

Stereo Processing by Semiglobal Matching and Mutual Information

Mutual information을 이용한 cost 정의

[1]의 논문에선 일반적인 block matching에서 사용하는 cost에 두 확률 분포에 대한 독립정도를 판단할 수 있는 mutual information을 추가해 cost를 정의한다. Mutual information이란 joint distribution인 $P(X, Y)$ 가 존재한다고 가정하고 X, Y 가 서로 독립적이라면 $P(X, Y) = P(X) * P(Y)$ 이다. 이때, $P(X, Y)$ 와 $P(X) * P(Y)$ 의 KL divergence를 구하면 X, Y 가 서로 얼마나 dependent한지 측정이 가능함을 이용한 것이다.

$$mi_{I_1, I_2}(i, k) = h_{I_1}(i) + h_{I_2}(k) - h_{I_1, I_2}(i, k) \quad (1)$$

$$C_{MI}(p, d) = -mi_{I_b, f_D(I_m)}(I_{bp}, I_{mq}), (d : disparity) \quad (2)$$

$$q = e_{bm}(p, d) \quad (3)$$

이를 통해 기존 cost function인 p 의 absolute minimum difference 즉, $C_{BT}(p, d)$ 에서 mutual information을 이용한 cost function인 Equation 2를 제안했다. 이는 C_{BT} 보다 14% 밖에 느리지 않고 성능은 더 뛰어나다. Pixelwise cost와 smoothness constraints는 energy $E(D)$ 즉, Equation 4 표현이 가능하다.

$$E(D) = \sum_P(C(p, D_p) + \sum_{q \in N_p} P_1 T[|D_p - D_q| = 1] + \sum_{q \in N_p} P_2 T[|D_p - D_q| > 1]) \quad (4)$$

이때, 첫 항은 pixel matching cost function이며 둘째 항은 neighborhood N_p 의 P1 constant penalty이다. 이제 $E(D)$ 를 minimize하면 된다. 최종 cost를 L_r 이라 하면 Equation 5라고 할 수 있고 이를 "Cost Aggregation"이라 한다. 이때, $C(p, d)$ 는 앞서 구한 pixelwise matching cost인 C_{MI} 또는 기존 방법인 C_{BT} 를 사용할 수 있다.

$$L_r(p, d) = C(p, d) + \min(L_r(p-r, d), L_r(p-r, d-1) + P_1, L_r(p-r, d+1) + P_1, \min_i L + r(p-r, i) + P_2) - \min_k L_r(p-r, k) \quad (5)$$

$$S(p, d) = \sum_r L_r(p, d) \quad (6)$$

마지막으로 disparity를 구하는 방법은 D_b, D_m 을 통해 Equation 7을 통해 구할 수 있다. Figure 1은 SGM algorithm의 overview이다.

$$D_p = \begin{cases} D_{bp} & \text{if } |D_{bp} - D_{mq}| \leq 1, \\ D_{inv} & \text{otherwise.} \end{cases} \quad (7)$$

$$q = e_{bm}(p, D_{bp}) \quad (8)$$

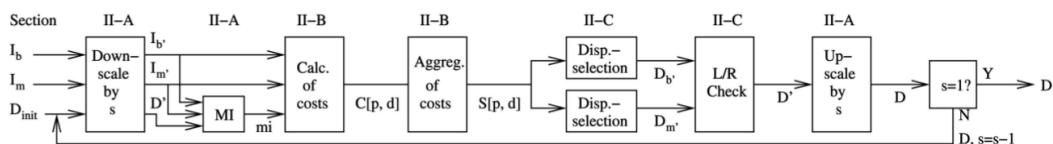


Figure 1: Overview of SGM algorithm

Experiments

Semiglobal matching algorithm은 python의 opencv에 내장 함수로써 존재한다. 일반적인 방법으로 stereoBM과 stereoSGBM을 사용하는데 stereoSGBM이 semiglobal matching algorithm에 해당하므로 이에 대한 서로 다른 sample을 비교하는 형식으로 실험하였다.

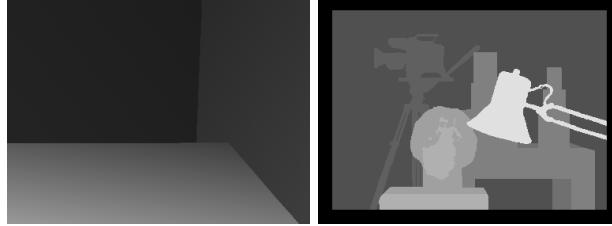


Figure 2: Ground Truth 좌 : Bull image, 우 : Tsukuba image

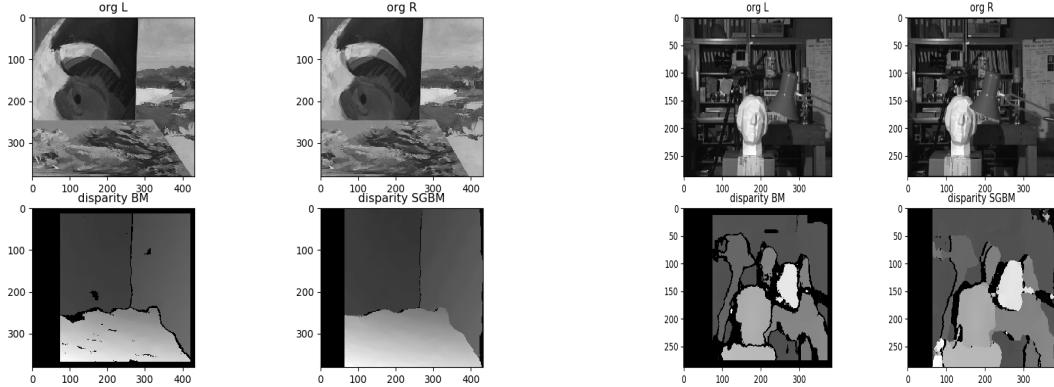


Figure 3: Stereo matching disparity 좌 : Bull image, 우 : Tsukuba image

Figure 2는 이상적인 disparity를 나타내는 ground truth라고 할 수 있고 Figure 3는 opencv 내장함수를 활용한 좌, 우 stereo image에 대한 disparity image이다. Sample끼리 비교한다면 Tsukuba image(우측)의 경우 image 자체가 복잡하고 다양한 물건들이 놓여져 있기 때문에 Bull image에 비해 disparity image가 비교적 복잡하게 나온 것을 알 수 있다. 기존 방법인 BM과 SGBM을 비교하면 Bull image의 경우 확연히 경계에서의 검은 부분도 많이 없어졌을 뿐만 아니라 중간중간에 hole 역시 없어 SGBM이 더 잘되었다고 할 수 있다. Tsukuba image의 경우도 비슷하게 경계에 검은 영역이 많이 없어지고 스탠드의 목부분도 조금씩 인식하는 것을 확인해봤을 때, SGBM 방식이 BM방식보다 잘된다는 것을 확인할 수 있었다.

References

- [1] H. Hirschmuller, “Stereo processing by semiglobal matching and mutual information,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 328–341, 2008.