A New Genetic Encoding for Edge Point Stereo Matching

H. ISSA¹, Y. RUICHEK², J.-G. POSTAIRE¹ ¹ Laboratoire d'Automatique I3D Université des Sciences et Technologies de Lille 59655 Villeneuve d'Ascq Cedex FRANCE

² Laboratoire Systèmes et Transports Université de Technologie de Belfort-Montbéliard 90010 Belfort Cedex FRANCE

Abstract: - Stereo vision is a well-known technique for obtaining depth information from two or more video images of a scene. The heavily investigated problem in this approach is how to find the corresponding features extracted from the stereo images. In this paper, we present a genetic scheme to the stereo correspondence problem where a new solution encoding is proposed. The problem is formulated in terms of finding pairs of true matches that satisfy local constraints and global ones. Local constraints are used to encode only compatible matches, whereas global constraints are used to define the fitness function such that best correspondences correspond to its minima. To demonstrate its effectiveness, the proposed genetic stereo matching is applied for extracting depth information of a scene seen by a linear stereoscopic sensor.

Key-Words: - Depth computation, Genetic algorithms, Genetic encoding, Image processing, Linear images, Object localization, Stereo matching, Stereo vision.

1 Introduction

Stereo vision is one of the most important topics in computer vision. The goal is to obtain depth information of objects from two or more stereo images taken at different viewpoints. The key problem in this approach consists of identifying features in stereo images that are generated by the same physical feature in the threedimensional world. The difference of the viewpoint positions causes a relative displacement, called disparity, of the corresponding features in the stereo images. Such relative displacement encodes the depth information, which is lost when the three dimensional structure projects to a retinal plane. Once the matching is established, the depth computation is reduced to a simple triangulation technique.

According to the considered application, the existing stereo techniques are roughly grouped into two categories. One is feature-based methods and the other is intensity- or area-based methods [1]. The feature-based methods use zero-crossing points, edges, line segments, etc. These methods are generally applied when only 3D information are required, as in object localization. Intensity-based methods use dense low-level features and intensity. This type of approaches are considered for 3D scene reconstruction applications. Genetic algorithms are efficient search methods based on the principles of population genetics, i.e., mating, chromosome crossover, gene mutation, fitness, and natural selection [2]. Recently, many genetic approaches have been proposed for the area-stereo matching problem [3-4]. To obtain a fittest disparity map, these methods use fitness functions defined from similarity and disparity smoothness constraint.

Previously, the authors have proposed a genetic algorithm for the feature-based stereo matching [5]. After generating a population of possible matches that respect two local constraints, a genetic search is applied to reach a solution for which the matches are compatible as well as possible with three competing global constraints. The genetic algorithm provides good matching results, but its major limitation is the time computation, which is long. To overcome this problem, we propose, in this paper, a new compact solution encoding, which allows exploring optimally the solution space and hence improves the convergence time.

To evaluate its effectiveness, the new genetic stereo matching scheme is applied to localize objects using linear stereo vision.

The outline of this paper is as follows. Section 2 describes the procedure used to extract edge points from linear images. The new solution encoding for the stereo matching problem is presented is section 3. The fitness

function for chromosome evaluation is described in section 4. The genetic stereo matching algorithm is summarized in section 5. Before concluding, experimental results are presented in section 6.

2 Edge Point Extraction

Low-level processing of a couple of two stereo linear images yields the features required in the correspondence process. Edge points appearing in these simple images, which are unidimensional signals, are valuable candidates for matching because large local variations in the gray-level function correspond to the boundaries of objects being observed in a scene.

Edge point extraction is performed by means of the recursive differential operator proposed by Deriche [6]. After derivation, we use first a low threshold value t in order to remove the very small responses of the differential operator lying in the range [-t, +t]. The adjustment of t is not crucial. Good results have been obtained with t adjusted at 10% of the greatest amplitude of the response of the differential operator.

A procedure is then applied to select the pertinent local extrema among the remaining edge points [7]. This is achieved by splitting the gradient magnitude signal into adjacent intervals where the sign of the response remains constant (Fig.1). In each interval of constant sign, the maximum amplitude indicates the position of a unique edge point associated to this interval when, and only when, this amplitude is greater than t.



Fig.1. Edge point extraction

The feature extraction method is applied to the left and right linear images and yields two lists of edge points as outputs. Each edge point is characterized by its position in the image, the amplitude and the sign of the response of Deriche's operator.

3 Solution Encoding

To solve the correspondence problem using a genetic algorithm, one must find a chromosome representation which codes the solution of the problem.

A classical solution encoding is to use a binary string where each element explores the hypothesis that an edge point in the left image matches or not an edge point in the right image. With this representation, a genetic algorithm manipulates chromosomes with large size: for N_L edge points in the left image and N_R edge points in the right image, the size of the chromosome is $N_L x N_R$. As a consequence, the convergence time of the genetic algorithm is long.

To overcome this limitation, we propose a new solution encoding which allows to explore optimally the solution space and hence improves considerably the convergence time. Let *L* and *R* be the lists of the edge points extracted from the left and right images, respectively. Let N_L and N_R be the numbers of edge points in *L* and *R*, respectively. The chromosome *CM* we use is encoded as a string of N_{max} integers whose values are between 0 and N_{min} , where $N_{max} = max (N_L, N_R)$ and $N_{min} = min (N_L, N_R)$. If $CM_i = 0$ then the edge point *i* have no match, otherwise, the edge points *i* and CM_i are matched.

In the chromosome, we authorize only compatible matches with respect to two local constraints. The first one, named geometric constraint, assumes that a couple of edge points l and r appearing in the left and right linear images, respectively, represents a possible match only if the constraint $x_l > x_r$ is satisfied, where x denotes the position of the edge point in the image. The second local constraint is the slope constraint, which means that only edge points with the same sign of the gradient are considered to be matched.

4 Fitness Function

A genetic algorithm needs a fitness function for evaluating the chromosomes. The fitness function representing the stereo correspondence problem is formulated from three global constraints. The first one is the uniqueness constraint, which assumes that one edge point in the left image matches only one edge point in the right image (and vice-versa). The second global constraint is the ordering constraint, which tends to preserve the order, in the images, between the matched edge points. This means that if an edge point l_1 in the left image is matched with an edge point r_1 in the right image, then it is impossible for an edge point l_2 in the left image, such that $x_{l_2} < x_{l_1}$, to be matched with an edge point r_2 in the right image for which $x_{r_2} > x_{r_1}$. The third constraint is the smoothness constraint, which assumes that neighboring edge points have similar disparities.

The fitness function is defined such that best matches correspond to its minima. With the encoding described above, it can be expressed as:

$$H = W_{I} \left(\sum_{i=1}^{N_{max}-1} \sum_{j=i+1}^{N_{max}} U(CM_{i}, CM_{j}) + \left(N_{max} - N_{min} - \sum_{i=1}^{N_{max}} Z(CM_{i}) \right)^{2} \right) + W_{2} \sum_{i=1}^{N_{max}-1} \sum_{j=i+1}^{N_{max}} O(CM_{i}, CM_{j}) + W_{3} \sum_{i=1}^{N_{max}} \sum_{j=i/j \neq i \text{ and } CM_{i} \neq CM_{j}}^{N_{max}} S(CM_{i}, CM_{j}) \right)$$

$$(1)$$

where W_1 , W_2 , and W_3 are weighting positive constants.

The first term of the fitness function corresponds to the uniqueness constraint, where the coefficient $U(CM_i, CM_j)$ represents a penalty when the constraint is violated:

$$U(CM_i, CM_j) = \begin{cases} 1 & \text{if } CM_i = CM_j \text{ and } CM_i \neq 0 \text{ and } CM_j \neq 0 \\ 0 & \text{otherwise} \end{cases}$$
(2)

The coefficient $Z(CM_i)$ is used to support the chromosomes encoding a maximum number of matches:

$$Z(CM_i) = \begin{cases} 1 & \text{if } CM_i = 0\\ 0 & \text{otherwise} \end{cases}$$
(3)

It can be seen that the term of the uniqueness constraint tends to its minimum 0 when each edge point in one image matches one and only one edge point in the other image.

The second term allows to respect the ordering constraint. The coefficient $O(CM_bCM_j)$ indicates if the order between the two pairs of edge points (i, CM_i) and (j, CM_i) is respected. It is computed as:

$$O(CM_{i}, CM_{j}) = \begin{cases} 1 & \text{if } CM_{j} < CM_{i} \text{ and } CM_{i} \neq 0 \text{ and } CM_{j} \neq 0 \\ 0 & \text{otherwise} \end{cases}$$
(4)

The third term is used to enforce the smoothness constraint. The coefficient $S(CM_i, CM_j)$ indicates how compatible are the two pairs of edge points (i, CM_i) and (j, CM_j) . This compatibility measure is computed as follows:

$$S(CM_i, CM_j) = C(X) = \frac{2}{1 + e^{(X - \mathbf{w})}} - 1$$
 (5)

where X is the absolute value of the difference between the disparities of the edge point pairs (i, CM_i) and (j, CM_i) , expressed in pixels.

The nonlinear function C(X) scales the compatibility measure smoothly between -1 and 1. The parameter **w** is adjusted so as to allow some tolerance with respect to noise and distortion. It is chosen such that a high compatibility is reached for a good match when X is close to 0, while a low compatibility corresponds to a bad match when X is very large. A satisfying value of this parameter is experimentally selected as w = 20.

5 Genetic Stereo Matching Algorithm

The genetic algorithm for the stereo correspondence works as follows. A population of chromosomes is randomly produced so that the local stereo constraints are satisfied. From a current population and after evaluation of each chromosome, the population is reproduced by selecting particular chromosomes with a selection probability proportional to the fitness value. Crossover and mutation operations are then performed to produce a new generation of chromosomes. To obtain the new generation, we use deterministic and stochastic selections. Deterministic selection, which uses the elitist strategy, represents 10% of the size of the population. The algorithm is iterated until a pre-specified number of generations is reached. Once the evolution process is achieved, the optimal chromosome, which corresponds to the minimum value of the fitness, gives the pairs of matched edge points.

6 Experimental Results

The performance of the genetic matching approach has been evaluated for localizing objects using linear stereo vision.

6.1 Stereo vision with liner cameras

A linear stereo system is build with two line-scan cameras, so that their optical axes are parallel and separated by a distance E (Fig.2). Their lenses have a same focal length f. The fields of view of the two cameras are merged in the same plane, called optical plane, so that the cameras shoot the same scene [8].



Fig.2. Geometry of the cameras.

Let us define the base-line joining the perspective centers O_l and O_r as the X-axis, and let Z-axis lie in the optical plane, parallel to the optical axes of the cameras, so that the origin of the $\{X, Z\}$ coordinate system stands midway between the lens centers (Fig.3). Let us consider a point $P(x_p, z_p)$ of coordinate x_p and z_p in the optical plane. The image coordinates x_l and x_r represent the projections of the point P in the left and right imaging sensors, respectively. This pair of points is referred to as a corresponding pair.



Fig.3. Pin-hole lens model.

Using the pin-hole lens model, the coordinates of the point *P* in the optical plane can be found as follows:

$$z_p = \frac{E \cdot f}{d}$$
(6)
$$x_p = \frac{x_l \cdot z_p}{f} - \frac{E}{2} = \frac{x_r \cdot z_p}{f} + \frac{E}{2}$$
(7)

where f is the focal length of the lenses, E is the baseline width and $d = |x_l - x_r|$ is the disparity between the left and right projections of the point *P* on the two sensors.

For object localization, the tilt angle of the stereo setup is adjusted so that the optical plane intersects the pavement at a given distance in front of the stereoscope (Fig.4).



Fig.4. Geometry of the linear stereo set-up.

6.2 Example of scene to be analyzed

One of the sequences shot by the linear stereo set-up is shown in Fig.5, where the linear images are represented as horizontal lines, time running from top to bottom. In sequence, a pedestrian moves towards the this stereoscope. The trajectory of the pedestrian during the sequence is shown in Fig.6. On the stereo images, we can see clearly the white lines of the pavement. A car located out of the vision field of the stereoscope projects a shade, which is visible on the right of the images as a black area.



Left sequence.



Right sequence. Fig.5. Example of stereo sequence.



Fig.6. Trajectory of the pedestrian during the sequence.

6.3 Depth computation

The stereo sequence, presented above, has been processed using the proposed genetic matching algorithm. The disparities of all matched edge points are used to compute the positions and distances of the edge points of the objects seen in the stereo vision sector. The results are shown in Fig.7, in which the distances are represented as gray levels, the darker the closer, whereas positions are represented as the positions along the horizontal axis. As in Fig.5, time runs from top to bottom.



Fig.7. Depth reconstruction.

The genetic parameters are selected as follows. Number of generations $N_G = 300$, size of the population of chromosomes *Size_Population* = 200, crossover probability $P_C = 0.6$ and mutation probability $P_M = 1/Size_Chromosome$.

The genetic approach provides good matching results. The edge points of the two white lines have been correctly matched and their localization is stable during the time. Indeed, their positions and distances are constant. The pedestrian is well localized as he comes closer and closer to the stereoscope. The transition between the pavement and the area of shade is also well localized. The presence of a few bad matches is noticed when occlusions occur, i.e., when the pedestrian hides one of the white lines to the left or right camera. These errors are caused by matching the edge points of the white line, seen by one of the cameras, with those representing the pedestrian. The convergence time of he genetic algorithm is considerably reduced. Indeed, with the new solution encoding, the processing time of the stereo sequence is about 7 minutes instead of about 8 days with the classical binary encoding.

6 Conclusion

A genetic approach for edge point stereo matching is presented. The correspondence problem is first formulated in terms of finding pairs of true matches that satisfy the constraints on the solution. A genetic search scheme based on a new solution encoding is then proposed. The stereo correspondence problem is characterized by two types of constraints: local constraints and global ones. Local constraints are used to encode only compatible matches. Global constraints are formulated to define the fitness function such that best matches correspond to its minima. The performance of the genetic matching approach is evaluated for extracting depth information from stereo linear images. The experiments show that the proposed method provides good matching results. Furthermore, with the new solution encoding, the convergence time is considerably reduced when compared to the convergence time required with the classical binary encoding.

References:

- R.M. Haralick, and L.G. Shapiro, *Image Matching, Computer and Robot Vision*, Part 2, Addison-Wesley, New York, 1992.
- [2] D. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading, MA, 1989.
- [3] K.P. Han, K.W. Song, E.Y. Chung, S.J. Cho, and Y.H. Ha, Stereo matching using genetic algorithm with adaptative chromosomes, *Pattern Recognition*, No.34, 2001, pp. 1729-1740.
- [4] H. Saito, and M. Mori, Application of genetic algorithms to stereo matching of images, *Pattern Recognition Letters*, 1995, No.16, pp. 815-821.
- [5] Y. Ruichek, H. Issa, and J.-G. Postaire, Genetic approach for obstacle detection using linear stereo vision, *IEEE Intelligent Vehicles Symposium*, 2000, pp. 261-266.
- [6] R. Deriche, Fast algorithms for low-level vision, Title of the Paper, *IEEE Trans. on Patt. Anal. Mach. Intel.*, Vol.1, No.12, 1990, pp. 78-87.
- [7] J.-C Burie, J.-L. Bruyelle, and J.-G. Postaire, Detecting and localizing obstacles in front of a moving vehicle using linear stereo vision, *Mathematical and Computer Modelling*, Vol.22, No.4-7, 1995, pp. 235-246.

[8] J.-L. Bruyelle, Conception et Realisation d'un Dispositif de Prise de Vue Stéréoscopique Linéaire – Application à la Détection d'Obstacles à l'Avant d'un Véhicule Routier, PhD, University of Sciences and Technologies of Lille, France, 1995.