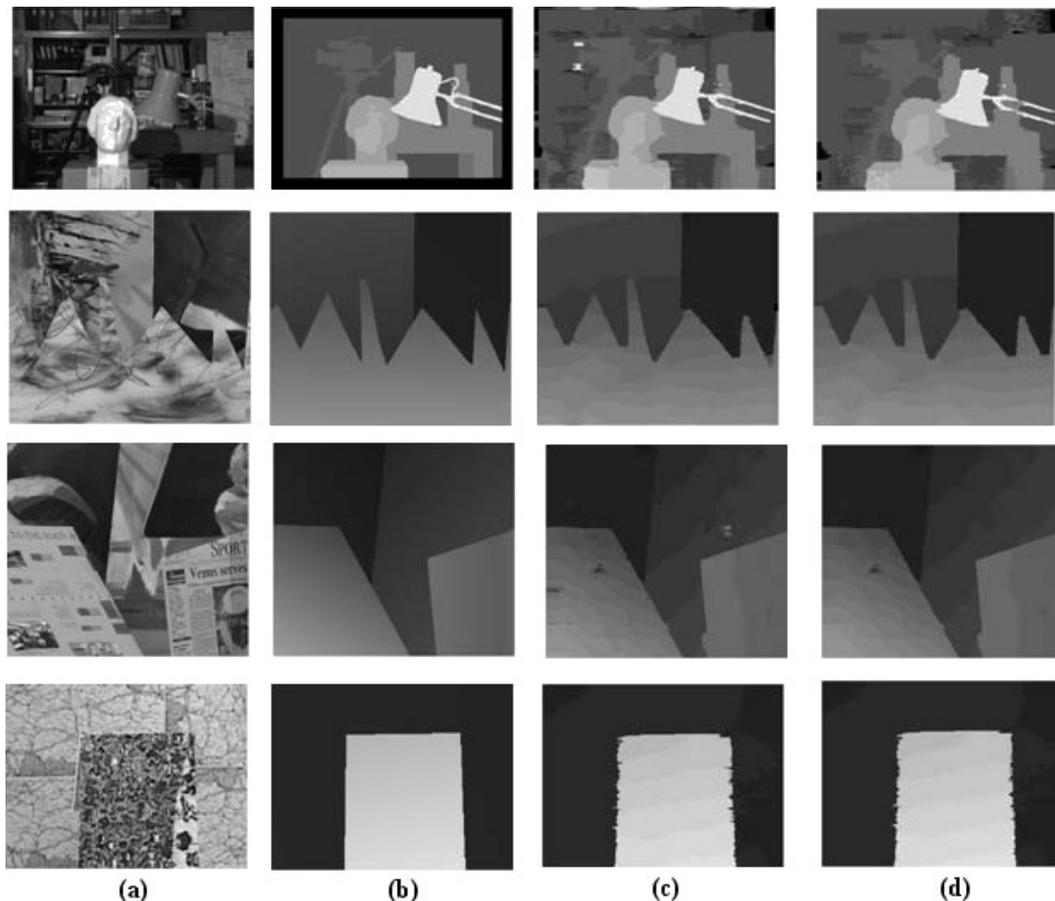


Fig. 3 (a): Left Image, (b): ground truth disparity, (c): Results without using SSR, (d): Results by using SSR.

Table 3 Performance of proposed algorithm in different regions based on WTA rule.

	Tsukuba			Sawtooth			Venus			Map	
	nonocc	untex	disc	nonocc	untex	disc	nonocc	untex	disc	nonocc	disc
without SSR	3.21	2.95	7.92	1.32	0.29	5.32	2.63	3.39	7.65	1.62	11.46
with SSR	2.49	2.76	8.57	1.12	0.23	5.04	1.04	0.94	6.88	0.66	7.70

Table 1 Performance of search space reduction algorithm.

	Tsukuba	Sawtooth	Venus	Map
ANCP before SSR	16	20	20	30
ANCP after SSR	1.94	1.87	1.91	1.85
ANCP Reduction Ratio	8.2	10.7	10.5	16.2

Table 4 Performance of proposed algorithm with LRCC.

	Tsukuba	Sawtooth	Venus	Map
without SSR	3.25	1.80	1.61	1.67
with SSR	3.60	1.62	1.01	1.12
without Val-Cor				
with SSR	2.92	1.60	0.80	1.06

Table 2 Percent of pixels without correct candidate.

	Tsukuba	Sawtooth	Venus	Map
Non-Occluded	0.27	0.19	0.07	0.0
All	0.29	0.57	0.12	0.0

that after ValCor process validated candidates are selected, while the process of correction is carried out. This in turn, may cause increasing the number of candidates to minimize pixels without correct answer.

To evaluate the second objective of SSR, note that the pixels without correct disparity candidate determine the lower bound of error in a WTA based rule algorithm; therefore, they are of great importance in evaluation of this stage. Most local matching algorithms estimate correct disparity values in occluded regions through the use of some post processing methods such as LRCC (Left-Right Consistency Check). Thus, in these cases the existence of correct disparity amongst remaining disparity candidates in occluded regions is of less importance. In the first row of Table 2 the percent of pixels without correct disparity in the non-occluded regions for different images is presented. In the second row, this percent is also presented for all pixels of images. Comparing percent of pixels without correct disparity candidate after SSR in Table 2, with the final error of algorithm in Tables 3 and 4, reveals that the share of SSR in final error is really negligible. For instance, considering the first row of Table 2 and the second row of Table 3 for non-occluded pixels, shows that the share of this stage is in average 10% of the final error for different images, meaning quite acceptable fulfillment of the second objective too.

In short, Tables 1 and 2 reveal that after SSR, by significant reduction of disparity candidates or search space, reliable information for final decision making are provided.

5.3 Evaluation of Estimated Dense Disparity Map

Weighted window, as described in Sect. 4, is used in the second stage (FDS) to estimate dense disparity map, with two promising objectives. The first one was expecting significant saving in massive computations of weighted window, due to SSR stage, which is now revealed to be true with the average of 11 times. The second promising objective was expected error reduction which is evaluated in the followings.

In weighted window method, proper value assignment to parameters γ_c , γ_p , T and window size, play important role on the performance, regarding the amount of the final error. In [7], while large windows are recommended, it is reported that for small windows, for instance 5×5 windows, percentage of error in dense disparity map estimation by weighted window is about two or three times more than the case of using simple SAD (Sum of Squared Differences) matching cost function. Since finding more suitable values for the parameters is out of the scope of this paper, the proposed values in [7], namely $\gamma_c = 7$, $\gamma_p = 36$, $T = 40$ and window size of 33×33 , are used in the implementation of FDS.

In Table 3 the performance of the proposed algorithm in the estimation of dense disparity map is shown, while the disparity of each pixel is selected simply based on WTA rule. Results without using SSR are also presented in Table 3, for comparison. Numbers in Table 3 represent the percents of bad pixels, i.e. pixels having absolute disparity error greater than one, in the regions of non-occluded, untextured and near the depth discontinuities.

To correct the disparity values in the pixels of occluded regions and border of image, LRCC process is used. Obtained dense disparity maps using SSR for different images are depicted in Fig. 3 (d). The maps obtained without using SSR are also presented in Fig. 3 (c). Table 4, in the first and third rows, shows the percent of the error for both versions in all of the pixels of different images, for more precise comparison. To see the effect of ValCor process in the proposed algorithm, the second row of Table 4 presents the results of algorithm without using ValCor process in SSR.

Considering experimental results reveals that applying SSR, more probably due to choosing the validated candidates and eliminating the noisy ones, leads to more precise results. This can be seen by comparing the results in Table 3, for non-occluded and different regions of the image, based on WTA rule. Also, comparing the results of the first and the third rows of Table 4 reveals that using SSR reduces the overall error of dense disparity map by the average percent of 25 on different images. The comparison of the second and the third rows of Table 4 with its first row confirms that the SSR having ValCor process always improves the results. On the other hand, the lack of ValCor process in SSR may in some cases; e.g. Tsukuba image in Table 4, lead to worse results. The better results caused by the using of ValCor pro-

cess may be attributed to the fact that correction phase has had a considerable effect on the compensation of the correct candidates elimination in earlier steps of SSR.

5.4 Overall Speedup Ratio

Up to this point, it is verified that the achievement of the objectives of this study has been realized except the one related to the computational cost of SSR and its effect on the overall speedup. Although it is to some extent clear, following evidences and calculations seem to be enough to validate that the computational overhead due to SSR stage, compared to the second stage is negligible. Therefore, the overall speedup factor would almost be the same as the reduction factor in SSR, with an average value of 11 for the bench marks used in this paper:

- Computational complexity, in terms of problem dimensions (N_p = the number of pixels in left image, D_r = $d_{\max} - d_{\min} + 1$ = possible range of disparity, size of square window, number and size of radial windows) for SSR compared to normal weighted window method is the same except for replacing the side size of square window with the number of radial windows. Though, since the methods are working for specific and narrow range of dimension values, not only the order of complexity but those valid values should be used to determine the performance. For example, in our problem considering both value and order, the number of pixels involved in cost function calculation in a square window is 1089, while in all eight radial ones is 360.
- All calculations in the processes of SSR are based on simple and fast fixed-point add, subtract, compare and shift operations, while cost function calculations for weighted window involve expensive time consuming floating-point operations and mathematical functions.
- Considering the number of required operations and functions and their execution times in C++, on the Intel P4 2.66 GHz PC, the ratio of computational cost per pixel for SSR (CP_{ssr}) and FDS (CP_{ww}) is found to be $CP_{ssr}/CP_{ww} = 25/10000$. if CC_{ww} and CC_{ssr-ww} represent respectively the total computational cost of weighted window method without and with search space reduction, then the following equations hold:

$$CC_{ww} = N_p * CP_{ww} * D_r \quad (12)$$

$$CC_{ssr-ww} = N_p * CP_{ssr} * D_r + N_p * CP_{ww} * D_r/R_c \quad (13)$$

$$CC_{ssr-ww}/CC_{ww} = (25/10000 + 1/R_c) \approx 1/R_c \quad (14)$$

Considering the range of values obtained for reduction factor, R_c (8.2–16.2), the overhead due to SSR stage in speedup determination is negligible, so R_c represents both of reduction factor and speedup factor at the same time. Using the proposed algorithm, the time needed to estimate the dense disparity map, for instance in

Tsukuba image, with $R_c = 8.2$, reduces from 12 minutes to about 1.5 minutes.

6. Conclusions

The proposed local matching algorithm has shown major performance improvement, regarding accuracy and speed, for weighted window method. The idea behind the work was that better results can be obtained if usual information and constraints are used in a better format and in a more proper sequence. Therefore, in the first stage of the algorithm, multiple radial windows with a new notion, intensity information, left-right consistency constraint, disparity continuity constraint and disparity consistency of the window pixels, in a reasonable manner and in a preprocessing search space reduction process were employed. The main properties of this stage were; significant reduction of search space, while retaining the correct answers, using low computational cost procedures. All these properties were validated using standard bench marks, resulting average reduction factor of 11, negligible computation time with respect to massive calculation of matching cost function of weighted windows and negligible pixels without correct candidate with respect to the overall error performance of WTA-based weighted window method. Although such an output would speedup many local and global computationally expensive algorithms, the weighted window was selected and examined for final disparity selection as the second stage in this paper. The presented experimental results indicate an overall speedup factor of 11 and 25% error reduction in dense disparity map estimation using weighted window method.

It seems worthwhile that continuing this path, using the information about remaining candidates in the reduced search space, other constraints and conditions and more proper sequence of them, could still improve the performance of weighted window as well as some other computationally expensive techniques.

References

- [1] T. Kanade and M. Okutomi, "A stereo matching algorithm with an adaptive window: Theory and experiment," IEEE Trans. Pattern Anal. Mach. Intell., vol.16, no.9, pp.920–932, 1994.
- [2] O. Veksler, "Stereo correspondence with compact windows via minimum ratio cycle," IEEE Trans. Pattern Anal. Mach. Intell., vol.24, no.12, pp.1654–1660, 2002.
- [3] A.F. Bobick and S.S. Intille, "Large occlusion stereo," Int. J. Comput. Vis., vol.33, no.3, pp.181–200, 1999.
- [4] H. Hirschmuller, P.R. Innocent, and J. Garibaldi, "Real-time correlation based stereo vision with reduced border errors," Int. J. Comput. Vis., vol.47, no.1-3, pp.229–246, 2002.
- [5] T. Darrel, "A radial cumulative similarity transform for robust image correspondence," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp.656–662, 1998.
- [6] Y. Xu, D. Wang, T. Feng, and H.Y. Shum, "Stereo computation using radial adaptive windows," Proc. Int'l Conf. Pattern Recognition, vol.3, pp.595–598, 2002.
- [7] K.-J. Yoon and I.-S. Kweon, "Adaptive support-weight approach for correspondence search," IEEE Trans. Pattern Anal. Mach. Intell., vol.28, no.4, pp.650–656, April 2006.

- [8] Q. Yang and L. Wang, "Stereo matching with color-weighted correlation, hierarchical belief propagation and occlusion handling," IEEE Conf. Computer Vision and Pattern Recognition, 2006.
- [9] L. Zitnick and S.B. Kang, "Stereo for image-based rendering using image over-segmentation," Int. J. Comput. Vis., vol.75, no.1, pp.49–65, 2007.
- [10] P.H.S. Torr and A. Criminisi, "Dense stereo using pivoted dynamic programming," Elsevier, Image and Vision Computing, vol.22, pp.795–806, 2004.
- [11] Kim J. Bobick, Kyoung T., Byoung T. and Sang L., "A dense stereo matching using two-pass dynamic programming with generalized ground control points," Proc. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 19/05: 1063-69.
- [12] L. Wang, M. Liao, M. Gong, R. Yang, and D. Nister, "High-quality real-time stereo using adaptive cost aggregation and dynamic programming," Proc. Third Int. Symposium on, 3D Data Processing, Visualization, and Transmission, pp.798–805, June 2006.
- [13] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," Int. J. Comput. Vis., vol.47, no.1-3, pp.7–42, 2002.



Ali M. Fotouhi was born in 1977 in Yazd, Iran. He received his B.S. and M.S. degrees in E.E. from Iran Univ. of Science and Tech. and Amirkabir Univ. of Tech. (Tehran Polytechnic) in 2000 and 2003, respectively. He has been a Ph.D. candidate at E.E. Dept. of Amirkabir Univ. of Tech. in Iran since 2003. His research interests are mobile robot navigation, robot vision, image processing, PC Interfacing and Hardware Design. Email:fotouhi@aut.ac.ir



Abolghasem A. Raie was born in Tehran, Iran (1951). He received his B.S. degree in E.E. from Sharif Univ. of Tech. (Iran 1973) and his M.S. and Ph.D. degrees in E.E. from Univ. of Minnesota (U.S.A. 1979 and 1982). From 1983 till now, he is a member of the faculty of E.E. Dept. at AmirKabir Univ. of Tech. in Iran. His research interests are Algorithm design and performance analysis, machine vision, sensor fusion and mobile robot navigation. Email:raie@aut.ac.ir