

Real-time Dense Disparity Estimation based on Multi-Path Viterbi for Intelligent Vehicle Applications

Qian Long¹

long@toyota-ti.ac.jp

Qiwei Xie¹

qw_xie@toyota-ti.ac.jp

Seiichi Mita¹

smita@toyota-ti.ac.jp

Hossein Tehrani²

hossein_tehrani@denso.co.jp

Kazuhisa Ishimaru³

kazuhisa_ishimaru@soken1.denso.co.jp

Chunzhao Guo⁴

czguo@mosk.tytlabs.co.jp

¹ Research Center for Smart Vehicles

Toyota Technological Institute

2-12-1 Hisakata, Tempaku, Nagoya, Aichi 468-8511, Japan

² DENSO CORPORATION

1-1, Showa cho, Kariya, Aichi, 448-8661 Japan

³ NIPPON SOKEN Inc.

Nishio, Aichi, Japan

⁴ Toyota Central R&D Labs., Inc.

Nagakute, Aichi, Japan

3D scene understanding plays an essential role for intelligent vehicle applications. In these applications, passive stereo vision systems offer some significant advantages to estimate depth information compared with active systems such as 3D LIDAR. To apply stereo vision in autonomous driving, a new real-time stereo matching algorithm paired with an on-line auto-rectification framework is proposed. This method uses a bi-directional Viterbi algorithm at 4 paths to decode the matching cost space and a hierarchical structure (as shown in Fig. 1) is proposed to merge the 4 paths to further decrease the decoding error. We introduce Total Variation [1] constraint into Viterbi path for approximately modeling 3D planes at different orientations to reach a similar effect as slanted-plane models. Structural similarity (SSIM)[3] is used to measure the pixel difference between left and right images at epipolar lines to improve robustness to luminance variation. The equation for one Viterbi path is expressed by:

$$e(p, u) = \min_{u' \in L_u} \{e(p-1, u') + \lambda e^{-|G|} |u - u'| + SSIM(p, u)\} \quad (1)$$

where $e(p, u)$ is the energy of Viterbi node at pixel p and disparity u , G is the gradient information of image, λ is the parameter, L_u denotes connected Viterbi nodes to the Viterbi node at pixel p and disparity u .

Based on the output of Viterbi process, a convex optimization equation is derived to estimate epipolar line distortion. we summarize the properties of the epipolar line distortion caused by normal factors in intelligent vehicle applications. Based on these properties and inspired by the famous optical flow problem, we convert this distortion estimation problem to an optimization problem and employ the convex optimization theory to solve it. The Viterbi process and convex optimization are integrated into an online framework (as shown in Fig. 2) and two parts benefit each other without losing speed in this framework. It can automatically keep the epipolar line constraint to avoid the degradation of stereo matching results, which usually happens when other stereo matching methods being applied for driving vehicles.

Extensive experiments were conducted to compare proposed algorithm with other practical state-of-the-art methods for intelligent vehicle applications. According to evaluation results at the KITTI [2] training dataset which includes total 194 images, our method has 7.38% average error rate compared to SGBM's 12.88% and ELAS's 11.99%. We also test the proposed algorithm in our experimental autonomous vehicle at real driving environments. For any 640x480 images with maximum 40 disparities, the running time is about 196ms with GTX TITAN GPU and Xeon E5-2620 CPU. Real driving videos including featured cases and typical failure cases can be found in the supplementary material.

[1] Antonin Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20 (1-2):89–97, 2004.

[2] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? The KITTI vision benchmark suite. In *CVPR*, pages 3354–3361, 2012.

[3] Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

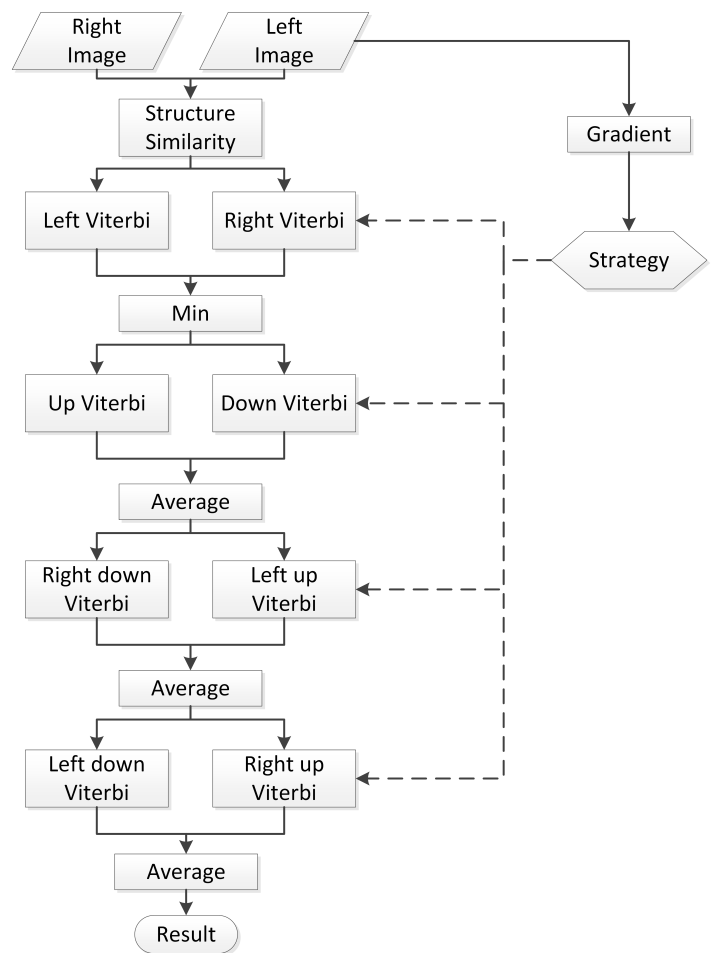


Figure 1: Hierarchical structure for the merging of multiple Viterbi paths.

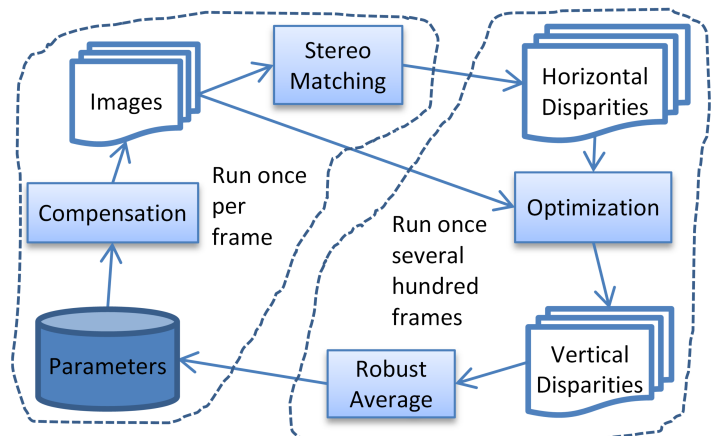


Figure 2: Framework of automatic online rectification.