Stripe Mesh Based Disparity Estimation by Using 3-D Hough Transform

Fatih M. Porikli
Electrical Engineering Department
Polytechnic University
Brooklyn, NY 11201

Abstract

This correspondence presents a novel matching technique for disparity estimation. The technique combines advantages of an adaptive stripe-based mesh structure and Hough transform. First, a mesh that is composed of triangular patches and fits to the depth changes is generated by using edge maps. In each patch, depth of the scene is approximated by a surface. A search space is built by difference maps that are obtained by subtracting left image from the shifted versions of right image. From search space, 3-D Hough spaces are produced such that a point of Hough space represents a surface in the search space. Then, a matching algorithm finds the combination of surfaces that gives minimum matching error in a neighborhood around patches by using Hough spaces. Continuity and smoothness constraints are formulated as probability density functions that modify the Hough spaces of the following patches by using the estimated surfaces for the previous ones. Our experiments demonstrate the accuracy of the proposed method.

1 Introduction

Human brain perceives depth by acquiring a pair of images of the surrounding three-dimensional world. In the same way, depth of a scene can be extracted from a pair of images by finding the relative displacement vectors, also referred as the disparity, that corresponds to the pair-wise related image points. Most of the disparity estimation algorithms use either feature-based or area-based matching methods [1],[2],[3]. In feature-based methods, the features derived from the images, such as corner points, edges, etc. are used as matching primitives, thus disparity is obtained only for those points. On the other hand, area-based methods aim to estimate the disparity of every image point. Generally, an area-based method divides the estimation problem into three fundamental parts: (1) determining the matching window size, (2) selecting constraints to minimize estimation errors, and (3) developing a search algorithm to optimize accuracy and speed.

Recently, we have developed an adaptive stripe-based mesh structure [4] that fits in the edges of the original intensity image. In this paper, an area-based disparity estimation technique that combines Hough transform which is an elegant technique to detect surfaces and adaptive mesh structure is proposed. The depth of the scene is assumed to be modeled by planar functions in the mesh patches. Then, the function parameters that give the best depth model of the scene are obtained by searching the Hough space.

One motivation of using stripe mesh is to apply epipolar line constraint more effectively. The set of all possible matches for an arbitrary patch builds up a stripe as all possible matches for a point constitute a line segment parallel to the epipoles. The uniqueness constraint asserts that a point in left image can have only one corresponding matching point in the right one, and it is essential to attenuate the effects of noise and to get a controlled and ordered search. Utilizing stripes simplifies managing search space since the mutual estimation range for neighboring patches is transformed into a simple stripe structure rather than an arbitrary region. Furthermore, the edges parallel to the epipoles are forced to coincide with the up and down sides of the patches, so that ill-conditioning caused by matching horizontal edges is minimized. Using Hough transform enables interpreting disparity map by the model surfaces even during the estimation stage as well as propagating pre-estimated disparity values to the next patches and stripes.

2 Stereo System

Let the baseline, B, of the stereo-scopic camera system be parallel to the x-axis and assume the imaging planes are coplanar as in Fig.1. Let (X, Y, Z) represent a point in the physical world and $(x_L, y_L), (x_R, y_R)$ be its perspective projections on the left and right image

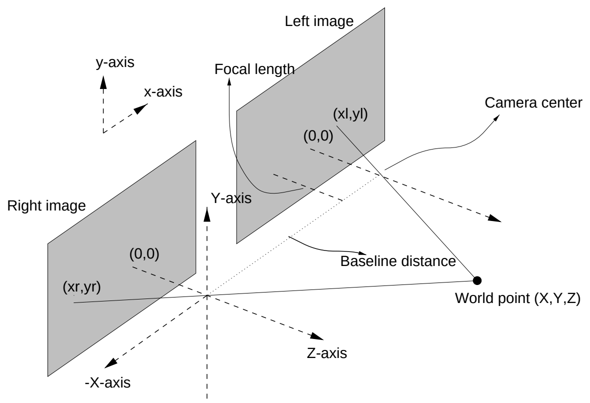

Figure 1: Stereo system

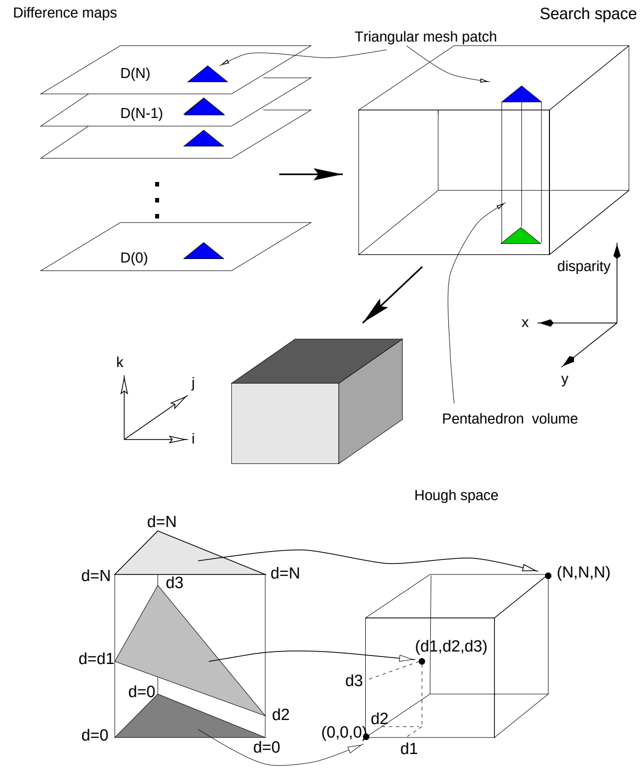

Figure 2: Hough space construction

planes respectively. Then, we obtain

$$\begin{bmatrix} x_L \\ y_L \end{bmatrix} = \begin{bmatrix} f\frac{X}{Z} \\ f\frac{Y}{Z} \end{bmatrix}, \qquad \begin{bmatrix} x_R \\ y_R \end{bmatrix} = \begin{bmatrix} f\frac{(X+B)}{Z} \\ f\frac{Y}{Z} \end{bmatrix}$$

where f is the focal length of the cameras. By modeling depth, Z, as a planar surface

$$Z = aX + bY + c$$

where a, b, c, are real constants, we get disparity as

$$d = x_R - x_L = \frac{fB}{Z} = -\frac{B}{c}(ax_L + by_L - f)$$

which is also a planar equation. In other words, if the depth of the physical world seen through a triangular patch is a planar surface then the disparity in the patch satisfies a planar equation.

3 Stripe-Based Triangular Mesh

The stripe-mesh, $M(s, p_n)$ where s stands for stripe number, p stands for patch number, and n is one of the three corners, i.e. nodes, is constructed from the edge maps of the left image. The vertical and horizontal edge maps are obtained by using 3×3 sobel operators, and later, refined by a confident algorithm to enhance the line continuity. For each row, an edge strength score, h_i, is calculated by adding values of the horizontal edges on the row i. This score provides an insight on the total magnitudes of the edges that are oriented horizontally and intersect the corresponding row. The up and down borders of a stripe, s, are decided by choosing the biggest h_i's by simultaneously constraining them with a closeness and a minimum stripe width criteria. Within each stripe, the strength of each possible abutting line segments, l_j, is calculated similarly by projecting horizontal and vertical edges on the possible line directions. The biggest l_j's, exceeding a strength threshold and representing the line segments that are sufficiently away from the previously selected ones, are chosen as quadrilateral patch borders. In the decision-making process, the selections causing irregular triangularization are also discarded. To achieve vertical continuity of the line segments from stripe to stripe, l_j's are weighted by such a function that if there is a line segment chosen as patch border in the previous stripe which is close to the lines corresponding to l_j's in the present stripe, the l_j's are increased. This weighting forces patch borders in conjoint stripes to align as much as possible. Then each quadrilateral patch is divided into two triangle patches, p's. $M(s, p_n)$ is the coordinates of the p^{th} patch of s^{th} stripe.

4 Application of Hough Transform

The Hough method [5] is an efficient implementation of a generalized matching and filtering strategy which has been used in a variety of domains including the object detection by template matching. Basically, the search space, which is processed to select the points that have a high likelihood of being on a template, is transformed onto a quantized parameter space called as Hough space. An accumulator array is then formed for the parameter space such that the value of each element is incremented by the strength of the search space point if this point matches with the corresponding template. The maximum in the accumulator array gives the best fitting template.

4.1 Generating Hough Space

The disparity range and image size determine the dimension of search space, $\mathcal{S}$, from which Hough spaces, $\mathcal{H}$, are generated. The right image is shifted horizontally up to a maximum disparity value and subtracted from the left image either pixel-wise or in a window around pixels in order to produce difference maps $D_n(x, y)$, $n = 1 \dots N$ where N is the maximum disparity. Since the surfaces that yield minimum difference error are what we want to detect, the magnitudes of the difference maps are reverted and indexed vertically, $D_n(x, y) \mapsto \mathcal{S}_{(x,y,n)}$, to obtain the search space as shown in Fig.2. By using the generated mesh, search space is divided into smaller pentahedra such that the top and bottom of each pentahedron correspond to a triangular patch in the mesh. Hough spaces are generated for each of these volumes. The points in the Hough space $\mathcal{H}$ represents a surface in the corresponding pentahedron. The depth of the physical world in patches is affine modeled hence the disparity corresponds to the planar surfaces which are parameterized by the disparity values of the three corner points. Hough spaces are cubicles and built by using three plane parameters quantized between the minimum and the maximum disparity values for each pentahedron. The indices of the cubicles sides represent the disparity value of the corresponding corners. If $\Psi_{p,\mathbf{d}}$ is a plane in the pentahedron p such that $\mathbf{d} = (d_1, d_2, d_3)$ and its corner points in $\mathcal{S}$ are $(M(s, p_1), d_1), (M(s, p_2), d_2), (M(s, p_3), d_3)$, then the corresponding point in $\mathcal{H}^p$ is (d_1, d_2, d_3). The accumulator array f_p is defined as a function in $\mathcal{H}^p$ and is equal to the sum of the search space points (x, y, n) on Ψ_p's.

4.2 Continuity Constraints

Vertical and horizontal smoothness constraints are used to confine Hough spaces by updating the values of accumulator arrays. Both constraints are defined as density functions. To achieve horizontal smoothness, the accumulator arrays of the Hough spaces belong to the next patches, $p+1, p+2$ in the stripe is scaled by a density function φ_v which is determined by the best estimation of the current patch p. φ_v is a 3-D Laplace function

$$\varphi_v(\mathbf{d}) = e^{-\kappa |T(\mathbf{d}, \tilde{\mathbf{d}})|}$$

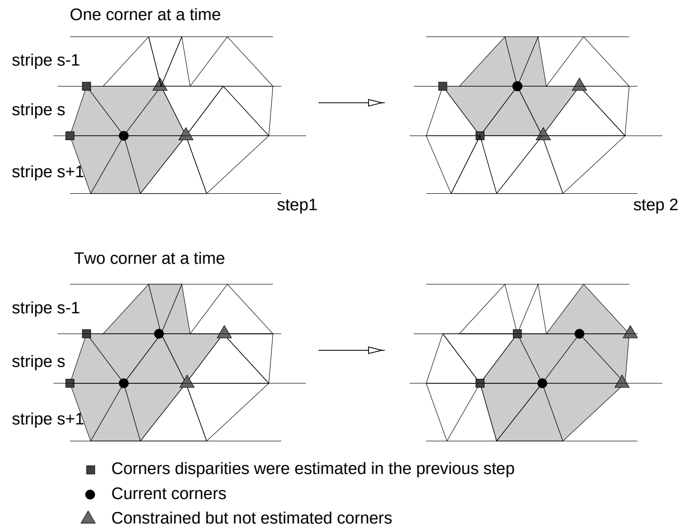

Figure 3: Matching schemes for continuous case

where $\tilde{\mathbf{d}} = (\tilde{d}_1, \tilde{d}_2, \tilde{d}_3)$ is the estimated point of the current patch and $\kappa > 0$ is normalization constant. For strictly smooth cases, κ has big values. $T(\mathbf{d}, \tilde{\mathbf{d}})$ is defined as

$$T = \begin{cases} (d_1 - \tilde{d}_1)^2 + (d_3 - \tilde{d}_2)^2, & p \text{ odd} \wedge \mathbf{d} \in p+1 \\ (d_1 - \tilde{d}_2)^2 + (d_3 - \tilde{d}_3)^2, & p \text{ even} \wedge \mathbf{d} \in p+1 \\ (d_3 - \tilde{d}_2)^2, & p \text{ odd} \wedge \mathbf{d} \in p+2 \\ (d_1 - \tilde{d}_2)^2, & p \text{ even} \wedge \mathbf{d} \in p+2. \end{cases}$$

Similarly, vertical smoothness is performed by another 3-D Laplace density function derived from the disparity values estimated in the previous *stripe*. Later, Hough spaces are filtered out to remove planes causing flip-overs of mesh nodes.

5 Matching Algorithm

Matching is done by finding the combination of the Hough space points that give the maximum of a confidence measure calculated over the accumulator arrays of the neighboring patches in the same stripe. Different schemes for the continuous and discontinuous disparity estimation are used. In the continuous case two scenarios are possible: (1) the patch corners that weren't estimated before spans the Hough spaces, hence these spaces are either in 2-D or lines, and at each step only such corners are estimated, (2) Hough spaces are not limited and estimation confidence is calculated in the neighboring patches such that the combination of Hough space points corresponds to connected surfaces. The neighborhood contains three adjoint patches for one corner point, and four patches for two connected corners in the same stripe, and abut-

Figure 4: Left image

ting patches in up and down stripes. At each stage, estimation can be done either for one corner or two connected corners as in shown in Fig.3. The horizontal density functions update Hough spaces of the next patches, and after estimation is completed for a stripe, Hough spaces of the following stripe are modified by the vertical density functions.

6 Experimental Results

The proposed algorithm is tested on an image pair sized 320×240 with disparity range 32 pixels. Hough spaces are generated at 1 pixel resolution. Original left image, generated stripe mesh, and estimated depth map by using continuous case are shown in Figs. 4, 5. The PSNR of the estimated right image from the left one is $25.82dB$. The experiments prove that proposed method produce accurate depth maps even for the texture-free image regions and it is free from the spark-like noise which happens in block matching.

7 Discussion

In this contribution, a disparity estimation algorithm that uses stripe mesh and Hough transform is introduced. After an adaptive stripe-based triangular mesh is generated, 3-D Hough spaces are established with accumulator arrays derived from the difference maps. Then, maximum of the combination of accumulator array elements of neighboring patches is chosen. The consequent patches are updated by the probability density functions. Thus far, only uncomplicated disparity models have been implemented. As a follow-up work, different probability density functions for continuity constraints and higher order depth models will be studied by using the same technique.

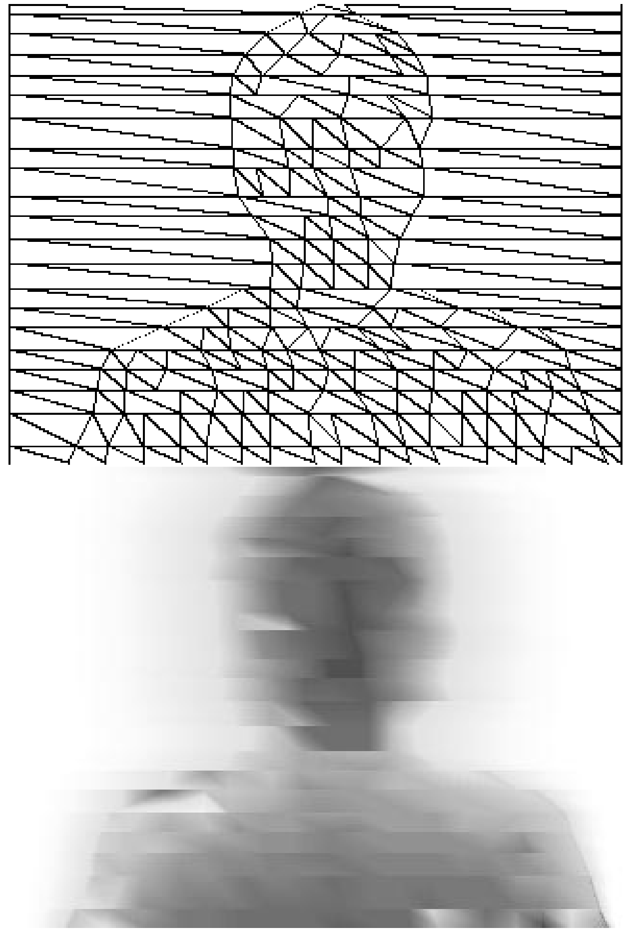

Figure 5: Generated mesh and estimated depth map

References

[1] S. Barnard and M. Fisher, "Computational Stereo", *ACM Comput. Surveys*, Vol. 14, no 4, pp. 553-572, December 1982.

[2] T. Kanade and M.Okutomi, "A stereo matching algorithm with an adaptive window", *IEEE Trans. Patt. Anal. Machine Intell.,* Vol 16, no. 9, pp. 920-931, September 1994.

[3] S. D. Cochran and G. Medoni, "3-D surface description from binocular stereo," *IEEE Trans. Patt. Anal. Machine Intell.*, vol 14, no. 10, October 1992.

[4] F.M. Porikli and Y. Wang, "Disparity estimation by patch matchin", *PCS*, Munich, September 1997.

[5] P.V.C. Hough, "Method and means for recognizing complex pattern", *U.S. Patent 3,069,654*, 1962.