PAPER

Fast Edge-Based Stereo Matching Algorithms through Search Space Reduction

Payman MOALLEM†, ***Student Member***, **Karim FAEZ**†, ***Regular Member***, ***and*** **Javad HADDADNIA**†, ***Nonmember***

SUMMARY Finding corresponding edges is considered being the most difficult part of edge-based stereo matching algorithms. Usually, correspondence for a feature point in the first image is obtained by searching in a predefined region of the second image, based on epipolar line and maximum disparity. Reduction of search region can increase performances of the matching process, in the context of execution time and accuracy. Traditionally, hierarchical multiresolution techniques, as the fastest methods are used to decrease the search space and therefore increase the processing speed. Considering maximum of directional derivative of disparity in real scenes, we formulated some relations between maximum search space in the second images with respect to relative displacement of connected edges (as the feature points), in successive scan lines of the first images. Then we proposed a new matching strategy to reduce the search space for edge-based stereo matching algorithms. Afterward, we developed some fast stereo matching algorithms based on the proposed matching strategy and the hierarchical multiresolution techniques. The proposed algorithms have two stages: feature extraction and feature matching. We applied these new algorithms on some stereo images and compared their results with those of some hierarchical multiresolution ones. The execution times of our proposed methods are decreased between 30% to 55%, in the feature matching stage. Moreover, the execution time of the overall algorithms (including the feature extraction and the feature matching) is decreased between 15% to 40% in real scenes. Meanwhile in some cases, the accuracy is increased too. Theoretical investigation and experimental results show that our algorithms have a very good performance with real complex scenes, therefore these new algorithms are very suitable for fast edge-based stereo applications in real scenes like robotic applications.

key words: *fast edge-based stereo, search spaces reduction, directional derivative of disparity, hierarchical multiresolution*

1. Introduction

Stereo vision refers to the ability to infer information on 3D structures and distances of a scene from at least two images (left and right), taken from different viewpoints. A stereo system must solve two essential problems: *correspondence* and *reconstruction* [1]. The correspondence consists of determining which item in left image corresponds to which item in right image. It is usually not a good practice to try to find corresponding points for all pixels. For example, a point in a uniform region in one image may correspond to many points in a corresponding region in other image. Thus, feature points or matching primitives are selected so that an unambiguous match could be resulted [2]. Depth in a stereo system is inversely related to disparity, which is the difference between position of corresponding points in left and right image. Disparity of image points forms so-called disparity map, which can be displayed as an image. If geometry of stereo system is known, the disparity map can be converted to a depth map of viewed scene. This process is called the reconstruction step in the stereo algorithms.

The correspondence, which is the most complex stage in stereo algorithms, is classified into three categories: area-based (or correlation based), feature-based, and pixel-based methods [3]. In the area-based methods, the elements to be matched are image windows of fixed or variable sizes, and similarity criterion can be a measure of correlation between the windows in two images. This process consists of extracting feature points in one image (for instance left one) and finding their corresponding points in other image. Feature-based methods match feature points in left image to those in right image, so these methods restrict search for the correspondence to a sparse set of feature points. Instead of using image windows in area-based approaches, they use numerical and symbolic properties of features, available from feature descriptors. Feature-based methods use a measure of distance between feature descriptors. In these methods, interpolation is a necessity for computing the depth for non-feature points. Pixel-based methods perform the matching at each pixel, using the measurements at a single pixel. In these methods, a dense disparity map can be obtained.

In some stereo vision applications, computational fast stereo techniques are required, especially for mobile robots [4], [5], autonomous vehicle [6], [7], virtual reality [8] and stereo image coding in 3D-TV [9]. Parallel algorithms [10], [11], multiprocessors [4], [11] and special purpose hardware [12] are posed to speed up the processing time, but in this paper we are going to discuss about fast algorithms and techniques. Dense depth map is not often required for many of the fast stereo applications because its computation is time consuming. Often, computation of distance between camera system and objects in scene is exclusive goal of the stereo systems. Thus, correspondence search in stereo images can be reduced to matching of the most fea-

Manuscript received September 11, 2001.
Manuscript revised February 25, 2002.

†The authors are with the Faculty of Electrical Engineering, Amirkabir University of Technology, Tehran 15914, Iran.

ture points or matching primitives [11]. Therefore the correspondence includes two stages that are executed sequentially: primitive extraction and matching. Generally, processing time in the matching stage is more than the primitive extraction, but in fast applications two stages are important, especially when the computations in the matching stage are reduced. Thinned edge points as primitives are used most often in fast stereo vision [11]–[15] and these stereo techniques are usually called edge-based stereo algorithms. In this paper, we suggested some fast edge-based stereo algorithms.

After edge extraction and thinning, the matching is established. Generally to find the matching of each edge in one image, all pixels in other image must be examined as the search region. The execution time of the matching can be reduced by two approaches: reducing examined points and reducing search space. If all pixels in other image are examined, the execution time is increased but the accuracy is high too, this method acts like as area-based [16]. If only the edge pixels are examined, the execution time is reduced very much and the accuracy is also reduced [15]. Generally in fast applications, the second method is chosen. The reduction of the search region can reduce complexity of matching and increase the accuracy. Most stereo matching algorithms narrow down number of possible matches for each feature by enforcing suitable constraints on feasible matches and proper matching strategies, which are discussed in the next subsections.

1.1 Matching Constraints

In the stereo correspondence, some matching constraints are generated based on underlying physical principles of world imaging and stereopsis. Some of the common constraints incorporated in stereo algorithms include:

• *Epipolar constraint:* Corresponding points must lie on corresponding epipolar lines. Epipolar lines are defined by intersection of epipolar planes and image plane. For each point in a 3D scene, the corresponding epipolar plane is a plane that contains the point and the stereo base line, which is the line connecting the two optical centers of the cameras [1].

• *Disparity limit constraint:* Regarding to maximum and minimum of depth and geometry of a stereo system, the maximum disparity range can be estimated [10].

• *Ordering constraint:* A left-right ordering relationship between two pixels in left image should have the same ordering correspondence pixels as in right image. In other words, matched primitives must occur in the same left to right orders along epipolar lines [17].

• *Uniqueness constraint:* Correspondence should be unique. In other words, each feature can have at most one match [17].

• *Figural continuity constraint:* Disparity along an edge contour changes smoothly, so there should be no disparity discontinuities along a contour. Moreover disparity allowed to abruptly changing across the contours [18].

• *Limit of directional derivative of disparity:* Maximum of directional derivative of disparity is limited, so this value between matched primitives is restricted [19]. The directional derivative of disparity is often incorrectly called the disparity gradient [20].

The epipolar constraints reduce the search region from the whole of the second image (two-dimensional space) to the epipolar line (one-dimensional space) [1]. Moreover, the disparity limit narrows down the one-dimensional search from the full search to a limited space [10]. In some special case, ordering constraint can also reduce the search region [17]. Moreover, probability density function of directional derivative of disparity can be used to reduce the search space in a disparity surface search technique [21] and edge matching [22], considering some geometric parameters of stereo systems. The uniqueness constraint [1], the figural continuity constraint [18] and the limit of the directional derivative of disparity are used to detect false matches and to correct them [23].

1.2 Matching Strategy

Besides the constraints and consistency checks, several control strategies have been proposed by many to further reduce the search region and the ambiguity, and to enhance stereo matching performances. Some of the common and popular ones are:

• *Coarse to fine strategy:* In a coarse to fine strategy, information obtained at a coarse scale is used to guide and limit the search for the matching of finer scale primitives or feature points. In this approach, an initial matching begins at a coarse scale where feature density is low due to scale change. This reduction in the feature density reduces the search space, which in turn makes the matching easier and faster, but not necessarily more accurate, because localization at coarse scale is less accurate. Such multistage strategy can be used with scale specific primitive representation and can be incorporated into the area-based, the feature based, and the pixel based correspondence techniques. These techniques are classified into two categories: Initial estimation at lowest scale called multiresolution and initial estimation at finest scale. In the multiresolution strategies, the initial matching begins at the lowest scale and then extended to finer ones [2]. This approach can be applied in both temporal space [24], [25] (gray level) and frequency space. Usage of the hierarchical Gaussian basis functions [3] (in temporal space) and complex multiwavelets [26] (in both temporal and frequency space) are common techniques in stereo vision. In the finest scale method, the disparity estimation of large blocks in the fine level is used as initial estimation of disparity for the entirety of blocks, and then the disparity is calculated within each block accurately [27], [28].

• *Structural and hierarchical multi primitives:* In hierarchical and structural stereo approaches, semantically rich primitive representations like regions, lines and edge segments are derived from an image and matched. Relational properties are used besides spectral properties in the structural methods to reduce the search space and to disambiguate the stereo matching [2]. In hierarchical systems, the matching takes place between more than one level of image descriptions. In these methods, the reduction in feature density and so in search space is achieved through an abstraction of higher level structures without a scale change [17], [29].

None of these strategies uses the stereo constraints to reduce the search region, but we will introduce utilization of a threshold on the directional derivative of disparity to do so. Hereinafter, we use sometimes the DDD for the directional derivative of disparity. In next section, we briefly discuss DDD, its concept and application in some stereo methods. Then we will formulate a reduction of the search region in the stereo matching of connected non-horizontal edge points in successive scan lines. Then we will develop a very fast stereo correspondence strategy based on this new formulation for the connected non-horizontal edge points as primitives. This strategy can be used with other primitives and methods. Finally, we will propose some very fast edge-based stereo algorithms and discuss about their results.

2. Basics of DDD

2.1 Definitions

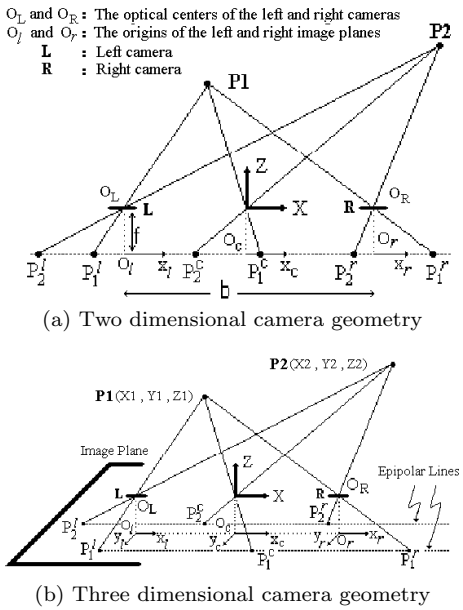
Figure 1 shows a cameras geometry for stereo vision where the cameras optical axes are parallel to each other and perpendicular to the baseline connecting the two cameras L and R. For a point $\mathbf{P}(X, Y, Z)$ in 3D scene, its projections onto the left image and the right image are $p^l(x^l, y^l)$ and $p^r(x^r, y^r)$. Because of this simple camera geometry $y^l = y^r$ and the disparity d is inversely proportional to the depth Z:

$$d = x^l - x^r = b.f/Z \tag{1}$$

Where f is focal length of camera lens and b is separation of two cameras or baseline [1]. Given the two points $\mathbf{P}_1(X_1, Y_1, Z_1)$ and $\mathbf{P}_2(X_2, Y_2, Z_2)$, there are two different definitions for the directional derivative of disparity (DDD). The first one which is simpler than the other one is defined as the difference in disparities divided by the distances between two points in the left image (p_1^l, p_2^l). Therefore, δd can be written as [19]:

$$\delta d = |d_2 - d_1|/\|p_2^l - p_1^l\| \tag{2}$$

Where $\|.\|$ denotes vector norm. The second definition, which is denoted by δd_c here, is defined as the difference in disparities divided by the cyclopean separation,



(a) Two dimensional camera geometry

(b) Three dimensional camera geometry

Fig. 1 Defining the directional derivative of disparity (DDD) in the stereo system with parallel cameras. In figure (a), only the xz plane is shown and figure (b) shows the three dimensions.

where cyclopean separation is a average distance between (p_1^l, p_2^l) and (p_1^r, p_2^r). Suppose a virtual camera is placed in the middle of the cameras L and R, i.e., at the position of origin (O_c in Fig. 1). Therefore we have:

$$\begin{aligned} &d_2 = x_2^l - x_2^r, \qquad d_1 = x_1^l - x_1^r \\ &p_2^c = (p_2^l + p_2^r)/2\ ,\ p_1^c = (p_1^l + p_1^r)/2 \end{aligned} \tag{3}$$

Considering these parameters, we can define δd_c as [2], [30]:

$$\delta d_c = |d_2 - d_1|/\|p_2^c - p_1^c\| \tag{4}$$

Note that from the definitions, DDD is always a non-negative number.

2.2 Applications of DDD in Stereo Correspondence

Bult and Julesz [19] provided evidence supporting the claim that, for binocular fusion of random dot stereograms by human visual system, δd must not exceed the unity. Pollard, Mayhew and Frisby [23] suggested that for most natural scene surfaces, including jagged one, this value (δd) between correct matches is usually less than 1. For discarding the ambiguity in the correspondence problem, they impose a threshold on directional derivative of disparity (δd) constraint among candidate matches. Pollard et al. [31] derived intrinsic relationship between the directional derivative of disparity (δd), surface orientation, and the depth in 3D scenes. Moallem and Faez [22] used probability density function of δd to establish some relations between search space and geometric parameters of stereo system. They also used an empirical threshold on δd_c to

restrict the search region for matching edge points [32]–[34]. Li and Hu [30] used δd_c as a basis for a unified cooperative stereo matching. They selected some families of neighborhood support functions based on δd_c.

2.3 Some Properties of DDD

In the previous subsection, there was some discussion about the empirical threshold on DDD. The empirical investigation showed that the threshold on DDD could be considered to be less than 2, admitting a little error. The threshold value of 1.0 [19], [23] for δd and 1.1 [30] and 1.2 [33] for δd_c were already reported. We want to restrict the search region for matching the edge points, based on the DDD threshold. The definition of δd is simpler but it has no upper limit. In fact δd can be varied theoretically between 0 to ∞. Considering the ordering constraint in stereo vision, it can be shown that δd_c has a bounded value. In fact, $\delta d_c > 2$ can be interpreted as a violation of ordering constraint [30], [33]. We like to show that in a normal condition of a stereo system like human vision (in which the baseline is about 5.5 cm and the depth is varied between 20 cm to 450 cm), the condition of $\delta d_c < 1.2$ results in accepting an error less than 2%. Therefore we used this condition to restrict the search region. On the other hand, the upper limit of 2 for δd_c is sometimes used in our suggestion to guarantee a valid full search region.

Now, we are looking for a 3D surfaces which yield $\delta d_c = 1.2$. Given two points $\mathbf{P}_1(X_1, Y_1, Z_1)$ and $\mathbf{P}_2(X_2, Y_2, Z_2)$ in the 3D coordinates, the value of δd_c can be described as [30]:

$$\delta d_c = \frac{b.|Z_2 - Z_1|}{\sqrt{(Z_1X_2 - Z_2X_1)^2 - (Z_1Y_2 - Z_2Y_1)^2}} \quad (5)$$

When $Y_1 = Y_2 = 0$, Eq. (5) reduces to:

$$\delta d_c = \frac{b.|Z_2 - Z_1|}{|Z_1X_2 - Z_2X_1|} \approx \frac{b}{Z_1}|\tan\theta_c| \quad (6)$$

Where θ_c is the slope with respect to X-axis and $\tan\theta_c$ is the slope of the small surface patch connecting $\mathbf{P}_1$ and $\mathbf{P}_2$. In fact, $\tan\theta_c$ can be varied between $-\infty$ to $+\infty$, thus $|\theta_c|$ is between 0° and 90°. We shall examine a simple situation in which $Y_1 = Y_2 = 0$. If $\delta d_c = 1.2$, $b = 5.5$ cm and Z_1=55 cm, then $|\tan\theta_c| = 12.0$, which indicate that the small surface patch or $|\theta_c|$ is 85.2°, a surface almost along the line of sight. Now, suppose that Z_1 is varied between 20 cm to 450 cm, by integrating θ_c over Z_1, the condition $\delta d_c = 1.2$ causes the mean of $|\theta_c|$ become 88.1°. Consider $|\theta_c|$ has a uniform distribution between 0° and 90°, so in this example the error is less than 2% according to $\delta d_c \leq 1.2$. Increasing the maximum of Z_1 means that $|\theta_c|$ approaches to 90° therefore the error reduces. On the other hand, in our computations the limit on δd_c is considered between 1.2 and 2, so the corresponding error becomes less.

The value of δd_c can be used to define various stereo matching constraints, which were explained in Sect. 1.1. A brief summary follows [30], [32], [33]:

- $\delta d_c > 2$ - Violation of non-reversal order constraint.
- $\delta d_c = 2$ - Violation of uniqueness constraint.
- $\delta d_c < 1.1$ or 1.2 - Empirical limit of DDD, the directional derivative of disparity.
- $\delta d_c \ll 1$ - Figural continuity constraint.

In the next subsection, we discuss the restriction of the search region for matching edge points, based on δd_c empirical limit ($\delta d_c < 1.2$) and its upper real limit ($\delta d_c < 2$) [32], [33], and then propose a fast stereo matching strategy [33], [34]. Then we combine this strategy with hierarchical multiresolution techniques to improve the performances and obtain some fast stereo matching algorithms.

3. Restriction on the Search Region through the DDD Limit

In this subsection, we are going to restrict the search region based on the δd_c limit, for edge based stereo matching. Therefore, we obtain the relationship between δd_c and the search space in stereo correspondence. If we substitute the relation 3 into Eq. (4), we have:

$$\delta d_c = \frac{2|(x_2^l - x_2^r) - (x_1^l - x_1^r)|}{\| (p_2^l + p_2^r) - (p_1^l + p_1^r) \|} \quad (7)$$

We can define Δx_l and Δx_r as the difference between the x position of p_2 and p_1 in left and right images. On the other hand, we obtain:

$$\Delta x_l = x_2^l - x_1^l \ , \ \Delta x_r = x_2^r - x_1^r \quad (8)$$

So δd_c can be changed to:

$$\delta d_c = \frac{2|\Delta x_l - \Delta x_r|}{\sqrt{(\Delta x_l + \Delta x_r)^2 + (\Delta y_l + \Delta y_r)^2}} \quad (9)$$

Suppose we want to find the correspondence of two points in successive scan lines, therefore we have $\Delta y_l = \Delta y_r = 1$, so we have:

$$\delta d_c = \frac{2|\Delta x_l - \Delta x_r|}{\sqrt{(\Delta x_l + \Delta x_r)^2 + (2)^2}} \quad (10)$$

As we discussed before, in a typical stereo system like human vision, a reasonable limit of δd_c is about 1.2, so we assume that $\delta d_c < 1.2$. By substituting in Eq. (10), we can solve the resulting non-equality:

$$0.64\Delta x_l^2 + 0.64\Delta x_r^2 - 2.72\Delta x_l \Delta x_r - 1.44 \leq 0 \quad (11)$$

Suppose the feature points in left image are connected non-horizontal edge points in subsequent scan lines. If we know the value of Δx_l in the left image, we can restrict Δx_r in the right image. Five cases are investigated: $\Delta x_l = 0$, $\Delta x_l = \pm 1$ and $\Delta x_l = \pm 2$. For example in case of $\Delta x_l = 0$ we have, $-1.5 \leq \Delta x_r 1.5 \leq$. This means that for two connected edge points in the subsequent scan lines of left image which have the same

Table 1 Relation between Δx_l,Δx_r and extended range of δd_c.

Δx_l	Rounded range of Δx_r	extended range of δd_c
-2	$-9,-8,-7,-6,-5,-4,-3,-2,-1,0$	$-1.41 < \delta d_c < 1.25$
-1	$-5,-4,-3,-2,-1,0,1$	$-2 < \delta d_c < 1.26$
0	$-2,-1,0,1,2$	$-1.41 < \delta d_c < 1.41$
1	$-1,0,1,2,3,4,5$	$-1.26 < \delta d_c < 2$
2	0,1,2,3,4,5,6,7,8,9	$-1.25 < \delta d_c < 1.41$

position in x direction, by context of $\delta d_c < 1.2$, the maximum allowable range of Δx_r will be 1.5 pixels in the right image.

Table 1 shows the relationship between Δx_l and Δx_r for the above five cases. Since the number of pixels can not have fractional part, we rounded Δx_r and then we have extended the range of δd_c. Third column of Table 1 shows the extended value of δd_c in both negative and positive limits. In all cases, the condition of $\delta d_c < 1.2$ is met and even in some cases (negative limit for $\Delta x_l = -1$ and positive limit for $\Delta x_l = +1$), the maximum allowable range of δd_c for non-reversal ordering constraint is achieved. In the next section, we propose a new fast matching strategy based on this restriction in the search region.

4. Fast Edge-Based Stereo Matching Strategy

In a calibrated stereo system with parallel optical axes, the area-based or the feature-based algorithms consist of two stages: feature point (or primitives) extraction and stereo matching [1]. Most of fast stereo algorithms use low-level primitives like edges, those that do not require sophisticated semantic analysis in their extraction [35]. Matching the horizontal edges in a stereo system with parallel optical axes is a problem [36], so some authors suggested non-horizontal edge points as feature points [14], [21], [32]–[34], [37]. In a correlation based framework, stereo matching for a pixel in a reference image (left) is obtained by searching in a predefined region of the second image (right).

Our stereo algorithms have two stages, *feature extraction* and *feature matching*. In the feature extraction stage, non-horizontal thinned edge points are extracted from the both of left and right image. The extracted edge points are first classified into two groups, positive and negative, depending on the graylevel difference between the two side of the edge point in x direction. The feature extraction stage will be explained accurately later. To explain the matching strategy of the extracted edge points, we would define a temporary concept which, is related to the extracted edge points. We call it successive connected edge points set or abbreviately SCE.

4.1 Successive Connected Edge Points

Some connected edge points in a left image can be grouped together as some sets that we call *successive connected edges points* (SCE) sets. Each SCE set ψ consists of n successive edge points:

$$\psi = \{p_1^l, p_2^l, \ldots, p_n^l\} \tag{12}$$

Its coordinates and type identify each edge point like p_k^l

$$p_k^l = (x_k^l, y_k^l, type) \tag{13}$$

The set ψ is said a *successive connected edges* set if these three conditions are met:
(1) The types of all p_k^l are the same.
(2) The successive p_k^l's are in successive scans lines. On the other hand we have:

$$y_{k+1}^l - y_k^l = 1 \tag{14}$$

(3) The absolute difference between x values of two successive points is less than 3 pixels, so we have:

$$x_{k+1}^l - x_k^l = \{-2,-1,0,+1,+2\} \tag{15}$$

Two samples of *successive connected edges* with 38 points are shown in Figs. 3 (a) and 3 (b).

4.2 Search Space in the Edge Matching

Suppose that we want to find correspondences of edge points of a sample SCE set in a left image. Considering the limit of δd_c and the connectivity of the edge points in a SCE set, if a value of Δx_l is known, Δx_r range can be computed directly by Table 1.

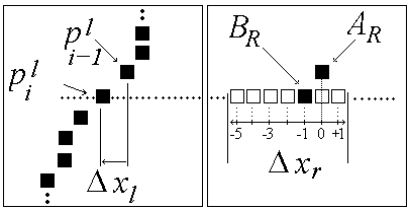
For example consider p_{i-1}^l and p_i^l are two successive edge points in a sample SCE set and A_R is the correspondence of p_{i-1}^l in the right image (see Figs. 2 (a) and 2 (b)). Considering Eq. (8), Δx_l is the difference between the x position of p_i^l and p_{i-1}^l in the left ($\Delta x_l = x_i^l - x_{i-1}^l$). For finding B_R (the correspondence of p_i^l) in the right image, the search space region or Δx_r range can be computed by Table 1. This search region in the right image (X_{Range}^r) is shown in Fig. 2 (b). Considering $d = x_{i-1}^l - x_A^r$ and $\Delta x_r = x_A^r - X_{Range}^r$, X_{Range}^r as the search space in the right image coordinate system, can be computed as:

$$X_{Range}^r = \Delta x_r + x_{i-1}^l - d \tag{16}$$

In Fig. 2, Δx_l is considered to be -1 so the Δx_r range is $\{-5,-4,-3,-2,-1,0,+1\}$, therefore the search region is restricted to 7 pixels only. The search region is also shown in Fig. 2 (b).

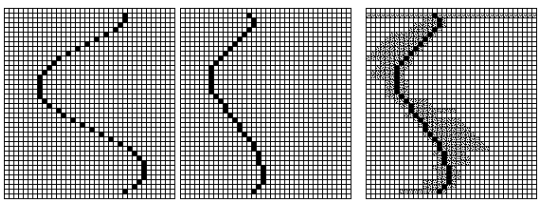
4.3 SCE Set Construction and Matching

In a process of the edge matching, the edge points of the left image are divided into two groups, checked and unchecked edge points. The checked edges are those edge points that were already examined in the process and the unchecked edges are those that are not examined yet. In the following edge matching strategy, the



(a) A part of SCE (b) Search region for B_R

Fig. 2 An example for the search space in the edge matching. Figure (a) is part of a sample SCE set in the left image, p^l_{i-1} and p^l_i are two successive edge points and Δx_l is the difference between their x positions. In figure (b), the search region for finding B_R is shown as the Δx_r range, and the epipolar line for B_R is also shown with dotted line.



(a) Left (b) Right (c) Restricted region

Fig. 3 An example of search space in the proposed matching strategy, the edge points are shown black. Figure (a) is the SCE set in the left image, (b) is the correspondences of the left edge points in the right image based on the epipolar line and finally (c) is the restricted search region in the right image. The search region is also marked gray.

SCE set constructing for edge points of left image and establishing correspondence are implemented concurrently. At first, all of edge points in a left image are marked as unchecked. The proposed matching strategy can be expressed briefly in two phases as follows:

• Phase I: Find the next unchecked edge in the left image by systematic scan, from the left to right and from the top to bottom. The found edge is considered as p^l_1, the first point of the corresponding SCE set. Consider X^r_{Range} in the right image coordinate as a full search region based on epipolar line and other constraints. Set $i = 1$ and go to phase II.

• Phase II:

(1) Find the correspondence of p^l_i in the right image, B_R (see Fig. 2 (b)) based on X^r_{Range} and mark p^l_i as checked edges. If B_R is not found, go to phase I.

(2) Compute $d = x^l_i - x^r_B$ (Eq. (3)).

(3) Find p^l_{i+1}, the next point of p^l_i in the corresponding SCE set based on the three conditions of SCE sets in Sect. 4.1 and go to next step. If p^l_{i+1} is not found, go to phase I.

(4) Compute $\Delta x_l = x^l_{i+1} - x^l_i$ (Eq. (8)).

(5) Compute Δx_r range based on Table 1.

(6) Compute $X^r_{Range} = \Delta x_r + x^l_{i-1} - d$ (Eq. (16)).

(7) Set $i = i + 1$ and got to the first step.

4.4 An Example

One sample SCE set with 38 points in a 40×36 image is shown in Fig. 3 (a) as the left image. Considering the epipolar line, Fig. 3 (b) (as the right image) shows the correspondence of each edge points of Fig. 3 (a). Considering the proposed matching strategy, the search space in the right image for the matching of the left SCE set is also shown in Fig. 3 (c). For the first point in the SCE set, the full search space is considered on the epipolar line, but for other points, the search spaces are restricted based on the above discussion. Effect of reduction in a search space is more evident for larger image dimensions.

In the next section, we propose some fast edge-based stereo matching algorithms based on this new and fast matching strategy and compare the results from point of view of the execution time and accuracy with those of other fast edge-based stereo matching algorithms.

5. Fast Edge-Based Stereo Matching Algorithms

The matching primitives, the strategy and the matching criterion are three essential parameters in the fast edge-based stereo matching algorithms. We consider these parameters in our proposed algorithms. Most of fast stereo algorithms use non-horizontal edge points as features or matching primitives. As we mentioned before, we use SCE sets in our fast stereo algorithms.

5.1 Detection of Thinned Non-horizontal Edge Points

At the first, non-horizontal edge extraction is performed and thinned edge points are classified into two groups, positive and negative, depending on the graylevel difference between the two sides of the edge points in x direction. Left and right images are convolved with a proper gradient mask. To detect non-horizontal edge points, we use a gradient of two-dimensional Gussian function with $\sigma = 1$ in x direction as the gradient mask. This mask is truncated to 5×5. A non-horizontal thinned positive edge in a left image, is localized to a pixel that the filter response or $\rho_l(x, y)$ has to exceed a positive threshold $\rho^+_\circ$ and has to obtain a local maximum in x direction, therefore [35]:

$$\begin{array}{ll} \rho_l(x,y) > \rho^+_\circ & \leftarrow Threshold \\ \rho_l(x,y) > \rho_l(x-1,y) \\ \rho_l(x,y) > \rho_l(x+1,y) & \leftarrow LocalMaximum \end{array} \quad (17)$$

The positive threshold $\rho^+_\circ$ is considered pre-defined or can be selected based on the filter response dynamically. Consider ρ^+_{mean} is the mean of positive values of the filter response, we choose $\rho^+_\circ = 1.5\rho^+_{mean}$, the coefficient of 1.5 is selected empirically.

The extraction of non-horizontal negative thinned edge points is similar to positive one. Consider ρ^-_{mean} is the mean of negative values of the filter response,

the negative threshold value is $\rho_{\circ}^{-} = 1.5\rho_{mean}^{-}$. Non-horizontal thinned negative edge points in a left image can be defined as points with these properties:

$$\begin{array}{ll} \rho_l(x,y) < \rho_{\circ}^{-} & \leftarrow Threshold \\ \rho_l(x,y) < \rho_l(x-1,y) \quad \rceil & \\ \rho_l(x,y) < \rho_l(x+1,y) \quad \rfloor & \leftarrow LocalMinimum \end{array} \tag{18}$$

5.2 The Matching Strategy

Search space reduction is very important to decrease processing time of matching. In the proposed SCE sets construction and matching strategy, from the second point up to end point of SEC sets, we could effectively reduce the search space for establishing the correspondence. On the other hand, considering a SEC set in a left image as $\psi = \{p_1^l, p_2^l, \ldots, p_n^l\}$, the search space can be reduced for the points of p_i^l $(i = 2, 3, \ldots, n)$ but for the first point that is found in the first phase of proposed matching strategy, i.e. p_1^l, we use the hierarchical multiresolution matching one to reduce the search space.

The hierarchical multiresolution matching strategy is a powerful tool to reduce the search space. This matching strategy belongs to a coarse to fine strategy. In this approach, several levels of image pyramid are made. This technique can be used to reduce a full search space to a limited region. Generally, stereo images are reduced by a convolution with a proper Gussian filter and subsampled by factor of two along each axis, to form a pyramid. Each pair of slices from this stereo pyramid is separately processed starting with the coarsest (most reduced) pair and produced the disparity maps. At each level, the disparity map of previous level is used to reduce the search space [24].

In the edge-based stereo, edge extraction is usually done at the finest level only. Then a series of coarser resolution edge's images is generated by series of reduction in size of finest edge images by factor of two. To establish the correspondence at the coarsest level, the full search space is usually selected. To decrease the processing time of matching in the fast edge-based stereo, for one edge point in the first image, only similar edge points in the other image are usually tested. However, non-edge points can be examined after testing the edge point. In this case, the matching rate increases but the execution time is decreased too.

In our stereo matching algorithms, the search spaces are reduced by both of the hierarchical multiresolution algorithm (for the first point of SCE sets) and our restricted search region one (for other points of SCE sets). We claim that our proposed restricted search region algorithm is more powerful than hierarchical multiresolution one. Therefore we should try to use our proposed restricted search region algorithm more oftenly than the hierarchical multiresolution one. In the edge-based stereo, as we mentioned before, usually the edge pixels are examined in the search space, but in our proposed edge-based algorithms, we use a different strategy. At first, we examine only the similar edge pixels, but if we can not find the correspondence in the search space, we try other non-edge pixels in the search space. Therefore, we increase probability of the matching for the second point, up to end points of SCE sets. For each edge that the correspondence is established, we update the disparity map in coarse and medium level, beside the fine level. This can help us to establish correspondence of the first point of SCE set faster.

5.3 The Matching Criteria

Most currently used fast stereo methods belong to the category of linear correlation methods, which include those based on sum of squared differences (SSD) and normalized cross correlation (NCC) [38]. In a real stereo system in which left and right cameras are different, NCC is preferable since it is invariant to linear brightness and contrast variations between perfect matching windows. The value of NCC is between -1 and $+1$, and a larger value indicates more similarity between windows [16]. The window size in the correlation-based methods is very important. As a window size decreases, discriminatory power of window based criterion is decreased and some local maximums in NCC could have been found in search region. Moreover, continually increasing the window size causes the performance to be degraded because of the occlusion regions and smoothing of disparity values across the depth boundary [16]. If a disparity search range could be automatically reduced to an effective range (about 10 pixels) then several local maximums in NCC would be eliminated form the selection process and therefore, the disparities found would be correct, even if the size of the matching block is small [27].

In our proposed algorithms, for the first point of SCE sets, the hierarchical multiresolution matching strategy is used. Therefore the window sizes of NCC are selected corresponding to the image size of the pyramid level. Since we could effectively reduce the search space for other points of SCE sets, we use NCC with window size of 3×3 as similarity measure.

6. Implementation and Results

For the first point of a SCE set, we use hierarchical multiresolution algorithm and for other points of that SCE set, the restricted search region method based on our matching strategy is used. For the first point, we examine the two approaches that we called HMEO and HMNE (HM means hierarchical multiresolution). In HMEO, only edge points are examined in each pyramid level but in HMNE, if the correspondence can not be found by examining the edge pixels, non-edge points are also tried. For the non-first points of SCE sets, as

we mentioned before, we use NCC with window size of 3×3 and, the threshold value of 0.7. Therefore we have a method for the non-first points that we called it RS (RS means restricted search region). Hence we have proposed two algorithms: RS-MHEO and RS-HMNE.

We compare results of our algorithms with those of fast edge-based, hierarchical multiresolution methods, HMEO and HMNE. To reduce the execution time in these methods, NCC with window size of 5×5 are used for the coarsest level and 3×3 for the other levels. Moreover, the NCC threshold value of 0.7 is selected for all pyramid levels. For comparing the results of the execution time, we chose a reference method that we called FSEO (Full Search, Edge only). This feature-based method finds the correspondence of non-horizontal edge points of left image by examining all of the similar edge points in the epipolar line of right image. As the similarity criterion, NCC with window size of 11×11 and the threshold value of 0.8 were selected.

These algorithms were tested on four different stereo scenes, House, Renault, Ball and Pentagon [39]. All the scenes used are gray levels 256×256 and their disparity ranges are shown in Table 2. To determine the error in the matching stage of the algorithms, we compute the accurate disparity map for the non-horizontal edge points of the left images by searching in the predefined region of the right images based on Table 2 and NCC with window size of 15×15. The errors in these computed disparity maps are corrected manually.

The codes of the algorithms were written by Watcom C and implemented by a PC under Windows operating system, with a Pentium II 450 MHz processor. As we mentioned before, our stereo algorithms have two stages, *feature extraction* and *feature matching*. Therefore at first the results of implementation of these two stages are investigated separately and then the execution times of the whole algorithms are introduced.

6.1 Feature Extraction

The edge point extraction stage of all algorithms is the same. Table 3 shows the results of the edge point extraction stage on the stereo scenes. In this table, the number of extracted edge points including positive and negative type (p for positive and n for negative) in the left and right images and the corresponding execution times (in milliseconds) are also shown.

The extracted edge points for House, Renault, Ball and Pentagon stereo scenes are shown in Figs. 4 to 8 respectively (part (c) and (d)). In these figures the red and blue colors are used to show the positive and negative edge points respectively.

6.2 Feature Matching

For House scene, the results of the investigated algorithms are shown in Table 4. Considering the total number of extracted features from the left image that are shown in Table 3, the number of the matched edges, failed edges and error in matching are shown in the matched, failed and error columns respectively. The matched and failed columns in Table 4 show the number of features which are matched and failed in the edge matching process. Therefore, the sum of these two columns should be the total number of extracted features from the left image. The percentages of matched and failed features, with respects to the total number of the left features are also shown in the parentheses. Considering the accurate disparity map for the feature points, the error of the matching stage can be computed. The error column shows this error and, the percentages of the error with respect to the total extracted left features are shown in the parenthesis too. ET (or execution time) column is in millisecond scale. Tables 5 through 7 show the results of the matching and the execution times of the algorithms on Renault, Ball and Pentagon stereo scenes, respectively.

Figure 4 shows the implementation results on House stereo scene. This figure shows the left and right images ((a) and (b)), the extracted features including non-horizontal edge points((c) and (d)), negative and positive type edges and the disparity maps obtained by the algorithms ((e) through (i)). In the disparity map figures ((e) through (i)), the failed edges and the error in the matching are shown with magenta and black color respectively. Figures 5 through 7 show the corresponding disparity maps and other results those computed by the algorithms, on Renault, Ball and Pentagon stereo scenes respectively.

Table 2 Disparity ranges of the tested scenes.

Disparity	House	Renault	Ball	Pentagon
Minimum	+35	0	−30	−10
Maximum	+60	+10	0	+10

Table 3 Implementation results of the edge point extraction on the stereo scenes, in last column, ET means execution time.

Scene	Left Edges(p,n)	Right Edges(p,n)	ET(msec.)
House	734(385,349)	731(343,388)	81.6
Ren.	3012(1466,1546)	3058(1462,1596)	84.4
Ball	3925(1903,2022)	4005(1939,2066)	84.6
Pent.	6459(3321,3138)	6690(3482,3208)	87.6

Table 4 Implementation results on House scene. 734 non-horizontal edge points are extracted from the left image. ET column shows execution time in milliseconds.

Algorithm	Matched(%)	Failed(%)	Error(%)	ET
FSEO	729(99.3%)	5(0.7%)	47(6.4%)	35.4
HMEO	696(94.8%)	38(5.2%)	46(6.3%)	10.2
HMNE	728(99.2%)	6(0.8%)	48(6.5%)	11.0
RS-HMEO	726(98.9%)	8(1.1%)	0(0.0%)	5.5
RS-HMNE	732(99.7%)	2(0.3%)	1 (0.2%)	5.8

6.3 Execution Time of the Algorithms

In a complete matching algorithm, processing time includes feature extraction time plus matching time. Table 8 shows the execution time of the algorithms including the execution time of the edge extraction and edge matching shown separately in the parentheses.

Figure 8 shows the execution time of each algorithm on the tested stereo scenes. In this chart, the execution time of edge extraction and matching are separately shown with the same scale. Therefore we can compare the execution time of these stages to each other. It can be seen that the edge extraction time is nearly constant but the edge matching time is very much dependent on the scene. This phenomenon has a great effect on the performances of our proposed algorithms. On the other hand, the execution time of the edge matching in the proposed algorithms is highly reduced with respect to the others (see Table 8 for more details).

Table 5 The implementation results on Renault scene. 3012 non-horizontal edge points are extracted from the left image.

Algorithm	Matched(%)	Failed(%)	Error(%)	ET
FSEO	2729(90.6%)	283(9.4%)	60(2.0%)	424.4
HMEO	2537(84.2%)	475(15.8%)	155(5.2%)	60.4
HMNE	2844(94.4%)	168(5.6%)	172(5.7%)	79.2
RS-HMEO	2573(85.4%)	439(14.6%)	118(3.9%)	33.0
RS-HMNE	2903(96.4%)	109(4.6%)	158(5.3%)	36.2

Table 6 The implementation results on Ball scene. 3925 non-horizontal edge points are extracted from the left image.

Algorithm	Matched(%)	Failed(%)	Error(%)	ET
FSEO	2879(72.6%)	1046(26.6%)	721(18.4%)	587.6
HMEO	2725(69.4%)	1200(30.6%)	743(18.9%)	99.0
HMNE	3041(77.5%)	884(22.5%)	800(20.4%)	146.0
RS-HMEO	2575(65.6%)	1351(34.4%)	627(16.0%)	71.4
RS-HMNE	3061(78.0%)	867(22.0%)	786(20.0%)	100.2

Table 7 The implementation results on Pentagon scene. 6459 non-horizontal edge points are extracted from the left image.

Algorithm	Matched (%)	Failed (%)	Error (%)	ET
FSEO	3250(50.3%)	3209(49.7%)	48(0.7%)	1177.6
HMEO	4379(67.8%)	2080(32.2%)	156(2.4%)	161.6
HMNE	5646(87.4%)	814(12.6%)	255(3.9%)	288.8
RS-HMEO	4310(66.7%)	2149(33.3%)	233(3.6%)	95.6
RS-HMNE	5575(86.3%)	884(13.7%)	314(4.9%)	140.6

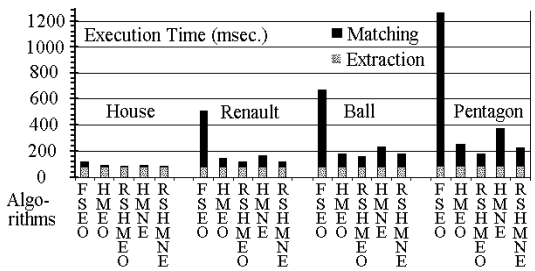


Fig. 8 The execution times of five considered algorithms on the tested stereo scenes are shown, the edge extraction and edge matching time are distinguished by gray and black columns respectively. The edge extraction time is nearly independent of the scenes, but the edge matching time is very dependent to the scenes. For FSEO algorithm, this time is varied between 117.0 to 1264.2 milliseconds. In all cases, execution time of RS-HMEO and RS-HMNE are less than HMEO and HMNE.

7. Discussion

The execution time and the failed rate are two essential parameters in the proposed algorithms. Therefore we first discuss these parameters and then investigate some effective factors in decreasing the processing time.

7.1 Speed Up and Failed Rate

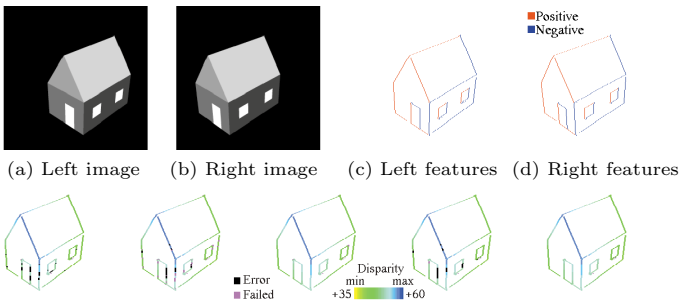
Considering FSEO algorithm as a reference one for the measuring execution time, speed up of the other algorithms could be computed and it is shown in Table 9. For all stereo scenes, RS-HMNE and RS-HMEO are faster than HMNE and HMEO. In this table, the speed up increasing of RS-HMEO and RS-HMNE with respect to HMNE and HMEO are also shown in RS-HMEO and RS-HMNE columns respectively. In the tested cases, the speed up is increased between 5.5% to 64.9%. The minimum increasing of the speed up is 5.5% for RS-HMEO in House scene, whiles the maximum one is 64.9% for RS-HMNE in Pentagon scene. In all cases, the speed up increasing in RS-HMNE is more than RS-HMEO.

Table 9 Comparing the speed up of the algorithms, the values in parentheses in RS-HMEO and RS-HMNE columns are the increasing of speed up respect to HMEO and HMNE respectively.

Scene	HMEO	RS-HMEO	HMNE	RS-HMNE
House	1.27	1.34(5.5%)	1.26	1.34(6.4%)
Renault	3.51	4.33(23.4%)	3.11	4.21(35.4%)
Ball	3.66	4.31(17.8%)	2.92	3.64(24.6%)
Pentagon	5.07	6.90(36.1%)	3.36	5.54(64.9%)

Table 8 Execution time of the algorithms in milliseconds, the first item in parentheses is the edge matching time and the second one is the edge extraction time.

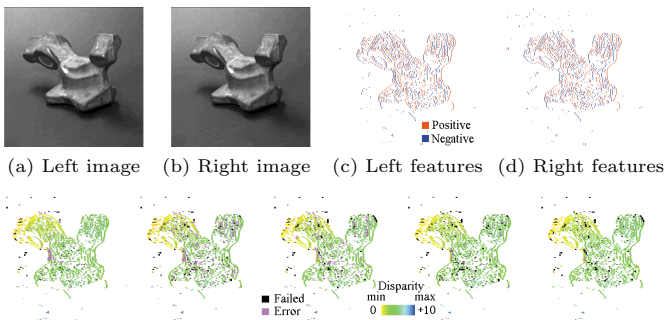
Algorithm	House	Renault	Ball	Pentagon
FSEO	117.0 (35.4 + 81.6)	508.8 (424.4 + 84.4)	672.2 (587.6 + 84.6)	1264.2 (1176.6 + 87.6)
HMEO	91.8 (10.2 + 81.6)	144.8 (60.4 + 84.4)	183.6 (99.0 + 84.6)	249.2 (161.6 + 87.6)
RS-HMEO	87.1 (5.5 + 81.6)	117.4 (33.0 + 84.4)	156.0 (71.4 + 84.6)	183.2 (95.6 + 87.6)
HMNE	92.6 (11.0 + 81.6)	163.3 (79.2 + 84.4)	230.6 (146.0 + 84.6)	376.4 (288.8 + 87.6)
RS-HMNE	87.4 (5.8 + 81.6)	120.6 (36.2 + 84.4)	184.8 (100.2 + 84.6)	228.2 (140.6 + 87.6)



(a) Left image (b) Right image (c) Left features (d) Right features

(e) FSOE (f) HMEO (g) RS-HMEO (h) HMNE (i) RS-HMNE

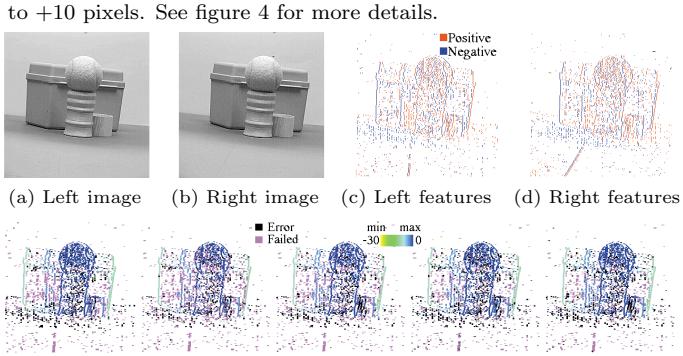
Fig. 4 Left image disparity maps obtained by the proposed algorithms on House stereo scene. (a) And (b) show left and right image, (c) and (d) show the non-horizontal edge points extracted as features for the left and right image. (e) through (i) show disparity maps computed by the algorithms, and failed edges and error in the matching are also shown with magenta and black color respectively. The disparity range is between +35 to +60 pixels.



(a) Left image (b) Right image (c) Left features (d) Right features

(e) FSOE (f) HMEO (g) RS-HMEO (h) HMNE (i) RS-HMNE

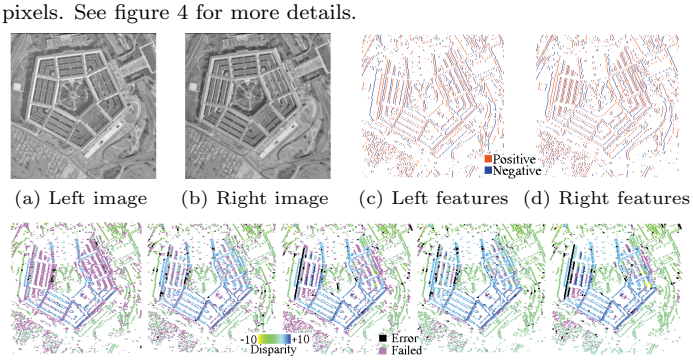
Fig. 5 The implementation results on Renault stereo scene. The disparity range is between 0 to +10 pixels. See figure 4 for more details.



(a) Left image (b) Right image (c) Left features (d) Right features

(e) FSOE (f) HMEO (g) RS-HMEO (h) HMNE (i) RS-HMNE

Fig. 6 The implementation results on Ball stereo scene. The disparity range is between −30 to 0 pixels. See figure 4 for more details.



(a) Left image (b) Right image (c) Left features (d) Right features

(e) FSOE (f) HMEO (g) RS-HMEO (h) HMNE (i) RS-HMNE

Fig. 7 The implementation results on Pentagon stereo scene. The disparity range is between −10 to +10 pixels. See Fig. 4 for more details.

Table 11 Some details of RS-HMNE algorithm results. The first point and non-first points columns mean the number of first point and non-first points in the SCE sets. N/F column means the ratio of non-first points to first point of SCE sets which matched.

Scene	Total Edges	Edge Density	Total Matched	First Point Total (Matched - Failed)	Non-first Points Total (Matched - Failed)	N/F that Matched	Decreasing of Matching Time
House	734	1.1%	732	39 (37 - 2)	695 (695 - 0)	18.8	47.3%
Renault	3012	4.6%	2903	696 (626 - 70)	2316 (2277 - 39)	3.64	54.3%
Ball	3925	6.0%	3061	1632 (931 - 701)	2293 (2130 - 163)	2.29	31.4%
Pentagon	6459	9.9%	5575	1855 (1000 - 855)	4606 (4508 - 96)	4.51	51.3%

Table 10 Comparing the failed rate of the algorithms.

Scene	HMEO	RS-HMEO	HMNE	RS-HMNE
House	5.2%	0.3%	0.8%	0.8%
Renault	15.8%	14.6%	5.6%	4.6%
Ball	30.6%	34.4%	22.5%	22.0%
Pentagon	32.2%	33.3%	12.6%	13.7%

Table 10 shows the failed rate of the algorithms. The failed rates in Ball and Pentagon scenes are more than others. In Ball scene, a large number of features in the left image are in a uniform region. On the other hand, the occlusion is nearly high in Pentagon scene. In our proposed algorithms, the failed rate are occasionally increased or decreased. Our algorithms decrease the failed rates in House and Renault and nearly in Ball scenes, but the failed rate is increased a little in Pentagon. House scene is a simple case while Pentagon is very complex one and the occlusions are more than the others. Renault and Ball scenes are nearly in the middle of these two scenes. Therefore in moderate complex scenes, our algorithms can decrease both the processing time and the failed rate, while the processing time is certainly decreased even in complex scenes in which the failed rate is nearly remains constant.

7.2 Important Factors in the Processing Time

In a complete matching algorithm, the processing time includes the feature extraction time plus the matching time. As we mentioned before, conventionally, for the non-first points of SCE sets, our proposed matching strategy works faster than for the first point. Therefore, when the number of first point of SCE sets are decreased relative to non-first points, we expect that the processing time should be lower relatively. Table 11 shows some details of matching stage for RS-HMNE algorithm. In this table, the results of matching for the first and non-first point are shown separately. N/F column is the ratio of non-first points to first point of SCE sets which are matched. Moreover the percentage of decreasing in the processing time (relative to HMNE) are also shown in the last column. In House scene, in which the matching is established more than 18 times for non-first points, the decreasing in the processing time is not very high (47.3%). The density of edge points is a very important factor in the matching time. HMEO and HMNE works very well when the density of edge points is low, but our algorithms have a different behavior. In our algorithms, the complexity of the matching for non-first point of SEC sets is nearly independent of the density of edge points, because the search space is narrowly restricted. Therefore the difference between the matching time of hierarchical multiresolution algorithms and ours will be higher, when the density of edge points is increased. Hence our algorithms work better in real scenes relative to synthetic scenes, for example comparing the results of HMNE and RS-HMNE in Pentagon, Renault and Ball scenes.

Considering the results of RS-HMNE for Renault and Ball scenes in Table 11, RS-HMNE in Renault scene decreases the matching time more than in Ball scene. In Renault scene N/F ratio is more than Ball, hence non-first points are relatively more than first point. Therefore, it can be concluded that the matching time is dependent mostly on N/F ratio. On the other hand, the connectivity of edge points is an important factor. If the number of edge points in each SCE set is increased, the matching time will be decreased, especially in complex scenes. In Ball scene, the connectivity of edge points is less than Renault (see Figs. 4 (c) and 5 (c)), therefore the decreasing of the matching time in Ball scene is less than Renault.

The other important parameter in the processing time is the feature extraction time. Generally, the feature extraction time is nearly independent of the density of edge points, because the main computation of this stage is the convolution of the images with a predefined proper mask. Table 3 shows that the maximum difference in the execution time is between House and Pentagon scene, which is only 6 milliseconds (less than 8%). Therefore it is clear that the execution time of feature extraction is nearly constant. On the other hand, the execution time of matching is mostly dependent on parameters like density of edge points and connectivity of edge points. Considering RS-HMNE algorithm, this time is varied between 5.8 milliseconds (House) to 140.6 milliseconds (Pentagon). Therefore in the scenes like House in which the density of edge points is very low, the execution time of the algorithms are dominant with feature extraction and not with matching. Hence, the privilege of the proposed algorithms is not evident in these cases. On the other hand, when the density of edge points is high, like Pentagon scene, the execution time of matching is high enough to dominate to fea-

ture extraction. Therefore in these cases, the decreasing times of our proposed algorithms are completely high light.

8. Conclusion

Edge-based stereo matching is a popular method in some fast stereo vision applications. Reducing the search space in the matching can increase the execution speed and accuracy. In this paper we introduced the concept of search space reduction in edge-based stereo by limiting the directional derivative of disparity, δd_c. Afterward, we formulated a search space region in the matching of the successive connected edge points, SCE sets. We could accurately restrict the search region in the matching of SCE sets. Then we posed a fast matching strategy for SCE sets in two phases, based on that restriction. We used the hierarchical multiresolution scheme to reduce further the search space and finally developed two fast edge-based stereo matching algorithms, RS-HMEO and RS-HMNE. The first one is very fast, but its fault rate is a little higher. The second one is both fast and accurate.

Most fast algorithms work better in simple and moderate scenes, but our algorithms have a better performance when the complexity of a real scene is increased. Connectivity of edge points are very important factor to decrease the execution time of our algorithms, so our methods might not work well in a scene like random dot stereogram, but work well in a real scene. In conventional hierarchical multiresolution methods, when the density of edge points is increased, the execution time is increased too, but the execution time of our proposed algorithms are not increased highly, since the search space is independent of the density of edge points for non-first point of SCE sets. Therefore, our proposed algorithms have very good performance in complex real scene where the connectivity of edge points is guaranteed as well as having high density of edge points. Since the proposed algorithms work very well in real scenes, these algorithms are very suitable for real time applications like robotics.

The execution time of edge extraction stage which is implemented separately for a left and right image, is nearly independent of scene complexity. Therefore to decrease the execution time, we suggest using two separate processors for implementation of edge extraction in the left and right image. Another processor can be used in pipeline manner to the others to implement the matching stage. Therefore the mentioned parallel architecture can be used to implement the proposed edge-based stereo algorithms in real time applications.

References

[1] E. Trucco and A. Verri, Introductory Techniques for 3D Computer Vision, Prentice Hall, 1998.

[2] U.R. Dhond and J.K. Aggarwal, "Structure from stereo — A review," IEEE Trans. Syst., Man. & Cybern., vol.19, no.6, pp.1489–1510, Nov.-Dec. 1989.

[3] G.Q. Wei, W.Brauer, and G.Hirzinger, "Intensity- and gradient-based stereo matching using hierarchical Gaussian basis functions," IEEE Trans. Pattern Anal. & Mach. Intell., vol.20, no.11, pp.1143–1151, Nov. 1998.

[4] D. Murry and C. Jennings, "Stereo vision based mapping and navigation for mobile robots," Proc. IEEE Int. Conf. on Robotics and Automation (ICRA'97), vol.3, pp.1694–1699, 1997.

[5] Y. Matsumoto, T. Shibata, K. Sakai, M. Inaba, and H. Inoue, "Real-time color stereo vision system for mobile robot based on field multiplexing," Proc. IEEE Int. Conf. on Robotics and Automation (ICRA'97), vol.3, pp.1934–1939, 1997.

[6] L. Robert, M. Buffa, and M. Hebert, "Weakly-calibrated stereo perception for rover navigation," Proc. IEEE Int. Conf. on Computer Vision, pp.46–51, 1995.

[7] C. Innocenti, G. Mondino, P. Regis, and G. Sandini, "Trajectory planning and real-time control of an autonomous mobile robot equipped with vision and ultrasonic sensors," Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS'94), vol.3, pp.1861–1866, 1994.

[8] S.W. Chan and A.F. Clark, "Periscopic stereo for virtual world creation," Proc. IEE Int. Conf. on Image Processing and Its Applications (IPA'97), vol.1, pp.419–422, 1997.

[9] W.H. Kim and S.W. Ra, "Fast disparity estimation using geometric properties and selective sample decimation for stereoscopic image coding," IEEE Trans. Consumer Electronic, vol.45, no.1, pp.203–209, Feb. 1999.

[10] H. Jahn, "Parallel epipolar stereo matching," Proc. IEEE Int. Conf. on Pattern Recognition (ICPR2000), pp.402-405, 2000.

[11] A. Koschan and V. Rodehorst, "Towards real-tim1e stereo employing parallel algorithms for edge-based and dense stereo matching," Proc. IEEE Int. Workshop on Computer Architectures for Machine Perception, CAMP'95, pp.234–241, 1995.

[12] H.K. Nishihara, "Real-time stereo- and motion-based figure-ground discrimination and tracking using LOG sign-correlation," Proc. IEEE Int. Conf. on Pattern Recognition (ICPR96), pp.95–100, 1996.

[13] I. Masaki, "Three dimensional vision system for intelligent vehicles," Proc. IEEE Int. Conf. on Industrial Electronics (IECON'93), vol.3, pp.1712–1717, 1993.

[14] M. Adjouadi, F. Candocia, and J. Riley, "Exploiting Walsh-based attributes to stereo vision," IEEE Trans. Signal Processing, vol.44, no.2, pp.409–420, Feb. 1996.

[15] Y. Ruicheck and J.G. Postaire, "A neural implementation for high speed processing in linear stereo vision," Proc. IEEE Int. Conf. on Systems, Men and Cybernetics (SMC'95), vol.5, pp.3902–3907, 1995.

[16] D.N. Bhat and S.K. Nayar, "Ordinal measure for image correspondence," IEEE Trans. Pattern Anal. & Mach. Intell., vol.20, no.4, pp.415–423, April 1998.

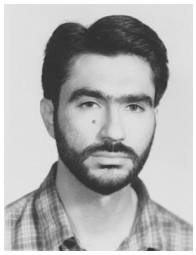
[17] S.B. Marapane and M.M. Trivedi, "Multi-primitive hierarchical (MPH) stereo analysis," IEEE Trans. Pattern Anal. & Mach. Intell., vol.16, no.3, pp.227–240, March 1994.

[18] R. Mohan, G. Medioni, and R. Nevatia, "Stereo error detection, correction and evaluation," IEEE Tans. Pattern Anal. & Mach. Intell., vol.11, no.2, pp.113–120, Feb. 1989.

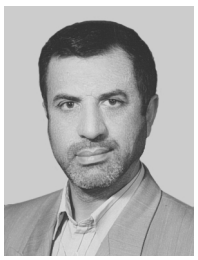
[19] P. Burt and B. Julesz, "A disparity gradient limit for binocular fusion," Science, vol.208, pp.615–617, 1980.

[20] C.V. Stewart, "On the derivation of geometric constraints in stereo," Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition (CVPR92), pp.769–772, 1992.

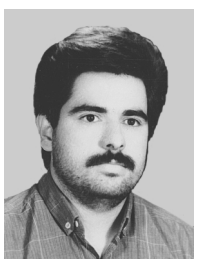
[21] C. Stewart, R. Flatland, and K. Bubna, "Geometric constraints and stereo disparity computation," Int. Jour. of Computer Vision, vol.20, no.4, pp.143–168, Kluwer Academic Publisher, 1996.
[22] P. Moallem, K. Faez, and J. Haddadnia, "Reduction of the search space region in the edge based stereo correspondence," Proc. IEEE Int. Conf. on Image Processing (ICIP2001), pp.149–152, Greece, 2001.
[23] S.B. Pollard, J.E.W. Mayhew, and J.P. Frisby, "PMF: A stereo correspondence algorithm using disparity gradient limit," Perception, vol.14, pp.449–470, 1981.
[24] S.D Cohran and G. Medioni, "3D surface description from binocular stereo," IEEE Trans. Pattern Anal. & Mach. Intell., vol.14, no.10, pp.981–994, Oct. 1992.
[25] W. Haff and N. Ahuja, "Surface from stereo, integrating feature matching, disparity estimation and contour detection," IEEE Trans. Pattern Anal. & Mach. Intell., vol.11, no.2, pp.121–135, Feb. 1989.
[26] J. Magarely and A. Dick, "Multiresolution stereo image matching using complex wavelet," Proc. IEEE Int. Conf. on Pattern Recognition (ICPR98), vol.1, pp.4–7, 1998.
[27] M.A. Torres and A.R. Guesalga, "Confidence factor and adaptive disparity range for reliable 3D reconstruction," in Computer Intelligence for Modeling, Control and Automation, ed. M. Mohammadian, pp.83–91, IOS Press, 1999.
[28] M. Accame, F.G.B. Natale, and D.D. Giusto, "Hierarchical block matching for disparity estimation in stereo sequence," Proc. IEEE Int. Conf. on Image Processing (ICIP95), vol.2, pp.374–377, 1995.
[29] Y. Hung, C. Chen, K. Hung, and Y. Chen, "Multipass hierarchical stereo matching for generation of digital terrain models from aerial images," Journal of Machine Vision and Applications, vol.10, pp.280–291, 1998.
[30] Z. Li and G. Hu, "Analysis of disparity gradient based cooperative stereo," IEEE Trans. Image Processing, vol.5, no.11, pp.1493–1506, Nov. 1996.
[31] S.B. Pollard, J. Porril, J.E.W. Mayhew, and J.P. Frisby, "Disparity gradient, Lipschitz continuity and computing binocular correspondence," Proc. 3rd Int. Symp. on Robot Research, ed. O.D. Faugeras and G. Giralt, pp.19–26, 1986.
[32] P. Moallem and K. Faez, "Reduction of the search space in stereo correspondence by context of disparity gradient limit," Proc. IEEE 2nd Int. Symp. on Image and Signal Processing and Analysis (ISPA'01), pp.164–169, Croatia, 2001.
[33] P. Moallem and K. Faez, "Search space reduction in the edge based stereo correspondence," Proc. Int. Workshop on Vision, Modeling, and Visualization 2001 (VMV'01), pp.423–429, Germany, 2001.
[34] P. Moallem and K. Faez, "Fast stereo matching algorithms based on reduction in the search space region by new analysis in the disparity gradient" (in Farsi), Proc. 9th Iranian Conf. on Electrical Engineering (ICEE2001), vol.2, pp.17.1-17.8, Iran, 2001.
[35] T. Frohlinghaus and J.M Buhmann, "Real-time phase-based stereo for mobile robot," Proc. EUROBOT96, pp.178–185, 1996.
[36] S. Tanaka and A.C. Kak, "Rule-based approach to binocular stereopsis," in Analysis and Interpretation of Range Image, ed. R.C Jain, A. K Jain, Springer Verlag, 1990.
[37] Y. Shirai, Three dimensional Computer Vision, Springer Verlag, 1987.
[38] R. Jain, R. Kasturi, and B.G. Schunk, Machine Vision, McGraw-Hill, 1995.
[39] CMU Stereo Images Database at `http://www.vasc.ri.cmu.edu/idb/html/stereo/index.html`.



Payman Moallem was born in 1970 in Tehran, Iran. He received his B.S. and M.S. degrees both in Electronics Engineering from Esfehan University of Technology and Amirkabir University of Technology (Tehran Polytechnic), Iran, in 1992 and 1995 respectively. Since 1994, he has researched in Iranian Research Organization, Science and Technology (IROST) on the topics like, parallel algorithms and hardware used in image processing, DSP based systems, visual target tracking and robot stereo vision. Meanwhile, he has joined to Iranian Aerospace Industrial Organization (AIO) since 1999. He has been a Ph.D. candidate at the Electrical Department of Amirkabir University of Technology since 1996. His interests include fast stereo vision, robot vision, target tracking, real-time video processing, image recognition and neural networks. He has published many papers in the mentioned areas. He is a member of SPIE. Email: m7523903@aut.ac.ir



Karim Faez was born in Semnan, Iran. He received his B.S. degree in Electrical Engineering from Tehran Polytechnic University as the first rank in June 1973, and his M.S. and Ph.D. degrees in Computer Science from University of California at Los Angeles (UCLA) in 1977 and 1980 respectively. Prof. Faez was with Iran Telecommunication Research Center (1981–1983) before joining Amirkabir University of Technology (Tehran Polytechnic) in Iran, where he is now a professor of Electrical Engineering. He was the founder of the Computer Engineering Department of Amirkabir University in 1989 and he has served as the first chairman during April 1989-Sept. 1992. Professor Faez was the chairman of planning committee for Computer Engineering and Computer Science of Ministry of Science, research and Technology (during 1988–1996). His research interests are in Pattern Recognition, Image Processing, Neural Networks, Signal Processing, Farsi Handwritten Processing, Earthquake Signal Processing, Fault Tolerant System Design, Computer Networks, and Hardware Design. He is a member of IEEE, and ACM. Email: kfaez@aut.ac.ir



Javad Haddadnia received the BS.C degree in Electrical and Electronic Engineering with the first rank in 1993 and MS.c in Electronic Engineering in 1995, respectively, all from Amirkabir University of Technology, Tehran, Iran. He is now a Ph.D. candidate in the Electrical Engineering Department of Amirkabir University of Technology, Tehran, Iran. His research interests include digital image processing, computer vision and face detection and recognition. He has published many papers in the area of image processinhg. He served as a visiting research at the University of Windsor, Canada. He is a member of SPIE and CIPPR.