

画素単位の大域的な木に基づくステレオマッチング

高瀬竜一[†] 西卓郎[†] 吉見隆[†] 河井良浩[†]

ステレオビジョンのための、対応コストの大域的かつ効率的な集約方法を提案する。一般に動的計画法にもとづくステレオマッチングアルゴリズムは、少ない計算量で良い結果が得られる。その一方で、線状のアーティファクトとして知られるノイズが発生することがある。提案手法は、対応コストの大域的な集約と動的計画法による対応点探索を統合する。最適化されたコストを、木構造によって画像全体から注目画素に向かって集約し、これにより線状のアーティファクトを抑制しつつ、効率的なステレオマッチングを実現した。実験結果によって本手法の有効性を示す。

Pixel-wise Global Tree based Stereo Matching

RYUICHI TAKASE[†] TAKAO NISHI[†]
TAKASHI YOSHIMI[†] YOSHIHIRO KAWAI[†]

We propose the global and effective aggregation method of matching cost for stereo vision. Dynamic programming based stereo matching algorithms produce fine results with low computational complexity. On the other hand, it may produce noises known as streaking artifacts. The proposed method integrates the global aggregation of matching cost with the dynamic programming based correspondence search. The optimized costs are aggregated from the whole image to the current pixel. Thus it suppresses streaking artifacts and achieves the efficient stereo matching. The effects of this method are explained through experiments.

1. はじめに

近年、3D ビジョンがエンターテインメント分野で普及し、距離データのアプリケーションが身近になっている。それら民生用の3D ビジョンには安価という長所がある。しかし供給が流行に左右されやすく、工業用に比べて誤差や個体差が大きい [1]。我々は距離データの歪みを補正した RGB-D カメラを用いて、ロボットによる工業部品のピッキングを試みている。しかし距離データの改善に加えて、データの解析にも技術を要する [2]。一つの対策は、誤差の小さい 3D ビジョンを、入手と交換が容易な機材で自ら組み上げることである。ハードウェアの構築は今後の課題とし、まず我々は 3D ビジョンのアルゴリズム開発に臨んだ。本稿では、アクティブ方式のステレオビジョンに適用できるステレオマッチングアルゴリズムを提案し、標準画像による定量評価と実データでの実験結果を示す。

広く知られているステレオマッチングの体系に従うと、アルゴリズムは次の四つのステップに分類できる [3]。

- 対応コストの計算 (Matching cost computation)
- 集約 (Aggregation)
- 最適化 (Optimization)
- 詳細化 (Refinement)

対応コストは注目画素と対応点候補の非類似性である。輝度の絶対差や輝度差の二乗がよく知られている [4, 5]。提案手法は輝度の 1 次微分と標準化係数 [6] を併用して、金属

部品の光沢に頑健な対応コストを実現した。

集約は対応コストの確からしさの向上である。これは局所的な手法と大域的な手法の二つに大別される。伝統的な局所的な手法は注目画素の周りに小さなウィンドウを設けて、その中の対応コストを合計する。この計算は高速だが、奥行きの不連続の境界があいまいになる短所がある [7, 8]。Graph Cut [9] や Belief Propagation [10] に代表される大域的な手法は広い範囲の対応コストを集約する。これらの手法は誤差が小さいものの、計算に多くの時間と資源を要する。

Semi-Global Matching [11] (SGM) は良いコストパフォーマンスを示す。すなわち誤対応が少なく、効率よく計算できる。このアルゴリズムは Scanline Optimization [3] の、多方向への自然な拡張であり、輝度による領域分割などを含まない。よってアクティブ方式の計測と相性が良く、対象にランダムドットパターンを投影しても問題を起こしにくい。しかし不完全な大域的な手法であるため、特有のノイズを生じることがある。そこで提案手法は、木構造に基づいて大域的に、かつ SGM と同程度に効率よくコストを集約する。さらにアクティブ方式に適用できる利点も維持する。

また、集約した対応コストの最適化によって視差を求め、詳細化によって最終的な視差画像を得る。これらは既存の手法の組み合わせによって実現した。

本稿で述べる提案手法の、主な利点を次に挙げる。

- (1) **ステレオ画像の輝度の不一致に頑健な対応コスト**
- (2) **大域的かつ効率的な集約方法**
- (3) **アクティブ方式にもパッシブ方式にも適用できる**

[†] 国立研究開発法人 産業技術総合研究所
National Institute of Advanced Industrial Science and Technology (AIST)

2. 微分と標準化係数による対応コスト

提案手法の第一の要素は対応コストの定義である。コスト関数を式 (1) に示す。式中の p は画像座標を、 d は視差を表す。対応コストの従来手法には、輝度の絶対差、LoGフィルタ、輝度の絶対差と水平1次微分の絶対差の組み合わせなどがある[12, 13]。我々は標準化係数フィルタ[6]を開発し、式 (2) のように水平1次微分フィルタと組み合わせ、標準画像[3]による実験で良好な結果を得た。

$$C(p, d) = \min(C'(p, d), \tau) \quad (1)$$

$$C'(p, d) = \alpha \cdot C_{BT}^{grad}(p, d) + (1 - \alpha) \cdot C_{BT}^Z(p, d) \quad (2)$$

ここで α は、2種類のコストの混合比を表す重み係数である。この重みは、後述する微分コストと標準化係数コストのバランスを画像に応じて調整する。また τ は対応コストの上限值である。このしきい値は、隠れなどがコストを過大にする場合にその値を切り詰める。

微分コストは、輝度 I について式 (3) から得られる微分値 ∇ に基づく。輝度はステレオ画像をなすので、微分値も左右二つの画像が得られる。

$$\nabla(p) = I(u+1, v) - I(u-1, v) \quad (3)$$

$$\text{ただし } p = (u, v)$$

二つの微分画像から、式(4)によって微分コストを求める。微分コストは、左右の微分値の絶対差に似ている。ただし一般の絶対差ではなく、画素のサンプリング誤差に対して頑健な Birchfield と Tomasi の関数である。ここで式中の L は左画像を、 R は右画像を、ハット記号は値を補間して求めることを意味する。実装は式 (4) そのままではなく、高速な近似計算を用いた[14]。

$$C_{BT}^{grad}(p, d) = \min\{a_{grad}, b_{grad}\} \quad (4)$$

$$a_{grad} = \min_{u-\frac{1}{2} \leq u' \leq u+\frac{1}{2}} \left| \hat{\nabla}_L(u', v) - \nabla_R(u-d, v) \right|$$

$$b_{grad} = \min_{u-\frac{1}{2} \leq u' \leq u+\frac{1}{2}} \left| \nabla_L(u, v) - \hat{\nabla}_R(u'-d, v) \right|$$

標準化係数 (z 値) は、輝度 I に式 (5) を適用して得られる。標準化係数コストはその非類似性である。 μ_p は点 p を中心とする局所領域の平均輝度を、 σ_p は同じく標準偏差を表す。標準化係数コストは式 (6) で表され、微分コストと同様に Birchfield と Tomasi の関数によって求める。

$$Z(p) = \frac{I(p) - \mu_p}{\sigma_p} \quad (5)$$

$$C_{BT}^Z(p, d) = \min\{a_Z, b_Z\} \quad (6)$$

$$a_Z = \min_{u-\frac{1}{2} \leq u' \leq u+\frac{1}{2}} \left| \hat{Z}_L(u', v) - Z_R(u-d, v) \right|$$

$$b_Z = \min_{u-\frac{1}{2} \leq u' \leq u+\frac{1}{2}} \left| Z_L(u, v) - \hat{Z}_R(u'-d, v) \right|$$

従来手法に対する提案手法の主な利点は、鏡面反射に対する頑健性にある。標準化係数コストは、輝度の平均値および分散の違いに対して頑健な性質を持つ[6]。この性質は、対象物の表面が鏡面反射成分を含む場合に有利に働く。これに対して、輝度の絶対差は輝度の平均値の違いに敏感になる。また微分の絶対差は輝度の分散の違いに対して敏感になる。提案手法の効果は後出の実データの実験で示す。

3. 画素単位の大域的な木

提案手法の第二の要素は、対応コストの集約である。ステレオ対応問題において、線状のアーティファクト (streaking artifact) として知られるノイズがある[15]、これは隣り合う画素列の間に視差の不整合が起きる現象であり、動的計画法に基づく手法にしばしば見られる。Scanline Optimization はこのノイズを生じる典型的な手法の一つである。これらの手法は比較的に計算量が小さいという利点があり、いくつかの改良が提案されている。線状のアーティファクトを抑える一つの方法は、大きい輝度変化を避けてコストを集約することである[16, 17]。線状のアーティファクトは、集約あるいは最適化する経路が奥行きの不連続をまたがるときに起きやすい。そこで輝度のエッジ付近に奥行きの不連続があると仮定して、エッジを避けるように経路を決定する手段が有効になる。しかし計測対象にランダムドットパターンを投影する場合は仮定が成り立たないため、期待された通りにならない。

別の方法は、一つの注目画素に対して多方向からコストを集約することである。Scanline Optimization は走査線と平行にコストを集約するので、線状のアーティファクトも走査線方向に伸びる。それに対して SGM は多方向から集約する[11]。特定の方向に奥行きの不連続があっても、他の方向の集約がステレオ対応を正しい視差に誘導し、ノイズが線状に伸びる現象を抑制する。対応コストを集約する方向は一律に決まっており、輝度のエッジに左右されない。

しかし SGM は、画像を網羅するほど大域的ではない。注目画素から伸びる一部の画素列だけが、コストの集約に反映される。その経路は直線的で隙間がある。そのため、まれに線状のアーティファクトを生じる (図 1)。

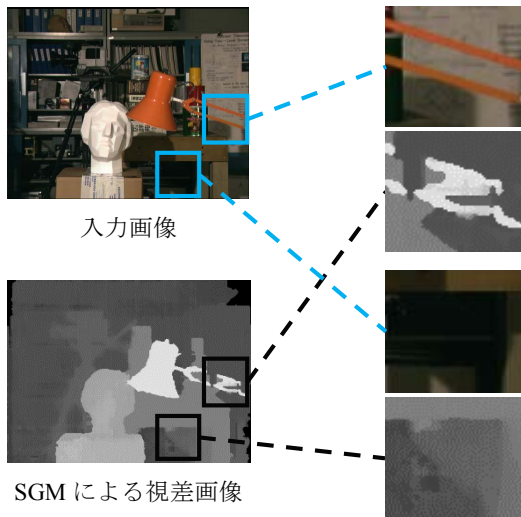


図 1 線状のアーティファクトの例. 左側の 2 画像は Middlebury Stereo Evaluation[21]から転載した. 右側の 4 画像は問題の部分拡大したもの.

提案手法は注目画素ごとに画像を網羅する. これには図 2 の木構造を用いる. 図中の u は画像の水平座標の, v は垂直座標の軸方向を表す. 太枠は注目画素を, 矢印は集約の順序を表す. この木によって対応コストの集約から隙間をなくし, 線状のアーティファクトを抑制する.

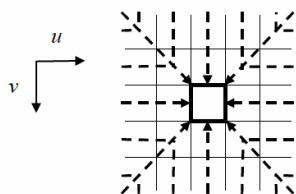


図 2 画素単位の大域的な木. 図中の太枠は注目画素, u の矢印は水平, v は垂直の座標軸を表す.

以降の説明では, 木の 1/4 にあたる左側の部分木を考える. 上下および右側の部分木も, 方向の違いを除いてアルゴリズムは同じである. 図 3 において水平に伸びる破線の左方向を, 左側の部分木の「主方向」と定義する. 同じく斜めに伸びる破線方向二つを, この部分木の「副方向」と定義する. 提案手法は, 1 画素ごとに何度も画像を網羅するのではなく, 方向ごとに画像を 1 回ずつ走査する.

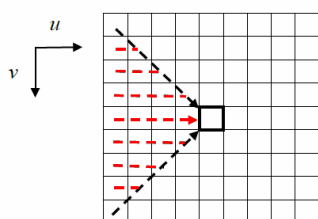


図 3 左側の部分木

ここで, 注目画素の周囲 8 方向に 0 から 7 の番号を割り当てる. 番号は図 4 のように反時計周りの順とする. 方向を x とするとき, O_x は x 方向で注目画素に隣接する画素の相対座標とする. これは式 (7) によって定義される.

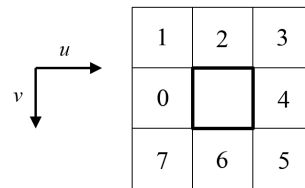


図 4 方向の定義. 番号によって注目画素から見た方向を表す. 太枠が注目画素.

$$O_x = \begin{cases} (-1,0) & : x=0 \\ (-1,-1) & : x=1 \\ \vdots & \\ (-1,1) & : x=7 \end{cases} \quad (7)$$

3.1 主方向コスト

画素ごとに, 木の枝に沿って画像を網羅すると計算コストが大きい. 計算式の再帰的な定義によって, 全ての木の, 特定方向の枝の処理が 1 回の画像走査ですむ. ここで主方向の集約で求める値を主方向コストとする. 主方向コストを式 (8) によって定義する. ただし座標 $p+O_q$ が画像の外であるとき, 式 (8) を $S_q(p, d) = C(p, d)$ で代えるものとする.

$$S_q(p, d) = C(p, d) + \min_{d' \in D'} (S_q(p + O_q, d') + w(d, d')) \quad (8)$$

$$w(d, d') = \begin{cases} 0 & : d = d' \\ P_1 & : |d - d'| = 1 \\ P_2 & : \text{otherwise.} \end{cases} \quad P_2 \geq P_1 \geq 0$$

q : 主方向 $q \in Q$

主方向の集合 $Q = \{0, 2, 4, 6\}$

D' : 視差の選択枝の集合 $D' = \{d-1, d, d+1, d_q\}$

D : 視差の探索範囲 $D' \subseteq D$

直前の最適な視差 $d_q = \arg \min_{b \in D} (S_q(p + O_q, b))$

ただし, $w(d, d')$ は視差の不連続に対するペナルティ関数である.

図 5 に主方向コストの計算を模式的に示す. 主方向コストが左方向の場合であり, 処理は矢印の方向に進める. 図中の太枠の画素に注目すると, 式 (8) より主方向コストは, 同じく注目画素の対応コストと, その直前 (この場合は左隣) の画素の主方向コストより求まる. これは注目画素と同じ水平座標の縦一列, 全ての画素について同様である. よって, 矢印方向に対して直前の縦一列の主方向コストが求めれば, 次の縦一列について式 (8) が計算可能となる.

これにより全ての画素の、左の主方向コストが1回の画像走査で求まる。

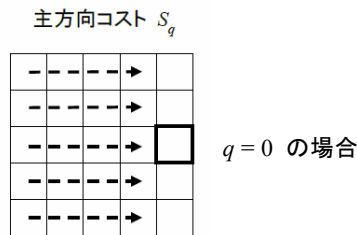


図 5 主方向コストの計算

3.2 副方向コスト

算出された主方向コストをもとに、副方向の集約を行う。この過程で求まる値を副方向コストとし、式(9)で定義する。式(8)と異なり、対応コストの代わりに主方向コストを加算する。ただし座標 $p+O_r$ が画像の外ならば、式(9)を $S_r(p, d) = S_q(p, d)$ で代えるものとする。

$$S_r(p, d) = S_q(p, d) + \min_{d' \in D''} (S_r(p + O_r, d') + w(d, d')) \quad (9)$$

r : 副方向 $r \in R_q$

副方向の集合

$$R_q = \{r \mid r \in \{q+1, q-1, q+7\} \wedge 0 \leq r \leq 7\}$$

D'' : 視差の選択枝の集合 $D'' = \{d-1, d+1, d_r\}$

直前の最適な視差 $d_r = \arg \min_{b \in D} (S_r(p + O_r, b))$

図6は副方向コストの計算の模式図である。左の主方向一つに対し、副方向は斜めに二つ存在する。主方向コストと同様に、左隣の縦一列の副方向コストが求まれば、注目画素と同じ水平座標の縦一列における式(9)が計算可能となる。副方向への計算は、それらが属する主方向とともに、同じ繰り返し処理の中で計算できる。よって全ての主方向コストと副方向コストを求めるには、画像の走査を主方向の数、すなわち4回だけ行えばよい。

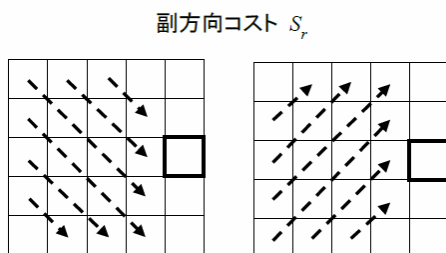


図 6 副方向コストの計算

3.3 集約コスト

次に主方向コストと副方向コストを集約する。図7の太枠の画素において、図左の実線の矢印で示される一つの主方向コストと二つの副方向コストを合計する。これにより図右の破線で示すような、左側の部分木の対応コストが集約される。上下および右側の部分木も同様に合計すれば、注目画素に集約されるコストは画像を網羅できる。

主方向コストと副方向コストの合計

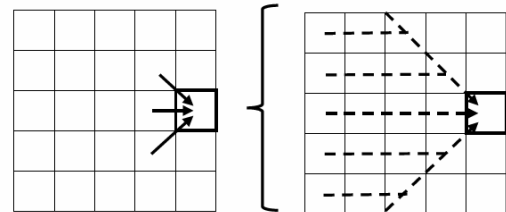


図 7 部分木について集約されたコスト

集約コストの計算を式(10)に示す。全ての主方向 q の、全ての副方向 r のコストが合計される。二つの副方向コストが同じ主方向コストを含むので、重複する主方向コスト1個を減算する。ここで n は主方向の数である。また、主方向コストの式中对対応コストが現れる。主方向コストを n 個合計するので $n-1$ 個の対応コストを減算する。これらの減算は、集約コストを整数のデータ型で実現するとき、加算によって値が容易に桁あふれしないように働く。

$$S(p, d) = \sum_{q \in Q} \left(\sum_{r \in R_q} S_r(p, d) - S_q(p, d) \right) - (n-1)C(p, d) \quad (10)$$

以上の処理で集約コストが求まる。集約コストは、次節で述べる視差の計算に用いられる。

SGMの場合、方向の数に関わらず画像を2回走査すれば、コストを集約できる。これに対し、提案手法は主方向の数だけ画像を走査しなければならない。しかし主方向が四つで画像を網羅できるので、計算量は定数倍である。画像の幅を W 、高さを H 、視差の範囲を D とするとき、提案手法の計算オーダは $O(WHD)$ となり、SGMと同様に効率よく計算できる。

4. 視差の計算

視差の計算はWTA (Winner Takes All) を用いる。式(11)によって求めた画素 p の最適解が視差 d_p となる。これにより、網羅的に集約した対応コストの合計がなるべく小さくなるように、そして隣り合う視差がなるべく滑らかに変化するように注目画素の視差が選ばれる。一連のアルゴリズムは、動的計画法の原理に基づいている。

$$d_p = \arg \min_{d \in D} (S(p, d)) \quad (11)$$

次に誤対応を削除する。算出された視差画像とコストを基に、対応の誤りが疑われる画素に削除マークを付ける。この処理は主に、一意性[18]に違反する視差と、周囲から突出して大きい（または小さい）小領域をなす視差の削除から構成される。

次に視差画像を適度に整え、詳細化する。まず視差画像に付けられた削除マークの領域に対して穴埋め処理を行う。そして視差をサブピクセル化し、画素単位より高い分解能で視差を推定する。視差の推定には等角フィッティングを用いた[19]。さらに対象が標準画像の場合は、視差画像にエッジ保存平滑化を施した。そのアルゴリズムには重み付きメディアンフィルタ[20]を採用した。これは提案手法を構成する中で唯一、輝度のエッジに依存する要素である。ランダムドットパターンを投影する場合は、視差に対するメディアンフィルタで代替する。

5. 実験

5.1 標準画像による実験

Middlebury ベンチマーク[21]によって、提案手法を評価した。提案手法の実装には C 言語と SIMD 命令 (SSE4.1) を用いた。またプログラムはシングルスレッドとした。

実験は 3.3GHz の Core i7 CPU と 8GB のメインメモリを搭載した PC で行った。標準画像の処理に要した時間は Tsukuba が 0.38 秒, Venus が 0.54 秒, Teddy が 1.13 秒, Cones が 1.9 秒であった。

図 8 に入力画像、真値、計算結果を示す。(a)列は入力されたステレオ画像の左画像であり、(b)は視差画像の真値、(c)は SGM による視差画像、(d)は提案手法による視差画像である。提案手法に用いたパラメータは、Tsukuba ($\tau=15$, $\alpha=1$, $P_1=6$, $P_2=6$), Venus ($\tau=15$, $\alpha=0.9$, $P_1=7$, $P_2=8$), Teddy ($\tau=10$, $\alpha=0.9$, $P_1=6$, $P_2=6$), Cones ($\tau=14$, $\alpha=0.7$, $P_1=5$, $P_2=6$) である。従来手法(c)は画像 Tsukuba のランプの支柱や机の脚に線状のアーティファクトが発生している。これに対し、提案手法(d)では線状のアーティファクトの発生が抑えられている。

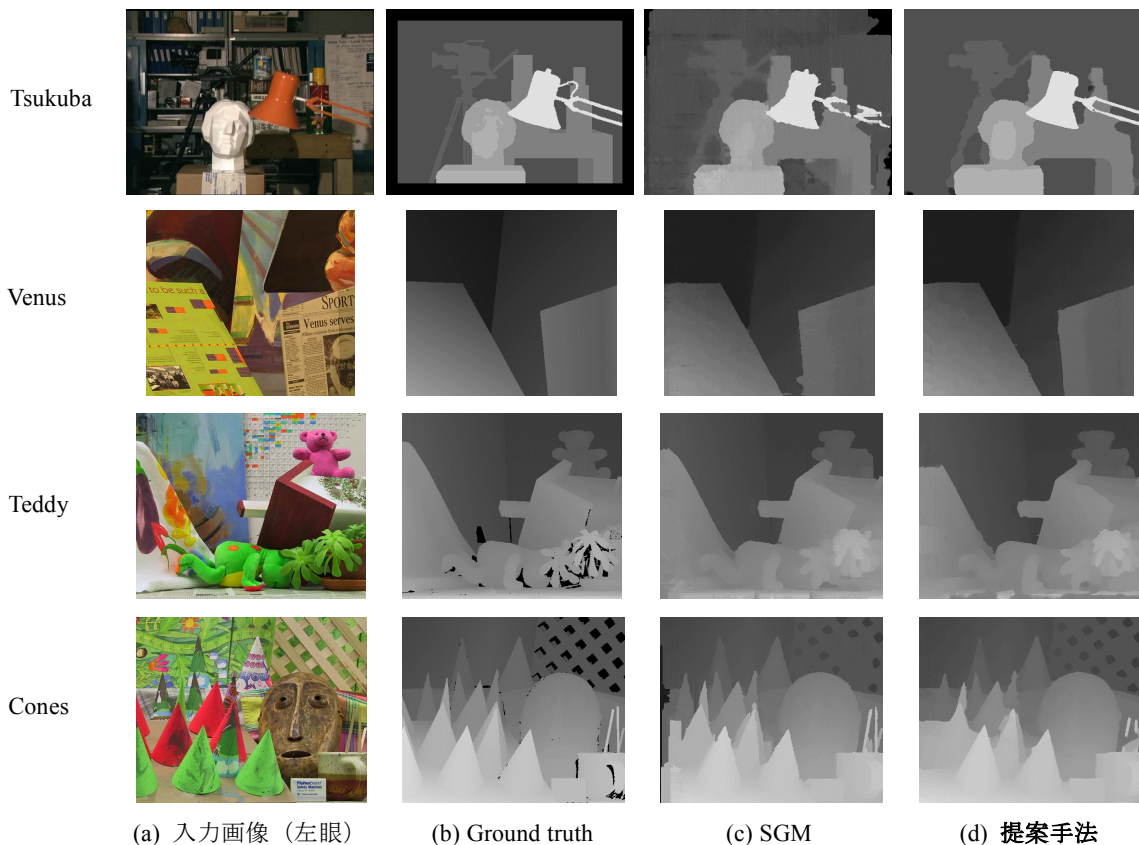


図 8 入力画像および視差画像

Error Threshold = 1		Sort by nonocc			Sort by all			Sort by disc			Average Percent Bad Pixels			
Algorithm	Avg.	Tsukuba ground truth			Venus ground truth			Teddy ground truth				nonocc	all	disc
	Rank	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc				
IGSM [159]	10.5	0.93	1.37	5.05	0.07	0.17	1.04	4.08	5.98	11.4	2.14	6.97	6.27	3.79
LCU [156]	12.4	1.06	1.34	5.50	0.07	0.26	1.03	3.68	9.95	10.4	1.63	6.87	4.82	3.89
PUTv3 [56]	87.1	1.77	3.86	122.9	0.42	0.95	101.5	7.02	14.2	114.18	2.40	9.11	6.56	6.64
GradAdaptWgt [53]	87.8	2.26	2.63	8.99	0.99	1.39	113.4	8.00	13.1	96.18	2.61	7.67	7.43	6.55
YOUR METHOD	88.9	1.67	2.13	8.98	0.38	0.62	5.18	6.97	11.1	19.5	3.94	9.58	11.3	6.77
RT-ColorAW [91]	89.2	1.40	3.08	105.5	0.72	1.07	118.3	6.69	14.0	112.15	4.03	11.9	10.2	6.55
AdaptWeight [12]	90.0	1.38	1.85	6.90	0.71	1.06	119.6	7.88	13.3	101.18	3.97	9.79	8.26	6.67
SegTreeDP [21]	90.0	2.21	2.76	10.3	0.46	0.95	112.7	9.58	12.9	15.2	3.23	7.86	8.83	6.82
InteriorPILP [33]	90.0	1.27	1.62	6.82	1.15	1.23	117.1	8.07	11.2	11.9	3.92	9.68	9.62	7.26
TwoStep [142]	95.3	2.91	3.68	120.1	0.27	0.45	53.2	7.42	9.4	12.6	4.09	10.1	10.3	7.14
ImproveSubPix [24]	95.8	3.00	3.61	117.1	0.88	1.12	147.1	7.12	9.2	12.4	2.96	8.22	8.55	6.90
RealTimeABW [73]	95.8	1.26	1.67	6.83	0.33	0.78	6.56	10.7	14.1	18.3	4.81	11.9	10.7	7.90
BP+DirectedDiff [54]	95.9	2.90	4.47	131.1	0.65	1.05	120.1	5.07	31.1	14.7	2.94	12.6	12.8	7.29
SemiGlob [6]	99.9	3.26	3.96	123.1	1.00	1.16	115.1	6.02	5.5	12.2	3.06	9.75	8.90	7.50
SDDS [115]	101.7	3.31	3.62	118.1	0.39	0.89	0.76	7.65	10.0	13.0	3.99	10.1	10.8	7.19
FastBilateral [61]	101.8	2.38	2.80	102.1	0.34	0.81	0.92	9.83	13.5	15.3	3.10	9.31	8.59	7.31

図 9 Middlebury Stereo Evaluation – Version2 による評価。
2015年3月2日時点におけるスクリーンショット。

図 9 は、真値との誤差を Middlebury Stereo Evaluation によって定量的に評価し、他の手法と比較した表である。表の各行はアルゴリズムに、縦の中央 4 列は標準画像に対応する。各セルの数値は評価結果である。評価対象の画素は、隠れなし(nonocc)、画像全体(all)、奥行きの不連続付近(disc)の 3 種類に分類される。それぞれの誤対応のパーセンテージが各セルに記載され、その横に小さく順位が示されている。Avg.Rank 列は順位の平均値であり、総合的な順位として扱われている。

太字で「YOUR METHOD」と書かれた行が提案手法の評価結果である。Middlebury Stereo Evaluation に掲載されている 161 件に対し、提案手法は 95 番目となり、順位で従来手法「SemiGlob」(SGM)を上回った。右端の行は誤対応のパーセンテージの平均値であり、提案手法は SGM より誤対応が平均で 0.73%ほど減少している。一方、視差の探索範囲が広い画像 Teddy や Cones において、奥行きの不連続付近の成績が従来手法より劣っている。この結果から、大きな視差変化が提案手法の弱点である事が見て取れる。

5.2 実データによる実験

実際のランダムピッキングでは、表面が無地の工業部品を扱う場合もある。このようなとき、ステレオビジョンで安定した結果を得るには、アクティブ方式の計測手法が必要となる。我々は工業部品にランダムドットパターンを撮影した場合も、提案手法が有効に働くかを実験した。実験

には IDS 社のカメラ Ensenso N10 [22]を用いた。この製品は 3 次元計測の機能を備えているが、今回は単にプロジェクタが付属したステレオカメラとして利用した。

計測の対象物を図 10 に示す。この金属部品は、スプリングワッシャ付き角座金である。形状は主に平面から構成されており、表面は無地で、鈍い金属光沢がある。対象物を撮影したステレオ画像を図 11 に示す。



図 10 計測に用いた金属部品

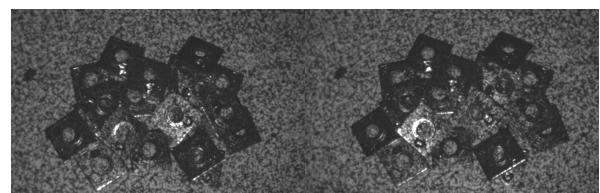


図 11 ランダムドットパターンを投影して撮影したステレオ画像

提案手法と比較するため、従来手法である SGM を実装した。この SGM は水平 1 次微分を対応コストに用いた。SGM の計測結果を 3 次元表示した画面を図 12 に、同じデータに対する提案手法の結果を図 13 に示す。比較のため SGM、提案手法ともに穴埋め処理を行っていない。パラメータは SGM ($\tau=20$, $P_1=15$, $P_2=30$), 提案手法 ($\tau=20$, $\alpha=0.1$, $P_1=15$, $P_2=30$) とした。SGM の P_1 と P_2 は、提案手法の主方向におけるペナルティに相当する。

二つの手法による結果の違いが、光沢が強く現れた部品と、一部が隠された部品に見られる。鏡面反射による光沢は、左右どちらか一方の画像で明るくなって現れる。そのような部品では従来手法のデータの欠損が多い。一方、提案手法では欠損が少ない。これは標準化係数フィルタの、輝度の分散の差に対する頑健性の効果と考えられる。

また一部が隠され、カメラから見える平面が少ない部品がある。それらは光沢が弱い場合でも、提案手法のデータ欠損が従来手法より少ない。その理由として、SGM は集約の経路が見える部分を十分に通らず、網羅的に集約する提案手法の性質が有利に働いたものと考えられる。

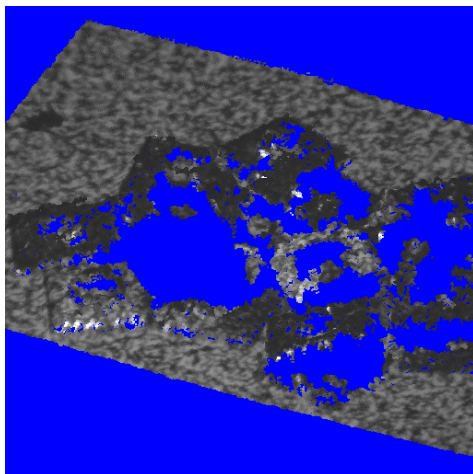


図 12 従来手法による 3 次元復元

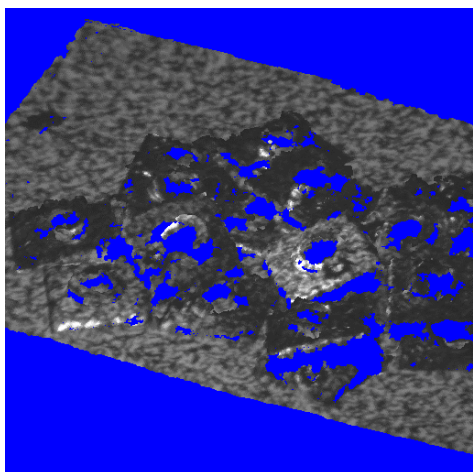


図 13 提案手法による 3 次元復元

6. おわりに

本稿では微分と標準化係数を用いる対応コスト、および木に基づく大域的なステレオマッチングを提案した。ランダムドットパターンを工業部品に投影し、3次元計測して、アクティブ方式のステレオビジョンでも提案手法が有効に機能することを示した。また、標準画像を用いて従来手法と性能を比較した。従来手法より線状のアーティファクトが抑えられ、定量評価において誤差がより小さかった。

ロボットビジョンには速度と正確さが求められる。従来手法はコストパフォーマンスに優れ、その派生技術はアクティブ方式のステレオビジョン製品にも採用されている[22]。しかし工業部品のランダムピッキングでは、金属光沢への頑健性やバラ積み部品に対する計測の安定性が、より求められる。それらの要求に対して、提案手法は有用であると目される。

一方、提案手法は大きな視差変化に対して不利となる可能性も示唆された。標準画像 Cones では、算出された視差画像から細長い棒の形状が一部失われている。定性的に見ても、視差の変化において誤差が大きいことが分る。その原因は、線状のアーティファクトを抑えるための、網羅的な集約方法自体にあると推測される。提案手法は、従来手法よりも奥行き連続性に対する拘束が強くなりすぎるのであろう。今後は、奥行き不連続に対するペナルティを適応的にキャンセルできるモデルについて検討する。

参考文献

- 1) 佐藤雄隆: RGB-D カメラから得られる Depth データの歪み補正. 映像情報 industrial, Vol.46, No.3, pp.66-71, 産業開発機構(2014).
- 2) 西卓郎, 吉見隆, 高瀬竜一, 原田研介, 永田和之, 新良貴陽平, 河井良浩: ビンピッキングのための RGB-D カメラを用いた三次元位置姿勢推定, および把持可能性を考慮したスコアリング手法. 情報処理学会研究報告, Vol.2014-CVIM-194, No.2, pp.1-6 (2014).
- 3) Scharstein, D. and Szeliski, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. International Journal of Computer Vision, Vol.47, No.1-3, pp.7-42 (2002).
- 4) Kanade, T.: Development of a video-rate stereo machine. In Proc. of Image Understanding Workshop, pp.549-557 (1994).
- 5) 金出武雄, 蚊野浩, 木村茂, 川村英二, 吉田収志, 織田和夫: ビデオレートステレオマシンの開発. 日本ロボット学会誌, Vol.15, No.2, pp.261-267 (1997).
- 6) 西卓郎, 高瀬竜一, 吉見隆, 河井良浩, 富田文明: 領域ベースのステレオ対応点探索のための「標準化係数フィルタ」. 計測自動制御学会論文集. Vol.50, No.10, pp.693-702 (2014).
- 7) 岡田慧, 加賀美聡, 稲葉雅幸, 井上博允: PC による高速対応点探索に基づくロボット搭載可能な実時間視差画像・フロー生成法と実現, 日本ロボット学会誌, Vol.18, No.6, pp.896-901 (2000).
- 8) 岡撰子, 片山保宏, 奥富正敏: 領域ベースステレオマッチングにおけるバウンダリオーバーリーチの解析と定量化. 情報処理学会研究報告, 2000-CVIM-123, Vol. 2000. No. 82. pp. 17-24 (2000).
- 9) Boykov, Y., Veksler, O. and R. Zabih. Fast approximate energy minimization via graph cuts. IEEE Trans. on PAMI, Vol.23, No.9, pp.1222-1239 (2001).
- 10) Yang, Q., Wang, L., Yang, R., Stewenius, H. and Nister, D.: Stereo matching with color-weighted correlation, hierarchical belief propagation and occlusion handling. IEEE Trans. on PAMI, Vol.31,

No.3, pp.492–504 (2009).

- 11) Hirschmuller, H.: Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information. IEEE Conf. on CVPR2005, pp.807–814 (2005).
- 12) Hirschmuller, H., Scharstein, D.: Evaluation of Cost Functions for Stereo Matching. IEEE Conf. on CVPR2007, pp.1–8 (2007).
- 13) Zhang, K., Fang, Y., Min, D., Sun, L., Yang, S., Yan, S. and Tian, Q.: Cross-Scale Cost Aggregation for Stereo Matching. IEEE Conf. on CVPR 2014, pp.1590-1597 (2014).
- 14) Birchfield, S. and Tomasi, C.: Depth discontinuities by pixel-to-pixel stereo. Sixth IEEE International Conference on Computer Vision, pp.1073-1080 (1998).
- 15) Bobick, A.F. and Intille, S.S.: Large occlusion stereo. IJCV, Vol.33, No.3, pp.1–20 (1999).
- 16) Veksler, O.: Stereo correspondence by dynamic programming on a tree. IEEE Conf. on CVPR 2005, Vol.2, pp.384-390 (2005).
- 17) Yang, Q.: A Non-Local Cost Aggregation Method for Stereo Matching, IEEE Conf. on CVPR 2012, pp.1402-1409 (2012).
- 18) Marr, D. and Poggio, T.A.: Cooperative computation of stereo disparity. Science, Vol.194, No.4262, pp.283–287 (1976).
- 19) 清水雅夫, 奥富正敏: 画像のマッチングにおけるサブピクセル推定の意味と性質. 電子情報通信学会論文誌 D-II, Vol.J85-D-II, No.12, pp.1791-1800 (2002).
- 20) Ma, Z., He, K., Wei, Y., Sun, J., Wu, E.: Constant time weighted median filtering for stereo matching and beyond. In Proc. of ICCV2013, pp.49-56 (2013).
- 21) Scharstein, D. and Szeliski, R.: Middlebury Stereo Evaluation - Version 2, The Middlebury Computer Vision Pages (online), available from <http://vision.middlebury.edu/stereo/eval> (accessed 2015-03-02).
- 22) Ensenso GmbH : Ensenso Stereo 3D Cameras (online), available from <http://www.ensenso.de/> (accessed 2015-04-15).