Probabilistic Correspondence Matching using Random Walk with Restart

Changjae Oh
ocj1211@yonsei.ac.kr
Bumsub Ham
mimo@yonsei.ac.kr
Kwanghoon Sohn
khsohn@yonsei.ac.kr

School of Electrical and Electronic Engineering
Yonsei University
Seoul, Korea

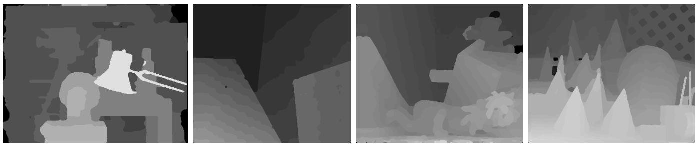

Figure 1: Disparity estimation results of the proposed method.

Introduction - Local correspondence matching methods are mainly composed of three steps: matching cost computation, cost aggregation, and disparity computation. Let us assume that a truncated absolute difference is used in matching cost computation. Each step can be represented as following probability inferring problem.

$$p_0(\mathbf{x},\mathbf{d}) = \max\left(\sigma - \|I_R(\mathbf{x}) - I_T(\mathbf{x},\mathbf{d})\|_1, 0\right) \quad (1)$$

$$p_{n+1}(\mathbf{x},\mathbf{d}) = \frac{\sum_{\mathbf{y}\in\mathcal{N}} w(\mathbf{x},\mathbf{y})\, p_n(\mathbf{y},\mathbf{d})}{\sum_{\mathbf{y}\in\mathcal{N}} w(\mathbf{x},\mathbf{y})} \quad (2)$$

$$\mathbf{d}(\mathbf{x}) = \underset{\mathbf{d}\in[\mathbf{d}_1,\ldots,\mathbf{d}_D]}{\arg\max}\; p_N(\mathbf{x},\mathbf{d}) \quad (3)$$

where $\mathbf{x} = [x,y]^T$ and $\mathbf{d} = [d,0]^T$ represent the position and disparity vector, respectively. The weight between $\mathbf{x}$ and $\mathbf{y}$ is defined as $w(\mathbf{x},\mathbf{y})$. First, a matching probability $p_0(\mathbf{x},\mathbf{d})$ is computed by an absolute difference between points on the reference image $I_R(\mathbf{x})$ and on the '$\mathbf{d}$'-shifted target image $I_T(\mathbf{x},\mathbf{d})$ with the threshold σ as in Equation 1. Then, in Equation 2, the probability $p_n(\mathbf{x},\mathbf{d})$ is iteratively aggregated with $\mathbf{y}\in\mathcal{N}$ where $\mathcal{N}$ is the neighborhood of $\mathbf{x}$. Finally, an optimal disparity $\mathbf{d}(\mathbf{x})$ is selected within a search range D for the aggregated probability $p_N(\mathbf{x},\mathbf{d})$ after the maximum iteration N, by winner-takes-all (WTA) as in Equation 3.

Proposed Method - Since correspondence matching method can be regarded as probability inferring problem, the probability optimization can be used as cost aggregation. In this paper, we present the cost plane optimization using RWR framework.

Consider the cost plane as an undirected weighted graph. We present steady-state probability computation by the RWR with the given graph. Let us denote the initial cost plane as $\mathbf{P}_n^k = [p_n(\mathbf{x}_i,\mathbf{d}_k)]_{M\times 1}$ and the adjacency matrix as $\mathbf{W} = [w_{ij}]_{M\times M}$, where M and w are the size of reference image and the weight between $\mathbf{x}_i$ and $\mathbf{x}_j$, respectively. The RWR can be formulated in an iterative manner as follows:

$$\begin{aligned}\mathbf{P}_{n+1}^k &= (1-\alpha)\mathbf{D}^{-1}\mathbf{W}\mathbf{P}_n^k + \alpha\mathbf{P}_0^k \\ &= (1-\alpha)\overline{\mathbf{W}}\mathbf{P}_n^k + \alpha\mathbf{P}_0^k\end{aligned} \quad (4)$$

where n denotes the number of iterations. The initial cost $\mathbf{P}_0^k$ is returned with the probability α at each iteration. The adjacency matrix $\mathbf{W}$ is normalized as $\overline{\mathbf{W}} = \mathbf{D}^{-1}\mathbf{W}$, where $\mathbf{D} = diag(D_1,\ldots,D_M)$, and $D_i = \sum_{j=1}^{M} w_{ij}$. When the solution reaches to the steady-state, $\mathbf{P}_n^k$ and $\mathbf{P}_{n+1}^k$ become identical, *i.e.*, the energy transition with respect to time approaches 0. Therefore, Equation 4 can be reformulated as follows:

$$\begin{aligned}\mathbf{P}_s^k &= (1-\alpha)\overline{\mathbf{W}}\mathbf{P}_s^k + \alpha\mathbf{P}_0^k \\ &= \alpha(I-(1-\alpha)\overline{\mathbf{W}})^{-1}\mathbf{P}_0^k \\ &= \mathbf{R}\mathbf{P}_0^k\end{aligned} \quad (5)$$

where $\mathbf{P}_s^k$ is the cost which reaches to the steady-state. $\mathbf{R}$ can be interpreted as affinity scores between two pixels in the initial cost plane $\mathbf{P}_0^k$. With the given steady-state solution in Equation 5, disparity can be simply selected by WTA measure as following:

$$\mathbf{d}(\mathbf{x}_i) = \underset{\mathbf{d}_k}{\arg\max}\, p_s(\mathbf{x}_i,\mathbf{d}_\mathbf{k}) \quad (6)$$

where $p_s(\mathbf{x}_i,\mathbf{d}_\mathbf{k})$ is an optimized steady-state probability obtained in Equation 5, and $\mathbf{d}_k \in \{\mathbf{d}_1,\ldots,\mathbf{d}_D\}$.

The correspondence matching within the RWR framework has the following advantages: 1) A non-trivial steady-state solution is guaranteed, which means that it is not needed to specify the number of iteration. In conventional methods, it is crucial to specify the number of iteration since the performance and the computation time largely depends on this parameter [1, 2]. 2) The global relationship between points or the steady-state solution can be captured by using an adjacent neighborhood only. Accordingly, the proposed method gives high quality matching performance in a semi-global manner with low complexity.

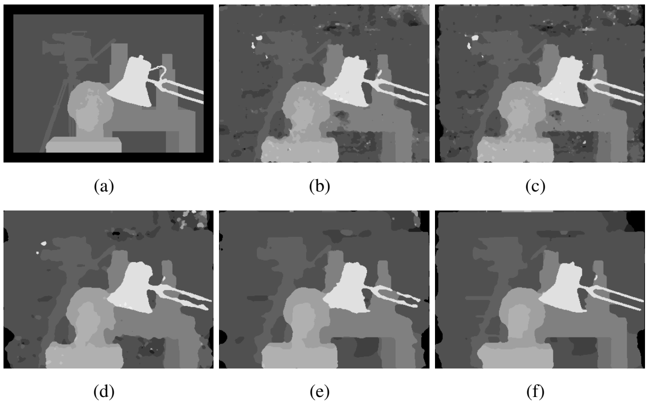

Figure 2: Advantage of our method. Disparity estimation results using the smallest window in cost aggregation. (a) The ground truth. 3x3 size window are used in (b) Adaptive weight [3], (c) Costfilter [4], and 4-neighbor pixels are used in (c) Anisotropic diffusion [1], (d) Geodiff [2], (e) Proposed method.

Experimental Results - The proposed method as shown in Figure 1 was compared with other state-of-the-art cost aggregation methods: adaptive weight (AW) [3], cost filter (CF) [4], anisotropic diffusion (AD) [1], and geodesic diffusion (GD) [2]. Note that AD is not the main proposal of [1], but just the part of their method. Table 1 shows the bad matching errors evaluated by the Middlebury website [5]. The symbol '*' indicates the results of the Middlebury evaluation website. It shows that the proposed method shows competitive results with state-of-the-art methods. The comparison of the computation times of AW, CF, GD, AD, and the proposed method are 12.1, 1.18, 0.93, 2.19, 1.0, respectively, when the computation time of the proposed method is normalized to 1.0. In order to compare the performance when the window size of each algorithm is similar, we conducted another experiment by changing the window size of AW and CF to 3x3 which is the similar to that used in AD, GD and the proposed method. Figure 2b and Figure 2c show the degraded results in AW and CF, which means that the results of these methods heavily depend on the window size.

Algorithm	*Tsukuba*	*Venus*	*Teddy*	*Cones*
AW [3]	2.77	0.46	13.2	8.60
AW [3]*	1.85	1.19	13.3	9.79
CF [4]	2.14	0.46	11.5	8.01
CF [4]*	1.85	0.39	11.8	8.24
AD [1]	3.85	1.78	14.2	8.83
GD [2]	2.96	0.45	12.4	8.65
GD [2]*	2.35	0.82	11.3	8.33
Proposed method	1.97	0.38	11.5	7.92

Table 1: Object evaluation for the proposed method

[1] D. Min and K. Sohn, "Cost aggregation and occlusion handling with WLS in stereo matching," *IEEE Transactions on Image Processing*, 17(8):1431-1442, August 2008.

[2] L. De-Maeztu, A. Villanueva, and R. Cabeza, "Near Real-Time Stereo Matching Using Geodesic Diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(2):410-416, February 2012.

[3] K. Yoon and I. Kweon, "Adaptive support-weight approach for correspondence search," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(4):650-656, April 2006.

[4] C. Rhemann, A. Hosni, M. Bleyer, C. Rother, and M. Gelautz, "Fast cost-volume filtering for visual correspondence and beyond," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 3017-3024, 2011.

[5] `http://vision.middlebury.edu/stereo`