Learning about SGM-Nets

Hongzhi Liu

May 7, 2018

1 Semi-global matching

As is known to all, *SGM* (Semi-global matching) is a widely used regularization method due to its high accuracy while keeping low computation cost. Even though *SGM* can obtain accurate results, tuning of *SGM*'s penalty-parameters which control a smoothness and discontinuity of a disparity map is uneasy and empirical methods have been proposed.

First I can learn a lot about *SGM* [1]. An energy function $\mathbb{E}$ for solving *SGM* is defined as Eq. 1 :

$$\mathbb{E}(D) = \sum_x \left(C(x, d^x) + \sum_{y \in N_x} P_1 T\left[|d^x - d^y| = 1\right] \right) + \sum_x \left(\sum_{y \in N_x} P_2 T\left[|d^x - d^y| > 1\right] \right). \tag{1}$$

$C(x, dx)$ represents a matching cost at pixel $x = (u, v)$ of disparity d^x. The first term represents the sum of matching costs at all pixels for the disparity map D. The second term represents slanted surface penalty P_1 for all pixels y in the neighborhood N_x of x. The third term indicates penalty P_2 for discontinuous disparity. P_2 is typically set small according to the magnitude of the image gradient, for example $P_2 = P_2' / |I(x) - I(y)|$ so that the discontinuities are easily selected.

Professor Seki proposes a new loss function in order to train neural networks which inputs are small patches and their location. New *SGM* parameterization that separates the positive and negative disparity changes in order to represent object structures discriminatively. *SGM-Nets* were able to outperform state of the art accuracy on KITTI datasets without the need for an explicit foreground shape prior such as a vehicle.

2 Semi-global matching with neural networks

During the training phase, *SGM-Nets* is iteratively trained by minimizing two kinds of costs, which are "Path cost" and "Neighbor cost". Figure 1 illustrates an overview of their proposed method. Their neural network which they call SGM-Net provides P_1 and P_2 at each pixel.

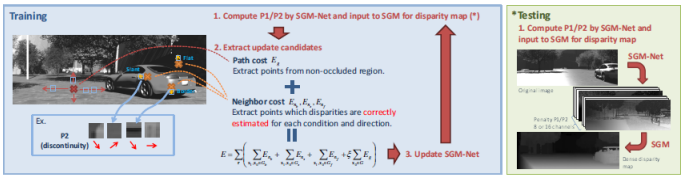

Figure 1: Overview of *SGM-Nets*. *SGM* estimates dense disparity by incorporating penalty P_1 and P_2 from *SGM-Nets*. *SGM-Nets* is iteratively trained on each aggregation direction with image patches and their positions.

References

[1] Heiko Hirschmuller. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 30(2):328–341, 2007.