

デプスマップ精度改善フィルタのフィードバックによるコストボリュームステレオマッチング

藤田 秀† (学生員) 中村 将大†
松尾 琢也† (学生員) 福嶋 慶繁† (正員)

Cost Volume Stereo Matching by Feedback of Refinement Image
Shu FUJITA†, Student Member,
Masahiro NAKAMURA†, Nonmember,
Matsuo TAKUYA†, Student Member, and
Norishige FUKUSHIMA†, Member

† 名古屋工業大学
Nagoya Institute of Technology

あらまし 本論文では、ステレオマッチングで推定、およびエッジ保持平滑化フィルタで改善した視差画像を、ステレオマッチングにおけるマッチングコストに反映することで、高効率にステレオマッチングを行う。結果、提案手法は従来手法に比べて高効率であることを示した。

キーワード ステレオマッチング, 視差画像, 精度改善, エッジ保持平滑化フィルタ

1. まえがき

ステレオマッチングにおいて、左右画像の対応関係を表した3次元のマッチングコストをエッジ保持平滑化フィルタ (Edge-Preserving Filtering: EPF) で平滑化することで、高精度な視差推定が可能となる [1], [2]. しかし、この手法は対応候補となる視差数分 EPF を行う必要があるため、計算コストが高い。一方、視差画像を推定後に、その視差画像を EPF で精度改善する手法がある [3]. もしマッチングコストの平滑化に単純な移動平均フィルタを用い、推定結果を精度改善すれば、EPF は一度だけ適用すればよい。しかし、この手法は2次元処理による3次元情報の高精度化に相当するため、前者に比べて精度が低下する。

本論文では、精度改善後の視差画像をマッチングコストにフィードバックし、コストを移動平均フィルタで平滑化することで、高効率に視差推定を行う、インループフィードバックマッチングを提案する [4].

2. インループフィードバックマッチング

インループフィードバックマッチング (In-loop Feedback Matching: IFM) では、従来の視差推定処理に、精度改善した視差画像のマッチングコストを構築するコストボリューム構築 (第 2.2.1 節) と、それをフィードバックして反映させるコストボリュームブレンド (第 2.2.1 節) の二つの処理を加える。IFM の

全体の流れを図 1 に示す。以下で、視差推定処理とフィードバック処理を説明する。

2.1 視差推定処理

マッチングコスト計算は、入力 of 左画像 I_l と右画像 I_r の絶対値差分 (Absolute Difference: AD) と、 I_l, I_r に対する、水平方向のソーベルフィルタから計算されるエッジ画像 $\nabla I_l, \nabla I_r$ の AD を用いる。つまり、ある画素 \mathbf{p} のマッチングコスト C は、以下のよう求められる。

$$C(\mathbf{p}, d) = \alpha \cdot C_A(\mathbf{p}, d) + (1 - \alpha) \cdot C_S(\mathbf{p}, d) \quad (1)$$

$$C_A(\mathbf{p}, d) = \min(|I_l(\mathbf{p}) - I_r(\mathbf{p} - d)|, T) \quad (2)$$

$$C_S(\mathbf{p}, d) = \min(|\nabla I_l(\mathbf{p}) - \nabla I_r(\mathbf{p} - d)|, T) \quad (3)$$

ここで、 d は取り得る視差値、 α はブレンド比、 C_A, C_S はそれぞれ、左右画像と左右のエッジ画像におけるマッチングコストである。コスト C_A, C_S は閾値 T で打ち切られる。

コスト集約では、計算されたコストを移動平均フィルタを用いて集約しており、以下の式で求められる。

$$\bar{C}(\mathbf{p}, d) = \sum_{\mathbf{q} \in N(\mathbf{p})} w(\mathbf{p}, \mathbf{q}) C(\mathbf{q}, d) \quad (4)$$

ここで、 $N(\mathbf{p})$ は注目画素 \mathbf{p} の近傍画素 \mathbf{q} の集合、 \bar{C} は集約されたマッチングコストであり、 $w(\mathbf{p}, \mathbf{q})$ は重み関数を表す。移動平均フィルタは近傍画素の平均値を出力するため、提案手法では任意の \mathbf{p}, \mathbf{q} に対して $w(\mathbf{p}, \mathbf{q}) = 1$ である。なお、 $w(\mathbf{p}, \mathbf{q})$ にエッジ保持平滑化フィルタの重みを適用すれば、文献 [1], [2] のようにコストボリュームフィルタリングとして展開することができる。

視差計算では、WTA (Winner-Takes-All) 方式を用いて、以下の式のように計算される。

$$D(\mathbf{p}) = \arg \min_d \bar{C}(\mathbf{p}, d) \quad (5)$$

ここで、 D は視差画像である。この際、式 (5) に代わってマッチングコストに対して、隣接した画素は滑らかに連結するなどの拘束条件を満たすように最適化を行うことで、高い精度で視差推定をすることが可能である。この最適化処理の代表例としては、確率伝播法 [5] やグラフカット [6] を用いた手法などが挙げられる。しかし、最適化手法は計算コストが高いことで知られており、提案手法は実行速度を考慮するために最適化処理を行わない。

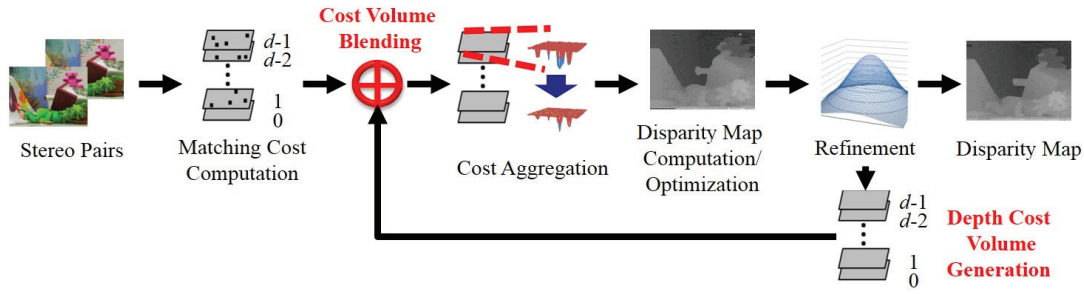


図 1: インループフィードバックマッチングの処理の流れ
Fig. 1 Procedure of proposed method.

視差画像の精度改善では、重み付きジョイントバイラテラルフィルタとジョイントニアレストフィルタ [3] を用いる。この精度改善された画像をフィードバック処理に用いる。

2.2 フィードバック処理

2.2.1 コストボリューム構築

精度改善した視差画像の情報をマッチングコスト計算にフィードバックするために、コストボリューム構築を行う。構築されるコストボリューム C_F は、以下の式で表される。

$$C_F(\mathbf{p}, d) = \frac{1}{\tau^2} \min((d - D'(\mathbf{p}))^2, \tau^2) \quad (6)$$

ここで、 D' は精度改善した視差画像、 τ は打ち切り定数である。

2.2.2 コストボリュームブレンド

コストボリュームブレンドでは、式 (1) で計算された初期のマッチングコスト C に、式 (6) で構築されたコストボリューム C_F を反映させる。この際、それぞれのコストを以下の式のようにして計算する。

$$C'(\mathbf{p}, d) = \beta \cdot C(\mathbf{p}, d) + (1 - \beta) \cdot C_F(\mathbf{p}, d) \quad (7)$$

ここで、 C' はフィードバック後のマッチングコスト、 β はブレンド比である。なお、本論文では、ブレンド比を 1:1 にしている ($\beta = 0.5$)。

3. 実験結果

実験では、提案手法と EPF をコストに適用した手法との比較を行う。EPF として、バイラテラルフィルタ (BF) [1]、ガイドフィルタ (GF) [2]、ドメイントランスフォームフィルタ (DTF) [7] を使用する。実験では、ステレオデータセットである “Teddy” (画像サイズ 450×375 、視差の候補数 64) を使用する。精度評価には、視差画像の誤り率 [8] を用いる。出力

された視差画像 D の誤り率 B は、正解視差画像 D_{GT} を用いて次の式から求められる。

$$B = \frac{1}{|\Omega|} \sum_{\mathbf{p}} f(D(\mathbf{p}), D_{GT}(\mathbf{p})) \quad (8)$$

$$f(x, y) = \begin{cases} 1 & (|x - y| > \varepsilon) \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

ここで、 Ω は画像全体の画素集合、 $|\Omega|$ は画像の全画素数、 ε は閾値である (本論文では $\varepsilon = 1$)。

縦軸を誤り率、横軸を処理時間とし、フィルタの各パラメータを変動させて作成した散布図を図 2 に示す。図 2 は各手法の比較であり、IFM については繰り返し回数による違いを表している。図 2 より、IFM の繰り返しの回数が 1, 2 回のとき、他の手法に比べて、高速かつ高精度である。図 3 は、IFM により推定した視差画像の結果である。図 3 (a) は繰り返しなし、図 3 (b) は 4 回繰り返した結果であり、繰り返し回数が多い方が、物体の輪郭が改善されることがわかる。また、図 2 と図 3 から、IFM の繰り返し回数を増やすことで、誤り率が減少し、推定視差画像の精度が向上することがわかる。

4. まとめ

本論文では、高速かつ高精度にステレオマッチングを行う手法を提案した。実験では、提案手法は、EPF をコスト関数に適用した手法よりも、高速かつ誤り率が低いことを示した。

なお、今回の実験では視差推定処理の実装に、比較手法 [1], [2], [7] にとって効率的な、視差ごとにコストスライスを構築し、それに対してフィルタするという方法を用いた。しかし、提案手法のようにコスト集約に移動平均フィルタを用いる場合、ある視差に対応する左右のブロックに対して、SAD (Sum of Absolute

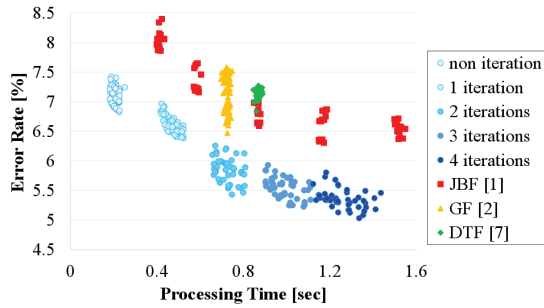
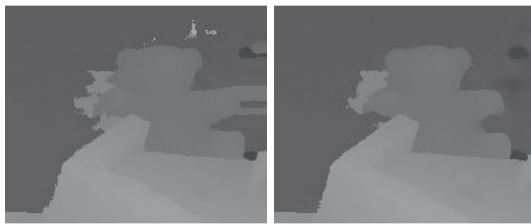


図 2: 誤り率と計算時間の結果.

Fig. 2 Results of error rate and processing time.



(a) Without feedback

(b) 4 iterations

図 3: IFM の出力結果.

Fig. 3 Estimated depth map by IFM.

Difference) を適用するブロックマッチング実装の方が計算効率は向上する. そのため, 今後の課題としては, 提案手法の更なる高速化のために, SAD を用いた実装や, GPU による実装が挙げられる.

謝辞 本研究は, 科研費 15K16023 による.

文 献

- [1] K.J. Yoon and I.S. Kweon, "Adaptive support-weight approach for correspondence search," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.28, no.4, pp.650–656, 2006.
- [2] C. Rhemann, A. Hosni, M. Bleyer, C. Rother, and M. Gelautz, "Fast cost-volume filtering for visual correspondence and beyond," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.35, no.2, pp.504–511, 2013.
- [3] T. Matsuo, N. Fukushima, and Y. Ishibashi, "Weighted joint bilateral filter with slope depth compensation filter for depth map refinement," *International Conference on Computer Vision Theory and Applications*, pp.300–309, 2013.
- [4] T. Matsuo, S. Fujita, N. Fukushima, and Y. Ishibashi, "Efficient edge-awareness propagation via single-map filtering for edge-preserving stereo matching," *Proc. SPIE*, pp.93930S–93930S–8, 2015.
- [5] J. Sun, N. Zheng, and H.Y. Shum, "Stereo matching using belief propagation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.25, no.7, pp.787–800, 2003.
- [6] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.23, no.11, pp.1222–1239, 2001.
- [7] E.S.L. Gastal and M.M. Oliveira, "Domain transform for edge-aware image and video processing," *ACM Trans. on Graphics*, vol.21, no.3, pp.69:1–69:12, 2011.
- [8] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol.47, no.1, pp.7–42, 2002.

(平成 xx 年 xx 月 xx 日受付)

Abstract In this paper, we propose an efficient stereo matching framework. We iteratively process the following steps: cost aggregation, depth map refinement, and feedback of the refined depth map. Experimental results show that our framework can estimate depth maps faster and more accurate than previous methods.

Key words stereo matching, depth map, depth map refinement, edge-preserving filter