

重み付きクロスバイラテラルフィルタによる奥行き推定精度の向上

Depth Map Refinement with Weighted Cross Bilateral Filter

松尾 琢也[†], 正会員 福嶋 慶繁[†], 正会員 石橋 豊[†]

Takuya Matsuo[†], Norishige Fukushima[†] and Yutaka Ishibashi[†]

Abstract In this paper, we propose a refinement filter for depth maps. The filter convolutes an image and a depth map with a cross-computed kernel. We call the filter a weighted cross bilateral filter. The main advantages of the proposed method are the filter fits outlines of objects in the depth map to silhouettes in the image, while the filter reduces Gaussian noise in other areas. Additionally, its computational cost is independent of depth ranges. Thus, we can obtain accurate depth maps at the lower cost than the conventional approaches, which require Markov random field-based optimization methods. Experimental results show that the accuracy of the depth map in edge areas goes up and its running time is low.

キーワード：クロスバイラテラルフィルタ, ステレオマッチング, 奥行き画像, 奥行き画像改善フィルタ, ポストフィルタリング

1. ま え が き

近年, Microsoft の Kinect* や ASUS の Xtion** といったような消費者レベルでのデプスセンサが発売され, 奥行き画像を用いた画像処理技術が注目を集めている. 例えば, 姿勢評価, 物体検出, 自由視点画像合成といったものが挙げられる. 特に, 自由視点画像合成においては高い精度の奥行き画像が要求される. 自由視点画像は原画像と奥行き画像を用いて Depth Image Based Rendering (DIBR)¹⁾ によって合成される.

一般に, 奥行き画像はステレオマッチングによって取得される. ステレオマッチングでは, 左右の視点の異なる原画像から各ピクセルにおける対応点を探し出し, その視差によって各ピクセルでの奥行き値を決定する. そのステレオマッチングは四つのステップにより成り立っており, コスト計算, コスト集約, 視差計算と最適化, 奥行き画像の改善というステップで構成されている²⁾. 現在, 主流となっているステレオマッチングは複雑なアルゴリズムにより最適化を行い奥行き画像を求めている. それらは, Markov Random Field (MRF) を基にした最適化を行っている. それらには, Semi-Global Block Matching (SGBM)³⁾⁴⁾, Belief

Propagation⁵⁾, Graph Cuts⁶⁾ といったものが例として挙げられる. これらの問題点は, アルゴリズムの複雑さにより実行時間が長いことである. これは, 計算時間が奥行き値の階調数に依存することも一因である. また, 全体の整合性を重視して奥行き画像を取得するため, 輪郭部分など局所的な部分が曖昧になるという問題がある.

一方, フィルタリングによる奥行き画像の改善によって精度を向上させる手法もある. 例えば, メディアンフィルタによりスパイクノイズを除去するといったことである. フィルタリングによる奥行き画像の改善では, 計算時間が奥行き値の階調数に依存しないものが多いため, 最適化と異なり, 高速な演算を望むことができる. また, 一度求めた Depth Map に対して適用するものであるため, どのような手法で求めた Depth Map に対しても適用可能となる. つまり, ステレオマッチングにより求めた Depth Map から Kinect など取得した奥行き画像まで, 広く適用可能である.

しかし, 従来のノイズ除去等に用いるフィルタとしては簡易なものが多く, 例えば, ガウシアンフィルタやメディアンフィルタ, バイラテラルフィルタは, ノイズは除去できるが輪郭の改善効果はない. メディアンフィルタやバイラテラルフィルタは, エッジを保持するフィルタではあるが, 現状のエッジを保持するのみであり, 物体輪郭を改善することはない. また, ガウシアンフィルタやバイラテラルフィルタなどは, フィルタカーネル内で重み付き平均を行うため, 奥行き値の混合が発生し, 奥行き画像にぼけが生じるという新たな問題も引き起こされる.

2011年8月30日受付, 2012年5月31日再受付, 2012年8月15日採録

[†]名古屋工業大学 大学院 工学研究科

(〒466-8555 名古屋市 昭和区御器所町, TEL 052-732-5113)

[†]Graduate School of Engineering, Nagoya Institute of Technology

(Gokiso-cho Showa-ku Nagoya 466-8555, Japan)

* <http://www.xbox.com/>

** http://event.asus.com/wavi/product/WAVI_Pro.aspx

輪郭の曖昧な奥行き画像を用いた画像処理では様々な弊害が予想される。例えば、奥行き画像を用いた背景分離では、前景と背景を正しく分離できないという問題が考えられる。また、このような奥行き画像を入力として、自由視点画像合成を行った場合、輪郭部分が正確に合成されない恐れがある。そのため、奥行き画像の輪郭の精度が重要となってくる。

そこで本論文では、ノイズ除去と輪郭の補正が可能な奥行き画像の高速な改善フィルタを提案する。ステレオマッチングの精度と速度はトレードオフの関係にあり、本論文では計算速度を重視した手法を検討する。本論文では、ステレオマッチングの中でも高速に動作する Block Matching (BM) を用いて Depth Map を取得する。BM による Depth Map は、平面部分・輪郭部分の精度がともに高くはないが、提案する奥行き画像改善フィルタにより大きく精度が改善可能なことを示す。また、BM 以外にも、高速な MRF 最適化が可能な SGBM や、特徴点ベースの手法 18) に対しても本フィルタにより精度改善が可能なことを示す。

以下では、まず第 2 章で従来の関連研究について説明を行い、それらの問題点と本論文の利点を明らかにする。次に第 3 章で提案する奥行き画像改善フィルタの説明を行う。そして、第 4 章で提案手法を評価する実験方法の説明、第 5 章で実験結果を報告する。最後に、第 6 章で結論となっている。

2. 関連研究

一般に奥行き画像はノイズが多いと言われている。そのため、フィルタを用いたノイズ除去が行われる。例えば、メディアンフィルタやバイラテラルフィルタ 7) が挙げられる。バイラテラルフィルタは、物体の輪郭を維持したまま、平面部分の平滑化が可能なフィルタである。しかし、輪郭維持とノイズ除去の性能は、入力画像に依存する。例えば、入力に多くのノイズがある場合、ノイズ除去の性能は大きく低下してしまう。また、このフィルタによりガウシアンノイズを除去することは可能でも、奥行き画像に現れるようなスパイクノイズを除去することはできない。

このような問題を解決する新しい手法として、クロスバイラテラルフィルタがある。このフィルタでは、同一視点の画像で、かつ、異なる情報を持つ 2 枚の画像を用いる。文献 8)9) では、フラッシュを焚いた場合と焚かない場合の 2 種類をペアとしている。フラッシュを焚いた場合、得られる画像のノイズを除去することができるが、蝋燭の明かりの下などでは照明の様子に変化してしまう。一方、フラッシュを焚かない場合は、照明の様子は変化しないが、ノイズが現れる。そこで、フラッシュを焚いた画像の情報を基にして、フラッシュを焚かない画像をフィルタリングする。その結果、照明の様子を維持したまま、ノイズの除去が可能であると述べている。ここで、この手法の重要点は、ノイズの多い画像の代わりに、ノイズの少ない画像をフィルタ

リングの計算に用いることで高い効果が得られる点である。

このクロスバイラテラルフィルタの効果は、奥行き画像の処理にも用いられている。文献 10)~12) では、クロスバイラテラルカーネルのコスト計算を基にした奥行き画像の超解像法を提案している。奥行き画像の解像度は低いという傾向があり、デプスセンサで取得した奥行き画像に効果を発揮する。これらは、入力にクロスバイラテラルカーネルのコスト計算に原画像を用いることで、原画像の輝度値を比較する。その結果、物体の輪郭を維持したまま奥行き画像の平滑化を行う。ここで問題となるのは、物体輪郭周辺の奥行き値が混合され、物体輪郭がぼけてしまうことである。これは、輝度値が大きく異なり、ピクセル間の距離が離れている場合でも、重みは少なからず与えられ、それにより奥行き値の混合が起こりやすいからである。

クロスバイラテラルフィルタの応用としてステレオマッチングの精度を向上させる手法 13) や、奥行き画像の改善を行う手法 14)~17) がある。文献 13) では、ステレオマッチングのコスト集約の段階においてクロスバイラテラルフィルタを適用させている。ここでは入力の原画像をカーネルのコスト計算に用い、ステレオマッチングの不安定な部分に対してコスト計算を行うことで、奥行き推定精度の向上を行っている。しかし、コスト集約の段階でクロスバイラテラルフィルタを用いるため、奥行き階調数に比例したフィルタリング回数が必要となり実行時間が長いという問題がある。また、クロスフィルタリングを用いた文献 14)~17) では、原画像をカーネルのコスト計算に用い、奥行き画像の改善を行っている。文献 14) では、クロスバイラテラルフィルタを複数回用い、その結果をピクセルごとに比較し、最も周囲と整合性のとれる結果を採択する。これにより物体輪郭を維持した奥行き画像の改善を可能としている。しかし、フィルタリングを複数回行うため、通常のフィルタリングより実行時間が長いという問題がある。

文献 15)~17) では、カーネル内の色の一致度、ピクセル間の距離、信頼度を用いてカーネルのコスト計算を行っている。信頼度は奥行き画像の精度がどの程度信頼できるかによって決定される。文献 15) では、ピクセル同士の奥行き値を比較し、その奥行き値が類似しているほど大きな重みが与えられる。その結果、クロスバイラテラルフィルタより大きな効果を得ることができる。しかし、物体輪郭を強調しようとすれば、クロスバイラテラルフィルタと同様に物体輪郭にぼけが発生し、奥行き値を強調しぼけの低減を試みると輪郭が補正されにくいといった問題点がある。これは、クロスバイラテラルフィルタは、原画像の輪郭を転写するようなカーネルを張ることで輪郭を補正することに起因する。輪郭を正しく補正するためには、ある程度の大きさのカーネルサイズが必要になる。しかし、クロスバイラテラルフィルタの混合によるぼけはカーネルサイズが大きいほど発生しやすい。そのため、混合を抑えるには、カーネルサイズを小さく抑えるしかない。つまり、ぼけ量と輪

郭補正の能力にはトレードオフの関係がある。文献16)17)では、左右視点の奥行き画像を用い、左右の奥行き画像を比較することで、各ピクセルの信頼度を決定している。左右で正しい奥行き画像を得られていると仮定すると、ある同一物体での奥行き値は等しくなる。そして、ある部分での奥行き値が左右の奥行き画像で大きく異なれば、その部分の奥行き値は信頼できないとして、信頼度を低く見積もる。また、繰り返し処理を行い、信頼度を更新していくことで精度の向上を図っている。これらの精度は非常に高いものになるが、これらは左右視点の奥行き画像を取得する必要があるため、ステレオマッチングに2倍の時間を要する。また、繰り返し処理を行うため、繰り返し回数に応じて実行時間も増大するといった問題点がある。

これらの従来手法では、精度が悪いが実行時間は短い簡単な改善手法か、精度はよいが実行時間が長い複雑な改善手法かのどちらかにほとんどが分類される。左右視点での奥行き画像を用い比較を行うか、繰り返し処理を行えば精度は向上させることができるが、実行時間が長くなり、このような手法が用いられる自由視点映像など、リアルタイム性が要求されるアプリケーションでは、用いることができない。また、デプスセンサを用いた場合では、左右視点の奥行き画像を取得すること自体が難しい。そして、一般にデプスセンサによって取得される奥行き画像の奥行き値の範囲は、ステレオマッチングによって取得されるものよりも大きいといわれている。例えば、Kinectで取得できる奥行き画像は11~16ビットである。このような理由から、奥行き値の階調数に依存して、実行時間が長くなる手法も好ましくない。したがって、本論文では一視点の奥行き画像と左右の原画像を用いて、繰り返し処理を行わずに改善を行うフィルタを提案する。従来手法に対する、提案手法である奥行き画像改善フィルタの主な優位点は次の点である。

- 左右の奥行き画像を用いず、また、繰り返し処理を行わないため高速であること
- 輪郭を維持しつつノイズ除去を行うクロスバイラテラルフィルタの特性を持ち、かつ、輪郭部分のぼけが抑制可能であること

以上から提案手法は、高速でありながら、比較的高精度な奥行き画像が取得可能である。

3. 提案手法

3.1 重み付きクロスバイラテラルフィルタ

提案手法では、ブロックマッチング (BM) で取得した奥行き画像をフィルタリングによって改善する。ブロックマッチングは簡単なステレオマッチングであり、実行時間は短くてすむが、平面部分の奥行き値は一定ではなく、また、輪郭部分に凹凸が多いなど精度が非常に悪い。提案するフィルタでは、物体の輪郭を維持した状態で、平面部分の平滑化を行う。これを重み付きクロスバイラテラルフィルタと呼ぶ。

重み付きクロスバイラテラルフィルタは、バイラテラルフィルタの拡張である。バイラテラルフィルタは次の式で表される。

$$O_p = \frac{\sum_{s \in N} w(X)c(Y)I_s}{\sum_{s \in N} w(X)c(Y)} \quad (1)$$

$$X = \|p - s\|_1 \quad Y = \|I_p - I_s\|_1$$

I : 入力画像, O : 出力画像, p : 注目ピクセルの座標, s : 比較ピクセルの座標, N : 比較ピクセル s の集合, $w(x)$: 位置重み, $c(x)$: 色重み, $\|\cdot\|_1$: L1 ノルム。

また、各重みはガウス分布を用いて以下のようになる。

$$w(x) = e^{-\frac{1}{2}(\frac{x}{\sigma_s})^2} \quad c(x) = e^{-\frac{1}{2}(\frac{x}{\sigma_c})^2} \quad (\sigma_s, \sigma_c: \text{定数}) \quad (2)$$

このフィルタでは、注目ピクセルと比較ピクセルの輝度値が近く、また、距離が近いとき、比較ピクセルに大きな重みが与えられる。輪郭をまたぐ部分は、輝度値が大きく異なるため、小さな重みが与えられる。このため、物体の輪郭を維持したまま、平面部分の平滑化が可能となる。

しかし、入力奥行き画像の多くが輪郭が正確でなく、そのような画像をバイラテラルフィルタでフィルタリングしても輪郭の補正はできない。そこで、入力画像に原画像を追加し、色重みの計算に原画像の輝度値を用いる。これにより、正しい物体輪郭の情報を用いて、輪郭を補正しつつ平滑化を行う。これをクロスバイラテラルフィルタといい、次の式で表される。

$$D_p = \frac{\sum_{s \in N} w(X)c(Y)D_s}{\sum_{s \in N} w(X)c(Y)} \quad (3)$$

$$Y = \|I_p - I_s\|_1$$

ここで、 D_p , D_s は注目ピクセルと比較ピクセルの奥行き値である。また、色重みの計算には原画像を入力とし、重みの計算は式 (2) を用いる。なお、ここにおける X および Y は式 (1) と同様である。これにより、物体の輪郭を維持したまま平滑化が可能となる。しかし、これにも一つ問題がある。それは、輝度値が大きく異なり、ピクセル間の距離が離れている部分に対しても、少なからず重みが与えられてしまうことである。そのため、輪郭線は原画像に近くなるが輪郭周りにぼけが発生する。

理想的には、輪郭をまたいだ先のピクセルには一切の重みを与えないという処理を行うことが望ましい。そこで、輪郭をまたいだ先か、そうでないかを判断する分類を設定し、輪郭をまたいでいない部分、つまり同一物体である平面部分にのみ大きな重みが与えられるよう処理を行う。その重みは事前に決定し、各ピクセルごとに重みが与えられる。また、重みは重み画像 R として取得される。そして、提案手法である重み付きクロスバイラテラルフィルタは次の式で表される。

$$D_p = \frac{\sum_{s \in N} w(X)c(Y)R_s D_s}{\sum_{s \in N} w(X)c(Y)R_s} \quad (4)$$

ここで、 R_s は平面部分を重視した重みである。また、ここにおける X および Y は式 (1) と同様である。この平面部分を重視した重みは、まず、ブロックマッチングで取得した奥行き画像と左右の原画像から、奥行き画像の信頼できる部分とそうでない部分を分類する。そして、分類によって信頼できるとされた部分に対してのみ重みを与える。その分類は次のように決定する。あるカーネルにおいて、その中心ピクセルを注目ピクセル p' 、その周りのピクセルを比較ピクセル s' とする。注目ピクセルと比較ピクセルの間で輪郭をまたがない場合、それぞれのピクセルは同一物体を指していると考えられることができるため、奥行き画像の奥行き値とその視点での原画像の輝度値は類似する。また、あるピクセルの奥行き値から、もう一方の視点での対応点を決定できるため、そのピクセルにおける奥行き値が正しければ、そのピクセルの原画像の輝度値と対応するもう一方の視点での輝度値も類似するはずである。これにより、分類は次の条件で行われ、重み画像が取得される。

- (1) まず、左視点を基準とした場合、左視点の奥行き画像において、注目ピクセル $D_{p'}^l$ と比較ピクセル $D_{s'}^l$ の奥行きを比較し、その差が閾値 α 以下であること。

$$\|D_{p'}^l - D_{s'}^l\|_1 \leq \alpha \quad (5)$$

- (2) 次に、左視点の原画像において、注目ピクセル $I_{p'}^l$ と比較ピクセル $I_{s'}^l$ の輝度値を比較し、その差が閾値 β 以下であること。

$$\|I_{p'}^l - I_{s'}^l\|_1 \leq \beta \quad (6)$$

- (3) そして、左右視点間の距離と、左視点の比較ピクセルでの奥行きから、右視点における比較ピクセルの対応ピクセルを一意に定められるため、右視点での原画像の対応ピクセル $I_{s'+D_{s'}^l}^r$ と、左視点での原画像の比較ピクセル $I_{s'}^l$ の輝度値を比較し、その差が閾値 γ 以下であること。

$$\|I_{s'}^l - I_{s'+D_{s'}^l}^r\|_1 \leq \gamma \quad (7)$$

- (4) 以上の条件を全て満たす部分は、取得した奥行き画像が信頼できるとする。

l : 左視点, r : 右視点, α, β, γ : 各任意の閾値.
 そして、信頼できるとしたピクセルに対してのみ、次で定める重みを与え、信頼できない部分は重みを 0 とする。

$$R_s = \sum_{s' \in N'} b(p', q') w(X) c(Y) e^{-\frac{1}{2} \left(\frac{\|D_{p'}^l - D_{s'}^l\|_1}{\sigma_r} \right)^2} \quad (8)$$

(σ_r : 定数)

$$b(p', q') = \begin{cases} 1 & (\text{satisfy conditions}) \\ 0 & (\text{otherwise}) \end{cases} \quad (9)$$

N' : 比較ピクセル s' の集合, σ_r : ガウス変数.

ここで、 $b(p', q')$ は、式 (5)~(7) の条件を満たした画素の重みだけを加えるための二値を取る関数である。なお、重み画像に与えられる $w(X)$ および $c(Y)$ は式 (1) と同様に与えられる。また、このときのステレオの入力視点は、重み画像の視点と同様であるため、この場合は左視点の原画像を入力とする。さらに、奥行き画像の輪郭上は不正確な奥行き値が割り当てられることが多いため、平面部分に多くの重みが割り当てられることが望ましい。もし、近傍画素の色が近く、そして、その奥行き値も近い値を取っていれば、平面である可能性が高い。そこで、式 (1) の $w(X)$ と $c(Y)$ を再利用し、また、新たに奥行き値の近さを指標に加えることで重みを設定する。これにより、ブロックマッチングのような輪郭が曖昧な奥行き画像では、輪郭部分の重みは小さくなり、輪郭をまたいだ部分には重みが与えられにくくなる。そして、輪郭をまたがない平面部分には大きな重みが与えられるため、同一物体となる部分にのみ大きな重みが与えられ、輪郭を維持したまま、平滑化が可能になる。

3.2 奥行き値の混合抑制フィルタ

重み付きクロスバイラテラルフィルタにより、輪郭部分の精度を改善し平滑化を行う手法を提案した。しかし、前景と背景の物体色が似ている場合など、奥行きが異なるにも関わらず、信頼できるとしてしまう部分が存在する。また、平面領域が小さく大きな重みが与えられないため、周囲の奥行き値の影響を受ける部分も存在する。さらに、物体が複雑で細かく奥行き値が変化する場合や、物体がフィルタのカーネルに対して小さい場合、周囲の奥行き値の影響を受ける可能性がある。そういった部分がある場合、その周辺の輪郭は複数の奥行き値を混合することでぼけが発生する。そこで、わずかに残る輪郭部分のぼけを解消するための処理を施す。

ぼけは重みによって、複数の奥行き値を混合することで出現する。本来、境界の奥行き値は前面と背面という二値的な奥行き値で存在するため、混合された中間値はふさわしくない。一方で、フィルタ前の奥行き値は、輪郭の位置は正しくなく、ノイズも含まれるが、混合された中間値は存在せず、正解となる前面、背面の奥行き値を多く含む。そこで、輪郭部分が二値的に表現されるフィルタリング前の奥行き画像を用いて、重み付きクロスバイラテラルフィルタで混合された中間値を前景と背景の奥行き値に振り分ける。ここでは、フィルタリング前の奥行き画像とはブロックマッチングで取得した奥行き画像である。その結果、混合部分は存在しなくなることで輪郭部分のぼけが除去され、正しい奥行き画像により近くなる。これを混合抑制フィルタと呼び、以下の式で表される。

$$D_{Proposed}(p) = \arg \min_{d \in D(W)} |D_{WCBF}(p) - D_X(s)| \quad (10)$$

D_X :手法 X で取得した奥行き画像 (例えば BM), $WCBF$:奥行き画像取得手法重み付きクロスバイラテラルフィルタ, BM :奥行き画像取得手法ブロックマッチング, W :比較ピクセル s の集合, $D(W)$:集合の要素 W が指す奥行き, d :取り得る奥行き値, D'_{WCBF} :混合抑制前の重み付きクロスバイラテラルフィルタで取得した Depth Map.

このフィルタでは, D_X のカーネル内にある奥行き値の候補に最も近い値を振り分ける. つまり, ぼけているが正しい値に近づいた奥行き画像のぼけ部分を, ぼけの生じる前の最も近い値に振り分けることでぼけを抑制する. 以上から, 奥行き画像を物体の輪郭を維持しつつ平滑化を行い, 輪郭がぼけた部分に対しては二値的な奥行きに変換する処理を行うことでぼけを低減する. ただし, この混合抑制フィルタにより一部劣化する可能性を持つ場合もある. まず一つ目は, 重み付きクロスバイラテラルフィルタにより生じたぼけと, ノイズの乗った奥行き値の値が偶然にも近かった場合である. 二つ目は, 隣り合う画素の奥行きが全て変化するような, 非常に急勾配で変化する領域であり, その場合は中間値のほうが望ましい. しかしながら, このような領域は, 奥行き推定を行うことも難しいため, 正解の奥行き値がほとんど求まらず, そもそも重み付きクロスバイラテラルフィルタによって奥行き画像が改善する可能性は低い. そのため, この混合抑制フィルタにより劣化する領域は非常に少ない.

4. 実験方法

実験では, 奥行き推定精度の比較を行う. 精度の比較は, The Middlebury Computer Vision Pages2) で提供されているステレオ画像のうち, tsukuba(図 1(a)), venus(図 1(b)), teddy(1(c)), cones(1(d)) を用いて奥行き推定精度を検証する.

提案手法の奥行き画像改善フィルタとしての効果を, 既存の奥行き画像改善フィルタと比較するために, メディアンフィルタ (MF), バイラテラルフィルタ (BF), クロスバイラテラルフィルタ (CBF) (8)~(10), 信頼度に奥行きの類似度を用いたクロスバイラテラルフィルタ (CTF) (15), Constant-Space Belief Propagation (CSBP) で用いられる奥行き画像改善フィルタ (14) を比較対象とする. これらフィルタへの入力奥行き画像にはまず, BM によって取得したものをを用いる. また, 提案手法の適用範囲を検証する実験として, BM により推定された Depth Map の入力だけでなく, 三つの奥行き推定手法で得られた Depth Map も入力する.

まず一つ目は, 高速な MRF 最適化を行うことができる (SGBM3) (4) である. 文献 (4) の手法のほうが文献 (3) の手法よりも高精度だが, 計算時間が非常に大きくかかるため, 後者の手法を基本に高速化を行った. 文献 (3) からの大きな変更は 3 点であり, コスト計算に Mutual Information (MI) を用いていた点を Sum of Absolute Difference (SAD) へ

変更し, LR チェック (2) を除去することで Depth Map 推定を 1 枚のみにして, 16 方向のパスからの MRF 最適化を行っていた点を 8 方向へと削減した.

二つ目は, 特徴点ベースの手法である Efficient Large-Scale (ELAS) (18) である. ELAS は, まず始めに, 疎に特徴点のみをマッチングさせ, その部分の奥行きを推定する. 残りは, その対応を基に補間する手法となっている. そのため, 奥行きの階調数に依存しづらい手法となっている. この手法では, 公開されているコード* を用い, 各パラメータのデフォルト値を用いた.

最後の比較手法は, セグメンテーションと Belief Propagation (BP) による最適化を行う Double Belief Propagation (DBP) である. この手法は, セグメンテーションと BP の各処理を計算するのに多くの時間を要し, 実行時間は分単位となる. しかしながら, 輪郭の精度も, 平面部分の精度も非常に高い手法となる. この手法では, 実装は行わず, ステレオマッチングの評価サイトである Middlebury Stereo Vision Page** から実行結果をダウンロードした.

以上の手法により取得した奥行き画像を入力として提案手法を適用し, その効果を検証する. なお, 提案手法である重み付きクロスバイラテラルフィルタと混合抑制フィルタ, それぞれの効果を検証するため, 重み付きクロスバイラテラルフィルタ (WCBF) 単体と, WCBF で取得した奥行き画像に混合抑制フィルタを適用した (Proposed) 場合の評価を行う.

そして, これらの手法から得られる奥行き画像の誤り率を比較し性能評価を行う. この誤り率は, 文献 (2) で提供されている正解奥行き画像と取得した奥行き画像の奥行き値を比較し, 誤っているピクセル数を画像の全ピクセル数で割った割合である. ここでは, 奥行き値の差が 1 以内の場合は正解と見なした. ただし, フィルタではカーネル内に正しい奥行き値が存在しないとき, 正しい奥行き値を求めることができない. 例えば, メディアンフィルタでは, カーネル内の中央値を選択するため, カーネル内に正しい奥行き値が存在しなければ, 決して正しい値にはならない. 元々, 正しい奥行き値が存在しない部分では, 基本的にフィルタで改善の余地がなく, どれだけ改善の余地があり, どれだけ改善されたかは, 誤り率を比較するだけではわかりにくい. よって, 奥行き画像改善の比較では, 改善の余地がある部分だけを比較するために, マスクを用いて改善の余地がある部分だけに絞った比較も行う. マスクは, あるピクセルの近傍 (13×13) の中の奥行き値を Ground Truth と比較し, 正しい値が一つも存在しない場合は評価から除外する.

5. 実験結果

前述の方法で取得した奥行き画像の各誤り率は表 1 のよ

* <http://www.rainsoft.de/software/libelas.html>.

** <http://vision.middlebury.edu/stereo/>.

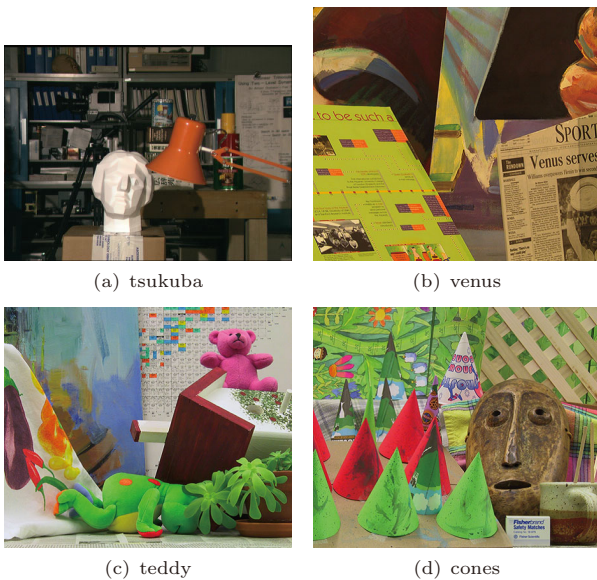


図 1 ステレオデータセット
Stereo data sets.

表 1 誤り率
Error rate.

Mask	Method	tsukuba	venus	teddy	cones
NoMask	BM	4.35	2.42	13.24	7.08
	WCBF	4.08	1.59	12.29	6.74
	Proposed	3.22	1.23	11.89	5.91
	MF	3.99	2.16	13.12	6.89
	BF	3.91	2.59	12.93	6.69
	CBF	4.36	2.77	14.04	8.49
	CTF	3.87	1.82	12.93	6.70
	CSBP	3.55	2.78	12.48	6.24
	SGBM	3.26	1.33	7.60	4.55
	SGBM with WCBF	2.97	1.11	6.98	3.78
	SGBM with Proposed	2.73	0.68	6.18	2.62
	ELAS	3.99	1.84	7.99	6.85
	ELAS with WCBF	3.43	1.42	7.00	5.88
	ELAS with Proposed	3.25	1.22	6.63	4.88
	DBP	0.88	0.13	3.53	2.90
	DBP with WCBF	0.86	0.15	3.60	2.96
	DBP with Proposed	0.85	0.12	3.72	2.89
Mask	BM	3.86	2.29	4.03	3.76
	WCBF	3.59	1.47	2.92	3.40
	Proposed	2.71	1.11	2.47	2.54
	MF	3.49	2.02	3.89	3.57
	BF	3.41	2.45	3.70	3.42
	CBF	3.85	2.64	4.92	5.24
	CTF	3.37	1.68	3.70	3.43
	CSBP	3.05	2.63	3.14	2.90
	SGBM	3.78	1.14	4.70	3.85
	SGBM with WCBF	2.74	1.02	3.86	3.11
	SGBM with Proposed	2.53	0.60	3.26	1.96
	ELAS	3.13	1.53	5.19	4.46
	ELAS with WCBF	2.50	1.98	4.16	3.36
	ELAS with Proposed	2.32	1.21	3.59	2.63
	DBP	0.90	0.13	1.81	1.95
	DBP with WCBF	0.88	0.14	1.86	2.02
	DBP with Proposed	0.87	0.11	1.86	1.94

うになる。表 1 では、上段が通常の誤り率評価、下段が実験方法で示した改善の余地がある部分だけに焦点を当てたマスクありの評価である。この際、用いたマスクは各データセットごとに図 2, 5- 7 の (b) となっている。マスクでは、黒で表示される部分は改善の余地あり、白で表示さ

表 2 実行時間
Running time.

Data Set	BM	Proposed	BM+Proposed	SGBM	ELAS
tsukuba	5.7	4.1	9.8	28.8	61.1
venus	8.4	6.7	15.1	45.9	110.0
teddy and cones	10.5	6.8	17.3	71.4	168.2

(ms)

れる部分は改善の余地なしとなる。ここでは、画像の端といった BM など簡単なステレオマッチングでは求められない部分が改善の余地なしとなっている。そして、各手法によって取得した奥行き画像は図 2(c)-(p) である。ここでは、データセット teddy を挙げる。ここで、WCBF の各データセットにおけるパラメータは、teddy では、重み画像作成時は $(\sigma_s, \sigma_c, \sigma_r, \alpha, \beta, \gamma) = (5.8, 3.0, 9.8, 1, 10, 205)$ であり、フィルタリング時は $(\sigma_s, \sigma_c) = (13.5, 16.4)$, tsukuba では、重み画像作成時は $(\sigma_s, \sigma_c, \sigma_r, \alpha, \beta, \gamma) = (6.4, 6.2, 13.2, 1, 10, 54)$ であり、フィルタリング時は $(\sigma_s, \sigma_c) = (10.9, 10.4)$, venus では、重み画像作成時は $(\sigma_s, \sigma_c, \sigma_r, \alpha, \beta, \gamma) = (8.5, 2.6, 10.3, 1, 5, 56)$ であり、フィルタリング時は $(\sigma_s, \sigma_c) = (14.4, 6.6)$, cones では、重み画像作成時は $(\sigma_s, \sigma_c, \sigma_r, \alpha, \beta, \gamma) = (6.1, 3.9, 6.2, 1, 10, 113)$ であり、フィルタリング時は $(\sigma_s, \sigma_c) = (10.9, 8.4)$ となっている。これらの値は経験的に決定し、その判断基準は原画像の輪廓を維持した状態で誤り率が低くなるように設定した。また、他データセットの結果は、簡単な改善フィルタである MF と提案手法と同程度の速度で動作する CTF を抜粋し図 5 - 7 となっている。

まず、表 1 より、BM を入力とした提案手法 (WCBF) の奥行き改善がすべてのデータセットにおいて効果があることがわかる。それぞれのデータセットで、BM で取得した結果を奥行き画像改善手法である提案手法によって改善を行い、改善による誤り率の減少値は (マスクなし, マスクあり) で表記すると、tsukuba では (1.13, 1.15), venus では (1.19, 1.18), teddy では (1.35, 1.56), cones では (1.17, 1.22) となった。提案手法と同じく、奥行き画像改善を行う各フィルタと比べてもマスクあり, マスクなしに関わらず提案手法が最も高い効果を得た。これは、奥行き画像改善フィルタにおいては改善余地のある部分はすべて同じためである。また、提案手法がもっとも高い効果を示した理由としては、MF や BF ではノイズは除去できるが輪廓の改善効果はなく、CBF や CTF ではノイズの除去や輪廓の改善は可能だが奥行き値の混合によりぼけが発生する。対して、提案手法ではノイズ除去や輪廓の改善、混合によるぼけの抑制を行うため、他の手法よりも高い効果を示した。

次に、提案手法を構成する WCBF と混合抑制フィルタの各手法ごとの効果を検証する。BM で取得した奥行き画像に WCBF を適用した段階での誤り率の減少値と、そこから混合抑制フィルタを用いた誤り率の減少値は (WCBF での減

少値, 混合抑制フィルタでの減少値)と記述すると, マスクなしの tsukuba では (0.27, 0.86), venus では (0.83, 0.36), teddy では (0.95, 0.40), cones では (0.34, 0.83) となった. マスクありの tsukuba では (0.27, 0.88), venus では (0.82, 0.36), teddy では (1.11, 0.45), cones では (0.36, 0.86) となった. この結果から, マスクの有無にかかわらず venus, teddy では WCBF の効果が高く, tsukuba, cones では混合抑制フィルタの効果が高いことがわかる. この値から, venus, teddy では平面部分のノイズ除去の効果が誤り率の減少に大きく寄与し, tsukuba, cones では, 輪郭部分のぼけの抑制が誤り率の減少に大きく寄与していることがわかる. これは, 輪郭部分の領域が venus, teddy に比べて, tsukuba, cones の方が多いため, 輪郭部分のぼけを抑制する混合抑制フィルタの効果が大きく発揮されたためだと考えられる. よって, 混合抑制フィルタの効果は, データセットの輪郭部分の占める割合に影響し, 輪郭部分が比較的多いデータセットに対しては, 重み付きクロスバイラテラルフィルタよりも効果を発揮するものと考えられる. しかし, 前提条件として, 混合抑制フィルタは輪郭線が正しく出ているがぼけてしまう領域に対して効果を発揮するため, 重み付きクロスバイラテラルフィルタをまず適用することが必要になる.

実際にどのような部分が改善したのかを図 2, 5 - 7 を観察し具体的に述べる. まず, 平面部分にある細かなノイズやムラは平滑化の効果で滑らかになっていることがわかる. そして, 提案手法の主な主張である輪郭部分の正確さは BM に比べて大幅に改善していることがわかる. BM で取得した奥行き画像では, 前景が背景に膨張している部分や, 前景が背景に浸食されている部分が多く, 輪郭部分に凹凸が多く存在している. しかし, 提案手法の WCBF で奥行き画像の改善を行うと, その凹凸が大きく改善されていることがわかる. これは図 3 に示す重みによって平面部分に大きな重みが与えられたことによる効果である. 図 3 の重みは, 色が白に近づくほど大きな重みが与えられる. 重みは平面部分など正しい奥行き値が容易に決定できる部分に大きな重みが与えられている. そして, 輪郭部分など奥行き値が不確定となりやすい部分では重みが小さくなる. その結果, 輪郭をまたぐ部分の平滑化において, 輪郭をまたいだ先の奥行き値には小さな重みが与えられ, 同一平面の奥行き値のみに大きな重みが与えられる. そして, 仮に輪郭をまたいだ先の重みが大きい場合でも, 色重み, 位置重みによって重みは小さく与えられる. その結果, 各図に見られるように, 物体の輪郭を維持したまま平滑化が可能となる. また, 平面部分に大きな重みを与えることで, 一度のフィルタリングで大きな効果が得られるため, 繰り返し処理の必要もない.

次に, BM 以外の手法で Depth Map を取得した場合について述べる. SGBM, ELAS, DBP で取得した奥行き画像を入力とした, 提案手法の適用範囲に関する検証では,

SGBM, ELAS に対しては提案手法を用いることによりすべてのデータセットで誤り率の低下が確認された. しかし, DBP に対しては誤り率の低下するものと増加するものがあり, また, 誤り率の低下するものでも, その差は微差となった. これは DBP ではセグメンテーションと高度な最適化を行うため, 輪郭精度が高くまた平面部分の精度も高い. そのため提案手法の輪郭補正やノイズ除去の効果が表れなかったことが原因の一つと考えられる. よって, 輪郭精度が高いセグメンテーションを用いた手法や, 平面の誤りが少ない高度な最適化を行う手法などには, 提案手法の適用は適切でないといえる. しかし, SGBM や ELAS のように効果が表れている手法もあるため, その最適化手法の特徴によって適用範囲が限られると考えられる. 提案手法の効果から, 平面部分にノイズの多い奥行き画像や輪郭精度が低い奥行き画像に対しては効果があると考えられる. そして, BM のようにノイズ量が多く, 輪郭精度も不十分な奥行き画像に対しては大きな効果が期待できる.

最後に, 奥行き画像取得までの実行時間を表 2, 図 4 に示す. ここで, 実験環境は Intel Core i7-920 の 2.93GHz である. 表 2 は, データセットから各手法により Depth Map を取得するまでに要した計算時間を示しており, BM の結果に奥行き画像の改善を行うフィルタを用いる方法が最も高速である. 最適化の中でも高速に動作する SGBM や奥行き階調数が多い場合にも高速に動作するとされている ELAS よりも高速である.

これは, 計算時間のオーダはフィルタリングである提案手法では $O(H \cdot W)$ であるのに対して, BM, SGBM は $O(H \cdot W \cdot D)$ となっている. ここで, H は画像の高さ, W は画像の幅, D は奥行き階調数である. BM と SGBM の視差に応じた計算量の増加を比べると, SGBM の増加量は大きい. 加えて, この奥行き階調数は画像サイズに応じて増加していくため, 画像サイズが増加するほど, 提案手法と最適化手法の実行時間の差は顕著になっていく. また, ELAS は D の影響は比較的小さいが, フィルタリングに比べると重たい処理となっている. 図 4 は, 解像度 128×128 , 奥行き階調数 16 を基本として解像度, 奥行き階調数を $256 \times 256, 384 \times 384, \dots, 32, 48, \dots$ と n 倍ずつしていったときに, 各手法の計算時間をシミュレートした結果を示している. 図中左は, 全ての手法の相対関係を示すために, 縦軸のスケールを上限 6000ms で表示している. また, 各手法の詳細を示すために, 右は上限 400ms で示している. この図から, 提案手法のフィルタと他の手法と比べると, 解像度や奥行き階調数が大きくなるほどその差が非常に大きくなることがわかる. また, 図 4 右から, 解像度や奥行き階調数が少ない場合には ELAS よりも, SGBM のほうが高速であることがわかった.

6. む す び

本論文では, 重み付きクロスバイラテラルフィルタと混

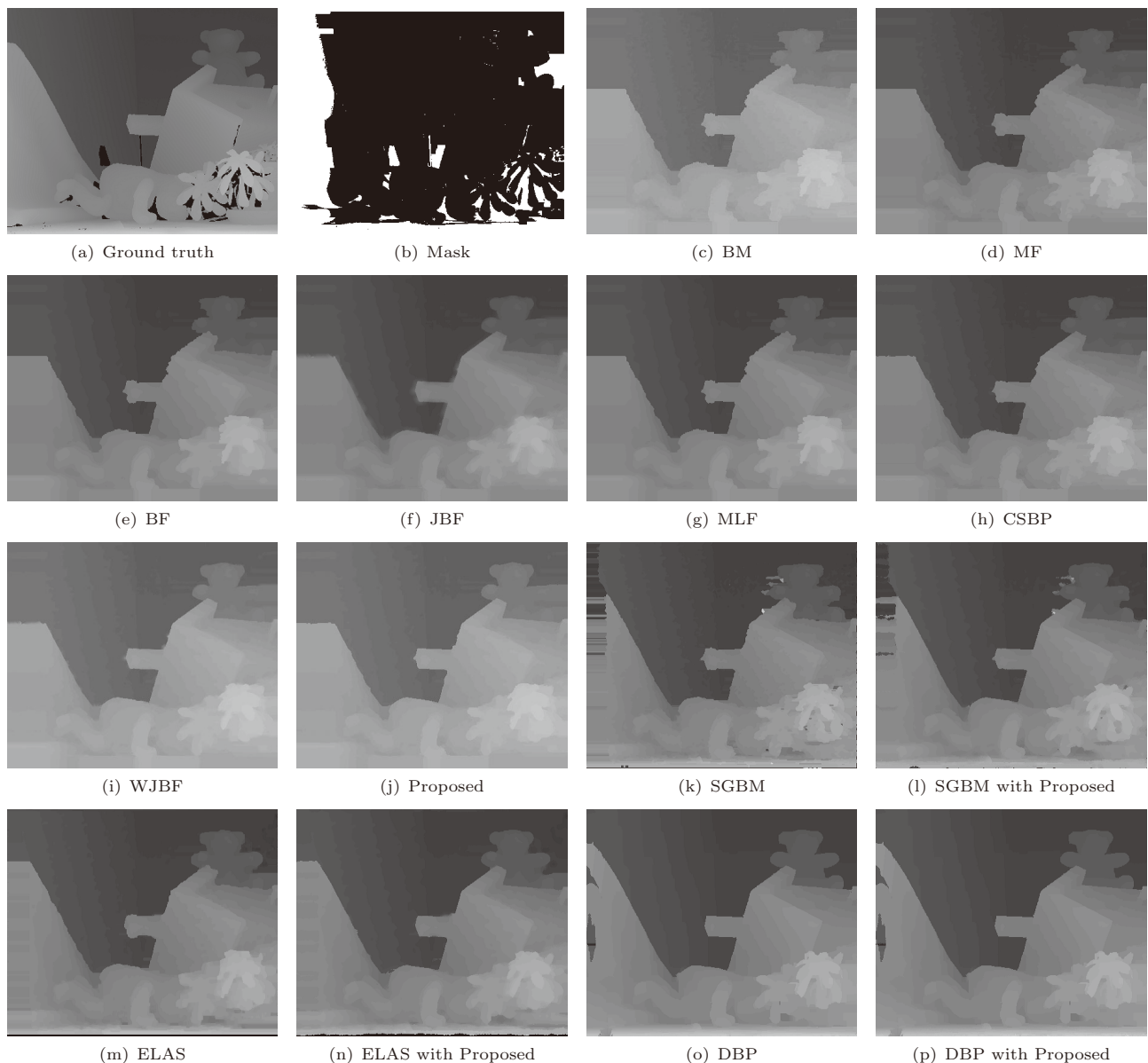


図 2 各手法の結果 (teddy)
Results of teddy's depth map.

合抑制フィルタを用いた奥行き画像の改善を提案した。本論文での主な主張は、繰り返し処理を行わず、輪郭部分のぼけを除去し、その部分を補間するということであった。提案手法により、それぞれのデータセットで、簡単であるが精度の悪い BM で取得した奥行き画像からの誤り率の減少値は、tsukuba で 1.13, venus で 1.19, teddy で 1.35, cones で 1.17 となり、いずれにおいても改善が確認された。同様に、フィルタによって改善可能な部分だけを取り上げた際の誤り率の減少値は、tsukuba で 1.15, venus で 1.18, teddy で 1.56, cones で 1.22 となり、こちらもすべてのデータセットにおいて改善が確認された。他の奥行き画像改善を行うフィルタと比較した結果は、提案手法が最も高い効果を得ることができた。また、誤り率の改善だけでなく、物体輪郭の改善が行われていることを画像から確認した。さらに、高速だとされているその他の奥行き推定手法

に対して、提案手法を適用した場合は、SGBM と ELAS とともに、改善が確認された。計算速度については、提案手法では繰り返し処理の必要がなく、高速に動作する SGBM や ELAS より高速であることを示した。

今後の課題は、物体輪郭が複雑な部分での輪郭精度の向上と、入力奥行き画像に対する依存性の検証、また、改善に及ぼすパラメータの依存性の調査が考えられる。

最後に、日ごろ、御討論頂く本学菅原真司准教授、布目敏郎准教授に感謝する。本研究の一部は、科学研究費補助金若手研究 (B) (22700174)、JST A-STEP、および、総務省戦略的情報通信研究開発推進制度 (SCOPE) によって行われた。

【文 献】

- 1) Y. Mori, N. Fukushima, T. Yendo, T. Fujii and M. Tanimoto, "View Generation 3D Warping Using Depth Information for FTV", *Signal Processing: Image Communication*, **24**, 1-2, pp. 65-

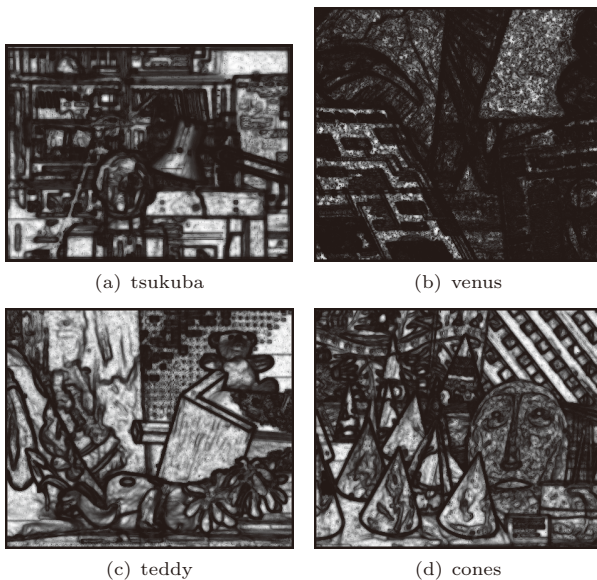


図 3 重み画像
Weight Maps.

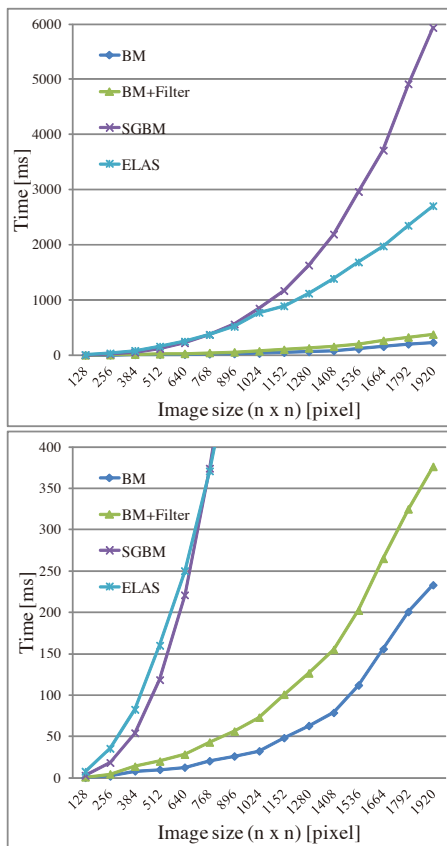


図 4 各手法の実行時間: 左) 縦軸 6000ms まで, 右) 縦軸 400ms まで.
Computational time of each method: left) Y axis is until 6000 ms, right) Y axis is until 400ms.

72 (Jan. 2009)

2) D. Scharstein and R. Szeliski, "A Taxonomy and Evaluation of Depth Two-Frame Stereo Correspondence Algorithms", International Journal of Computer Vision, **47**, 1, pp. 7-42 (May 2002)

3) H. Hirschmüller, "Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information", Proceedings of IEEE Computer Vision and Pattern Recognition (CVPR'05), pp. 807-814 (June 2005)

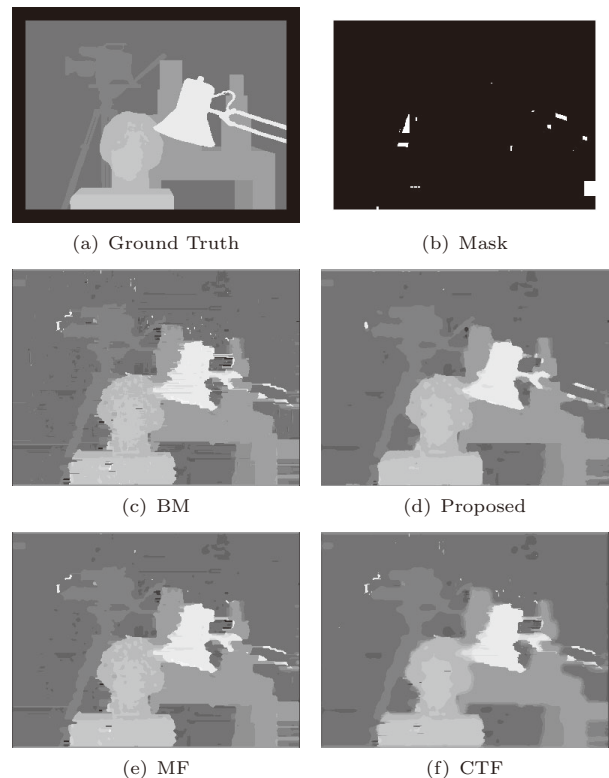


図 5 各手法の結果 (tsukuba)
Results of tsukuba's depth map.

4) H. Hirschmüller, "Stereo Processing by Semi-Global Matching and Mutual Information", IEEE Transaction on Pattern Analysis and Machine Intelligence, **30**, 1, pp. 328-341 (Feb. 2008)

5) J. Sun, N. N. Zheng and H. Y. Shum, "Stereo Matching Using Belief Propagation", IEEE Transaction on Pattern Analysis and Machine Intelligence, **25**, 7, pp. 787-800 (July 2003)

6) Y. Boykov, O. Veksler and R. Zabih, "Fast Approximate Energy Minimization via Graph Cuts", IEEE Transaction on Pattern Analysis and Machine Intelligence, **23**, 11, pp. 1222-1239 (Nov. 2001)

7) C. Tomasi and R. Manduchi, "Bilateral Filtering for Gray and Color Image", Proceedings of IEEE International Conference on Computer Vision (ICCV'98), pp. 839-846 (Jan. 1998)

8) G. Pestschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe and K. Toyama, "Digital Photography with Flash and No-Flash Image Pairs", ACM Transaction on Graphics, **23**, 3, pp. 664-672 (Aug. 2004)

9) E. Eisemann and F. Durand, "Flash Photography Enhancement via Intrinsic Relighting", ACM Transaction on Graphics, **23**, 3, pp. 673-678 (Aug. 2004)

10) J. Kopf, M. F. Cohen, D. Lischinski and M. Uyttendaele, "Joint Bilateral Upsampling", ACM Transaction on Graphics, **26**, 3, pp. 96 (July 2007)

11) Q. Yang, R. Yang, J. Davis and D. Nistér, "Spatial-Depth Super Resolution for Range Images", Proceedings of IEEE Computer Vision and Pattern Recognition (CVPR'07) (June 2007)

12) D. Chan, H. Buisman, C. Theobalt and S. Thrun, "A Noise-Aware Filter for Real-Time Depth Upsampling", Proceedings of European Conference on Computer Vision (ECCV'08) Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications (Oct. 2008)

13) K. J. Yoon and I. S. Kweon, "Adaptive Support-Weight Approach for Correspondence Search", IEEE Transaction on Pattern Analysis and Machine Intelligence, **28**, 4, pp. 650-656 (Apr. 2006)

14) Q. Yang, L. Wang and N. Ahuja, "A Constant-Space Belief Propagation Algorithm for Stereo Matching", Proceedings of IEEE Computer Vision and Pattern Recognition (CVPR'10), pp. 1458-1465 (June 2010)

15) P. L. Lai, D. Tian and P. Lopez, "Depth Map Processing with Iterative Joint Multilateral Filtering", Proceedings of Picture Coding Symposium (PCS'10), pp. 9-12 (Dec. 2010)

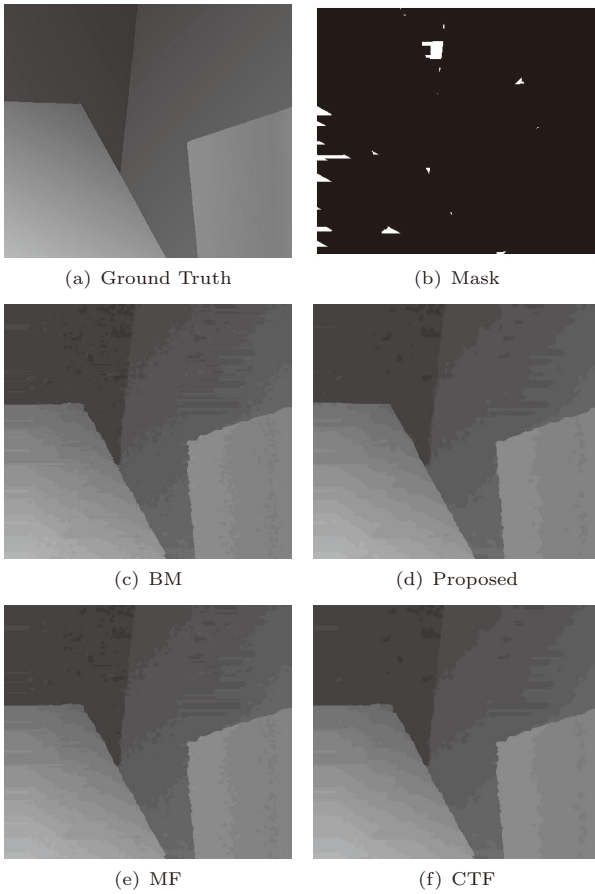


図 6 各手法の結果 (venus)
Results of venus's depth map.

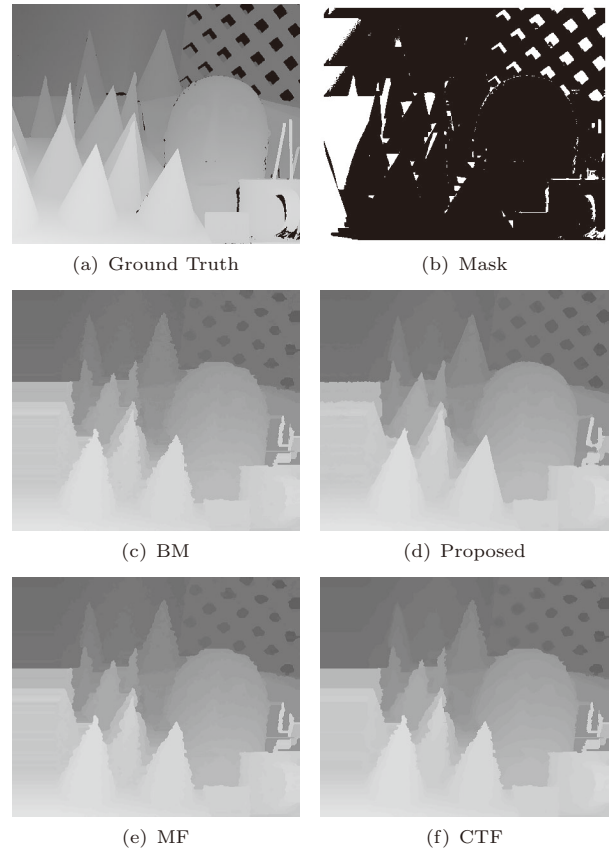


図 7 各手法の結果 (cones)
Results of cones' depth map.

- 16) M. MueTller, F. ZillTy and P. KauffT, "Adaptive Cross Trilateral Depth Map Filtering", Proceedings of 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON'10) (June 2010)
- 17) J. Jachalsky, M. Schlosser and D. Gandolph, "Confidence Evaluation for Robust, Fast-Converging Disparity Map Refinement", Proceedings of IEEE International Conference on Multimedia and Expo (ICME'10), pp. 1399-1404 (July 2010)
- 18) A. Geiger, M. Roser and R. Urtasun, "Efficient Large-Scale Stereo Matching", Proceedings of Asian Conference on Computer Vision (ACCV'10), **6492**, pp. 25-38 (Sep. 2010)
- 19) Q. Yang, L. Wang, R. Yang, H. Stewénus and D. Nistér. "Stereo Matching with Color-weighted Correlation, Hierarchical Belief Propagation and Occlusion Handling", IEEE Transaction on Pattern Analysis and Machine Intelligence, **31**, 3, pp. 492-504 (Mar. 2009)



まつお たくや
松尾 琢也 2011 年, 名古屋工業大学工学部情報工学科卒業。現在, 同大学院工学研究科創成シミュレーション工学専攻博士前期課程に在学中。多視点映像からの奥行き推定手法の研究に従事。



ふくしま のりしげ
福嶋 慶繁 2004 年, 名古屋大学工学部電気電子情報工学科卒業。2009 年, 同大学大学院工学研究科電子情報システム専攻博士課程後期課程修了。同年, 名古屋工業大学大学院助教。現在, 多視点映像処理の研究に従事。博士 (工学)。正会員



いしはし ゆたか
石橋 豊 1981 年, 名古屋工業大学工学部情報工学科卒業。1983 年, 同大学大学院修士課程修了。同年, 日本電信電話公社入社。NTT ヒューマンインターフェース研究所主任研究員を経て, 1993 年より名古屋工業大学工学部電気電子情報工学科助教授。現在, 同大学大学院教授。分散マルチメディアの研究に従事。工学博士。正会員。