Stereo ambiguity index for semi-global matching

Mathias PAGET
mathias.paget@ifsttar.fr

Jean-Philippe TAREL
jean-philippe.tarel@ifsttar.fr

Pascal MONASSE
monasse@imagine.enpc.fr

UNIVERSITÉ PARIS-EST — IFSTTAR — IMAGINE École des Ponts ParisTech — CSTB le futur en construction

We propose a stereo ambiguity index for semi-global matching that measures the ambiguity of the SGM solution. It can be used :
- to detect occlusions
- to refine the stereo reconstruction

SGM method, [1]

Stereoscopic reconstruction problem is often written as a 2D Markov's chain and amounts to minimise the following energy:

$$E(D) = \sum_{p \in P} \text{Data}(p, d_p) + \sum_{(p,q) \in N} \text{Prior}(d_p, d_q)$$

where P is a set of pixels and N the 4-connexity neighbor relation. SGM (Semi-Global Matching) consists in decomposing the grid into 1D paths in R vector directions v from the considered pixel and in minimising the energy along these paths by dynamic programing. C_V costs are recursively computed with the rule:

$$C_v(p, d) = \text{Data}(p, d) + \min_{d'} C_v(p - v, d') + \text{Prior}(d, d').$$

SGM energy is obtained by summing C_V costs over all directions and solution is obtained by selecting the minimal cost per pixel over disparities:

$$SGM(p, d) = \sum_v C_v(p, d) - (R - 1)\text{Data}(p, d)$$

$$d_p = \arg\min_d SGM(p, d).$$

Index derived from SGM energy

The SGM energy at pixel p correponds to the minimal energy over the cross around p, x_p, knowing disparity in p, d_p, is egal to d :

$$SGM(p, d) = \min_{d_x, x \in X_p | d_p = d} \{ \sum_{x \in X_p} \text{Data}(x, d_x) + \sum_{(x,y) \in N \cap X_p^2} \text{Prior}(d_x, d_y) \}.$$

The proprosed index at p is the number of disparities whose energy is lower than the minimal energy plus a threshold T_1:

$$\text{Index}(p) = \sum_d \mathbb{1}(SGM(p, d) \leq SGM(p, d_p) + T_1)$$

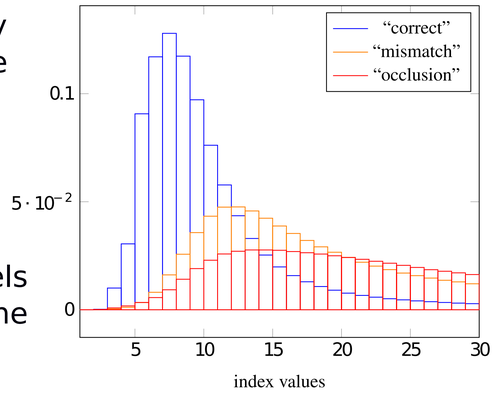
The minimal index value is 1. The higher the index, the more ambiguous the SGM energy is.

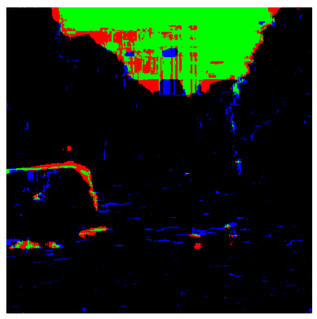
Index as occlusion detector

We perform the left-right consistency check to classify pixels into 3 categories, D_l and D_r being the left and the right image disparities :

$$\text{if } |D_l(p) + D_r(p - D_l(p))| \leq 1 \text{ correct},$$
$$\text{else if } |d + D_r(p - d)| \leq 1, \text{ for some } d \neq D_l(p), \text{ mismatch},$$
$$\text{otherwise occlusion.}$$

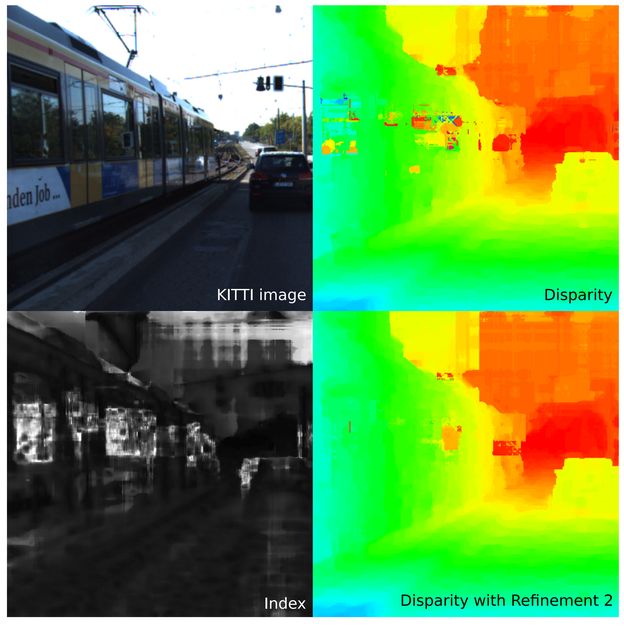
The proposed index is able to indentify "non-correct" pixels with a precision of 77% for a 50% recall using only one disparity image.





The image, at left, represents in green "correct" pixels with high index value (over-detection), in red "non-corect" pixels with high index value (good detection), in blue "non-correct" pixels with low index value (under-detection) and in black "correct" pixel with low index value.
Thin objects are under-detected due to the regularisation effected. Over-detections correspond to the sky and right point of view occlusions.

Index as confience mesure in the model



We propose two ways to refine the disparity map by using the index:
- 1) smoothing disparities where index is high,
- 2) reweight data term with index and reprocess.

Improvements have been shown in KITTI12 and KITTI15 datasets:

	Error rate (%) KITTI12/KITTI15
Original	6.12/5.23
Refinement 1)	5.38/4.77
Refinement 2)	5.25/4.55

In the images, we oberve improvement in tramway glasses and in the car, with Refinement 2).

[1] Heiko Hirschmüller, "Stereo processing by semiglobal matching and mutual information," *IEEE Transactions on patternanalysis and machine intelligence*, vol. 30, no. 2, pp. 328–341,2008