

超解像処理によるライトフィールドカメラ画像の高画素化 Super Resolution of Light Field Camera Images

蚊野浩† 中島弘喜†

Hiroshi Kano†, Hiroki Nakajima†

†京都産業大学大学院先端情報学研究科

†Kyoto Sangyo University, Division of Frontier Informatics

E-mail: kano@cse.kyoto-su.ac.jp

Abstract

ライトフィールドカメラの入力画像群を超解像処理によって高画素化する手法を提案する。入力画像群は多数のステレオ画像を構成するため、マルチベースラインステレオ法を用いて画像間の位置ずれ量を視差として検出する。その視差に基づいて入力画像を位置合わせし、MAP型超解像処理によって高画素化する。その結果、良好に視差が計算される領域では超解像度処理による高画素化が有効であった。一方、高周波成分の折り返しが強く現れる領域は視差を計算することが難しく、超解像度処理による高画素化は困難である。

1 はじめに

デジタルカメラは撮像素子表面に鮮明な光像を形成し、その光像に忠実なデジタル画像を生成する装置である。ここでは、画像再現のための写真画像処理技術や顔検出・動被写体の追跡などのコンピュータビジョン技術を中心に、高度なデジタル画像処理技術が利用されている。

現在のデジタルカメラと異なった原理に基づくカメラとしてライトフィールドカメラ(以下LFCと略記)が知られている。これはライトフィールド(光線空間)として多数の画像を取得し、この画像群に画像処理を施すことで最終画像を得るカメラである。代表的な機能は、撮影後の後処理によってカメラ焦点距離を変更した画像を生成できるリフォーカスである。初期のLFCとして、多数のカメラを縦横に並べたカメラアレイによる実装例がある[1]。

カメラアレイによるLFCは機能検証に有効であるが、実用的ではない。Ngらは撮像素子の前面にマイクロレンズを配置し、コンパクトなカメラ筐体で画像群を取得可能なLFCを発明した(以下マイクロレンズLFCと略記)[2]。これは優れたアイデアであるが、最終画像が非常に小さくなるという課題がある。例えば、 $3,840 \times 3,840$

(1,474万)画素の撮像素子で、縦横 $15 \times 15 = 225$ 個の画像群を取得する構成では、個々の小画像は 256×256 画素になり、最終画像もこの大きさになる。

本論文は、最終画像の画素数が小さいことが課題となるマイクロレンズLFCへの実用化を想定して、LFC画像を高画素化する手法を提案する。基本的な考え方はMAP(Maximum A Posteriori)型超解像処理を利用[3][4]することであるが、超解像処理の前処理である複数画像の位置合わせに、マルチベースラインステレオ法(以下MBLS法と略記)[5]に基づくステレオ計測を利用する手法を提案する。

2 MAP型超解像処理と位置合わせ手法

2.1 MAP型超解像処理

複数画像からの超解像処理は、高周波成分の折り返しを含むように標本化された画像群を処理することで、高周波成分を含む高解像度画像を推定する技術である。MAP型超解像処理は、最も一般的な複数画像からの超解像処理である。MAP型処理は、図2.1に示すように、高解像度画像(\mathbf{I})から低解像度画像($\mathbf{I}_1, \dots, \mathbf{I}_n$)を生成する作用素 \mathbf{M} とともに、 \mathbf{I} に関する先見知識を用いて高解像度画像を推定する。低解像度化の作用素 \mathbf{M} は位置ずれの作用素 $\mathbf{T}_1, \dots, \mathbf{T}_n$ およびサブサンプル作用素 \mathbf{S} などから求める。良く利用される推定式は、 \mathbf{M} による誤差の2乗和と、 \mathbf{I} のラプラシアン(2次微分)の2乗和の合計を最小化するものである。ラプラシアンの2乗和の項を正則化項と呼ぶ。

図2.1において、高解像度画像 \mathbf{I} および i 番目の低解像度画像 \mathbf{I}_i を、画素値を1次元化した縦ベクトルによって表現すると、位置ずれ作用素 \mathbf{T}_i 、サブサンプル作用素 \mathbf{S} を行列表現することができ、これらの関係を式(2.1)と記述することができる。

$$\mathbf{I}_i = \mathbf{S} \cdot \mathbf{T}_i \cdot \mathbf{I} \cdots (2.1)$$

$i=1 \sim n$ について $\mathbf{S} \cdot \mathbf{T}_i$ を縦に並べると高解像度画像か

ら n 個の低解像度画像を生成する行列 \mathbf{M} を求めることができる。高解像度画像に対してラプラシアンを計算する行列を \mathbf{L} とする。 \mathbf{M} による誤差の2乗和と、 \mathbf{I} のラプラシアン²の2乗和の合計を最小化する条件では、 i 個の低解像度画像から高解像度画像を推定する行列を、式 (2.2)として得ることができる。

$$(\mathbf{M}^T \cdot \mathbf{M} + \lambda \mathbf{L} \cdot \mathbf{L})^{-1} \cdot \mathbf{M}^T \dots (2.2)$$

ここで \mathbf{M}^T は \mathbf{M} の転置行列を示す。 λ は正則化項の重みである。

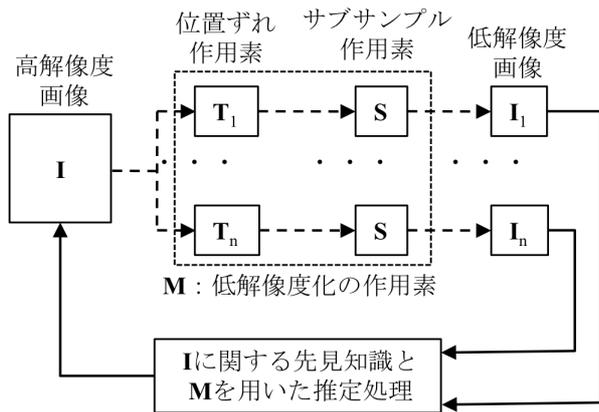


図 2.1 MAP 型超解像処理

2.2 複数画像の位置合わせ手法

超解像処理を行なう前処理として、複数画像を精度良く位置合わせする必要がある。画像の位置合わせ手法に、勾配法と領域ベースのマッチング法がある。勾配法はMAP型超解像処理でよく用いられる方法であるが、次の課題がある。

- 画素単位での位置合わせが難しい。
- 濃度勾配を利用するため、高周波成分の折り返しが誤差要因になる。

これらの問題を避けるため、

- 処理対象領域を、比較的大きな平面領域とする、
- ローパスフィルタ処理を施し、折り返し成分の影響を軽減する、

をなどが行なわれる。しかし、折り返し成分が低解像度画像に与える影響は予測できないため、それを強く含む低解像度画像を勾配法で位置合わせすることは難しい。

図 2.2 は解像度チャート画像の一部と、それに適当な位置ずれを加え、1/8 に間引いた低解像度画像である。低解像度画像間の位置ずれ量が既知であれば、このように折り返しが強く現れていても、超解像処理によってくさび型解像度チャートが完全に分離した高解像度画像

を生成することが可能である。一方、図示した低解像度画像のように、折り返しを非常に強く含む場合、その領域を勾配法で位置合わせすることはできない。

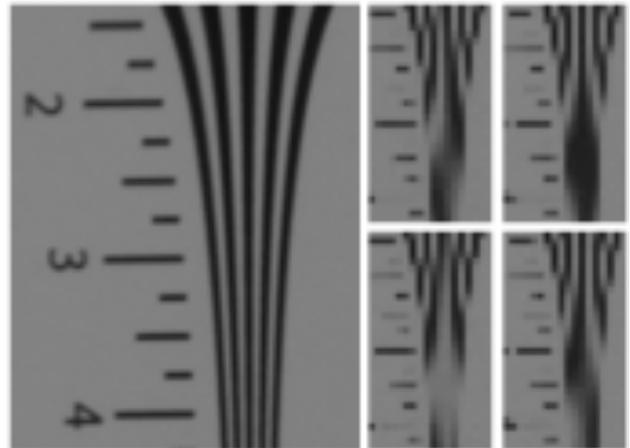


図 2.2 解像度チャートとそれを 1/8 に間引いた画像

領域ベースのマッチングはステレオ計測などでよく利用される方法である。勾配法に比べて、狭い領域(例えば 3×3 画素)間の位置合わせが可能である。本論文では、MBLS 法で位置合わせする手法を提案する。

3 ライトフィールドカメラ

ライトフィールドの考え方、LFC の原理およびマイクロレンズ LFC について説明する。

3.1 ライトフィールド

図 3.1 が示すように、ピンホールカメラで撮影した画像の個々の画素は、対応する被写体が発する一本の光線に対応する。ピンホールカメラを3次元空間に密に配置すれば、それらの画像群によって空間の光線状態を精度良く記録することができる。このように、光線集合によって記述される3次元空間をライトフィールドと呼ぶ。なお画像群の取得は、水平・垂直・前後の3次元的位置で行なう必要はなく、水平・垂直の2次元的に離散化された位置で行なえば十分である。

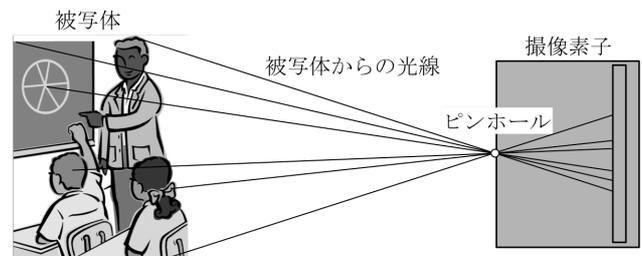


図 3.1 ピンホールカメラによる光線の取得

3.2 LFC の原理

通常のカメラは図 3.2 で示すように、被写体表面から発する光線の中でレンズに到達するものを集光し、撮像素子表面に像を作る。被写体からさまざまな方向に発する光線は撮像素子上で積算されるため、元の光線を分離することは不可能である。

レンズに到達する光線を分離して記録することは、図 3.3 に示すように、元のレンズ位置に多数の小口径カメラ(またはピンホールカメラ)を配置して画像を取得することで実現できる。各画像の各画素は、3次元空間での画素位置とカメラの投影中心、被写体上の1点を通る光線に対応する。このようにしてライトフィールドを取得すると、カメラによる撮影現象を後処理によって計算することができる。その代表的な処理が、リフォーカスと呼ばれるカメラ焦点位置を自由に変えた画像の生成である。

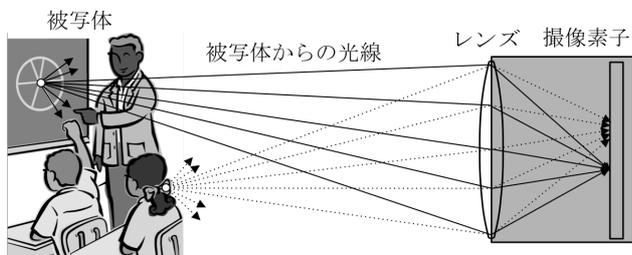


図 3.2 通常のカメラによる光線の結像

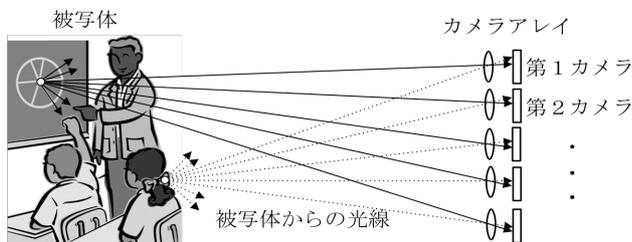


図 3.3 カメラアレイによるライトフィールドの取得

3.3 マイクロレンズ LFC

カメラアレイによる LFC は、原理を理解するには好都合であるが、実用的なカメラではない。そこで図 3.4 に示す、撮像素子の直前にマイクロレンズアレイを配置する構成が考案された。一つのマイクロレンズは撮像素子の 15×15 画素程度をカバーする大きさである。マイクロレンズアレイは撮像素子の 0.5mm 程度前、主レンズの焦点位置近傍に配置される。

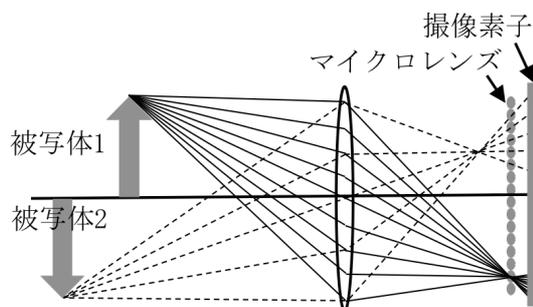


図 3.4 マイクロレンズ LFC の構成

図 3.5 を用いて、マイクロレンズ LFC とカメラアレイ LFC の関係を説明する。例えば、マイクロレンズ LFC の撮像素子の画素数が $3,840 \times 3,840$ で、一つのマイクロレンズが 15×15 画素をカバーし、マイクロレンズアレイの素子数が 256×256 とする。これを、カメラアレイ LFC に換算すると、 256×256 画素の小画像を取得する小カメラを、主レンズの口径上に 15×15 個配置するのと等価になる。この小カメラは、主レンズ口径を 15×15 に分割したレンズで 256×256 画素の画像を撮影するカメラと見なすことができる。実質的なレンズ口径が小さいため、その被写界深度は深くなる。ここではマイクロレンズ LFC の撮像素子が無駄無く利用されるように説明したが、実際には撮像素子の利用効率率は 100% ではない。したがって、マイクロレンズ LFC の最終画像はさらに小さいものになると予想される。

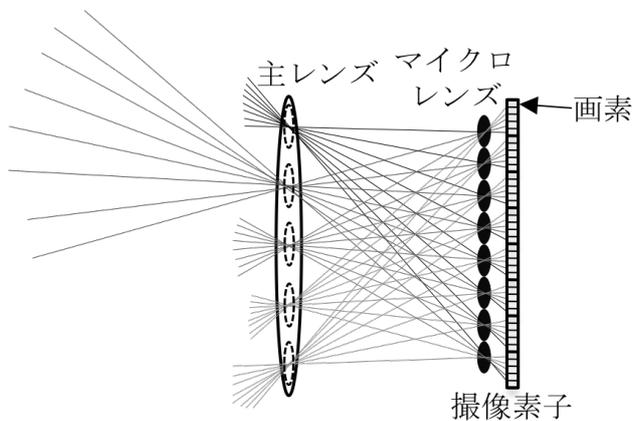


図 3.5 マイクロレンズ LFC とカメラアレイ LFC の関係

3.4 マイクロレンズ LFC の模擬装置

マイクロレンズ LFC とカメラアレイ LFC が 3.3 節で説明した関係にあるため、カメラアレイ LFC を使ってマイクロレンズ LFC の機能を模擬することができる。ただし、主レンズの口径内に複数のカメラを配置する必要があるため、カメラを併置することは難しい。今回利用したマイクロレンズ LFC の模擬装置を図 3.6 に示す。

精密メカニカル XZ ステージ(横・縦に移動)でデジタル一眼レフカメラを高精度に位置決めする。デジタル一眼レフの撮影パラメータをパソコンから制御し、撮影画像をパソコンに保存する。デジタル一眼レフの撮像素子は $3,456 \times 5,184$ (1,790 万) 画素であるが、今回は $1,728 \times 2,592$ (447 万) 画素の Small-RAW と呼ばれる RAW フォーマット画像を使用した。

マイクロレンズ LFC の機能を模擬するため、メカニカルステージによる単位移動量を 2mm に設定した。これは $15 \times 15 = 225$ 個の画像を取得する場合、カメラの全移動量が $28\text{mm} \times 28\text{mm}$ となり、使用するレンズの開放口径に収まる。従って、マイクロレンズ LFC で複数の低解像度画像をワンショット撮影するのと同様に撮影位置と見なすことができる。撮影時のカメラパラメータは絞り F36、シャッター速度 2.5sec 、焦点距離 100mm とした。F36 は設定可能な最も大きな F 値であり、計算上の有効口径は 2.8mm になる。カメラの特性的に F36 は絞りすぎであり、光の回折によるぼけが生じる。しかし、使用する生画像が比較的画素数の少ない Small-RAW であるためぼけの影響が軽微であり、 2.8mm の有効口径がカメラの移動量である 2mm に近い値であることから F36 の絞りで実験を進めた。

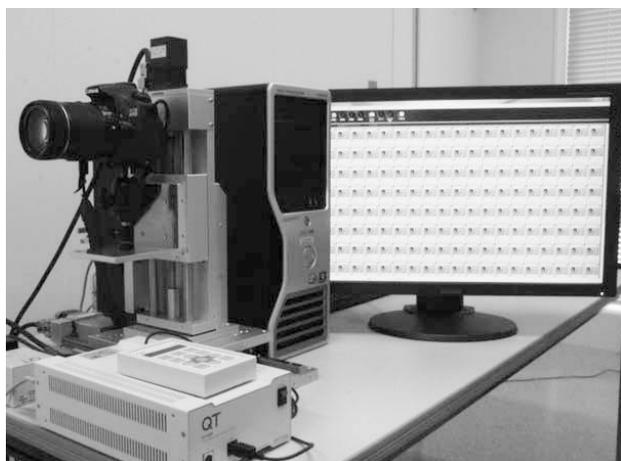


図 3.6 マイクロレンズ LFC の模擬装置

15×15 の各位置で撮影した $1,728 \times 2,592$ 画素の画像を $1/8$ に間引き、 216×324 画素の画像に縮小する。このように生成した 225 個の 216×324 画素の小画像を、マイクロレンズ LFC を模擬する入力画像とした。その画像の一枚を図 3.7 に示す。解像度チャートの部分に強い折り返し歪みが発生していることが確認できる。



図 3.7 マイクロレンズ LFC を模擬する入力画像の一枚

4 実験結果

15×15 に配列した 225 個の小画像から MBLS 法で視差を計算する。この視差情報に基づいて、入力画像群を位置合わせし、MAP 型超解像処理を施すことで高画素化する。

4.1 MBLS 法による視差の計算

MBLS 法は、複数のステレオ対から得る情報を統合することで、信頼性良く視差(画像間のずれ)を得る手法である。本論文における MBLS 法の設定を示す。

- 15×15 に配列する入力画像の中央画像を基準画像とし、その他の画像とペアをなす 224 個のステレオ対の情報を統合した。
- ステレオマッチングの領域を 1 画素とし、画素単位で視差を求めた。
- 視差探索範囲を $0 \sim 15$ 画素とし、その範囲を 1 画素単位で探索する。画素単位で得る視差の評価関数にパラボラフィティングすることでサブピクセル精度の視差を得た。
- パラボラフィティングにおける 2 次項の係数値がある閾値以下の場合、視差の信頼性が不十分であると判定する。

図 3.7 を基準画像として求めた視差画像を図 4.1 に示す。図 4.1 は視差 $0 \sim 15$ を輝度 $0 \sim 255$ に対応させて表示した。なお、信頼性が不十分な視差も輝度 0 として表示している。ここでの視差は、基準画像と最も外側の画像との視差である。

計算した視差の分布を図 4.2 に示す。今回の被写体は平面画像であり、図 3.6 の撮影装置に生対して設置した。したがって、視差は画面全体でほぼ等しいと考えられる。一方、別の手段で確認した真の視差は 7.25 画

素であった. 図 4.2 から視差の誤差は 2σ で ± 0.5 画素程度であり, また 8.5~10 画素に見られる視差が誤計測であることが分かる.

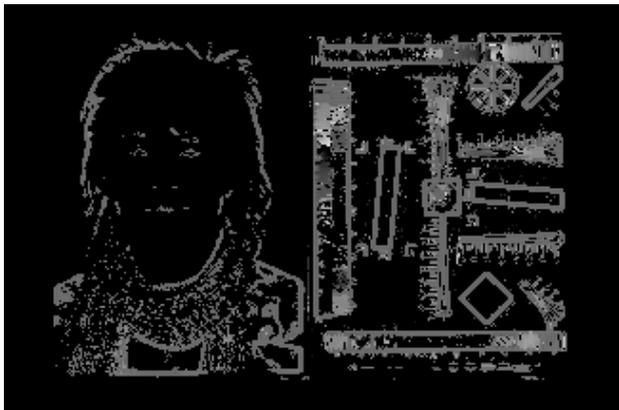


図 4.1 視差画像

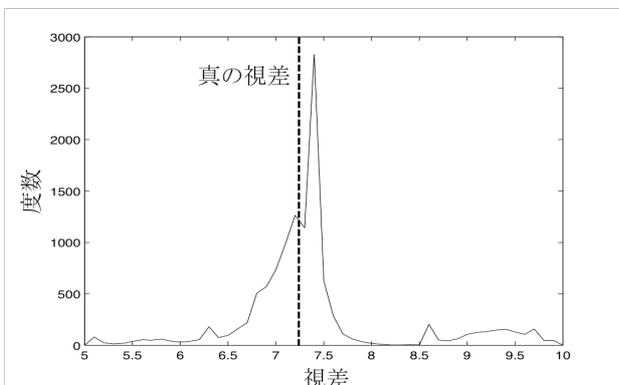


図 4.2 視差の分布

4.2 視差情報を用いた超解像処理

4.1 節の手法で得た視差情報を用いて 225 個の低解像度入力画像を位置合わせし, MAP 型超解像処理を用いて 8 倍に高画素化した. 図 4.3 に処理の流れを示す. ここでは, 入力となる 15×15 に配列した入力画像の中央に位置する基準画像を高画素化する場合について説明する.

基準画像のある画素に対する視差を d とする. 基準画像の画素および, それに対して視差 d を有するその他の画像の 224 個の画素を選択する. ここで, 視差が小数以下の桁を有する場合, バイリニア補間で画素値を求める. これによって MAP 型超解像処理に必要な 225 個の画素を準備する. 視差 d の信頼性が高い場合, d に応じた低解像度画像から高解像度画像への変換式を式(2.2)から求め, 225 個の低解像度画素にその変換式を適用することで, 基準画像の画素を 8 倍に高画素化する. 視差 d の信頼性が低い場合, 基準画像をバイリニア補間し, 画素を 8 倍に拡大する. この処理を基準画像の画素に順次適用することで, 基準画像全体を 8 倍に高画素化する.

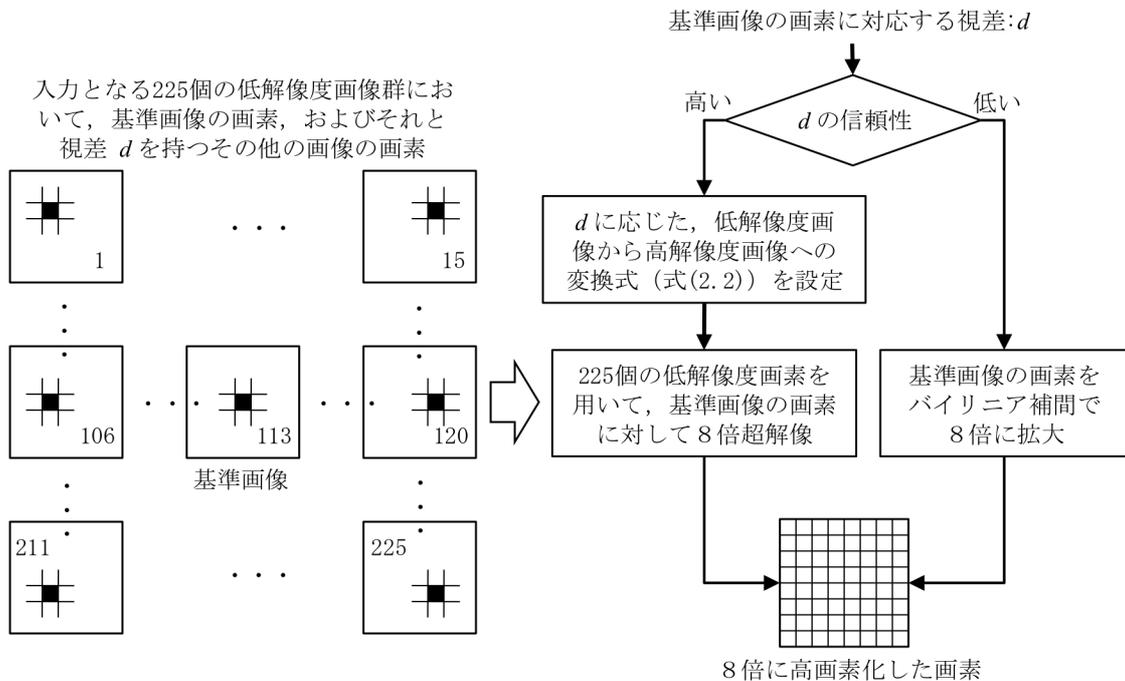


図 4.3 視差を用いた超解像処理手法

図 4.3～図 4.5 に、提案手法の効果を示す8倍超解像画像と、その他の比較画像を示す。いずれの比較も、低解像度の基準画像、基準画像をバイリニア補間で8倍に拡大した画像、提案手法で8倍に超解像処理した画像、真の視差を用いて8倍に超解像処理した画像の4枚を示した。

図 4.3 を観察すると、くさび形解像度チャートの低周波領域において、低解像度画像とそれをバイリニア補間した画像では、階段状の濃淡変化が見られる。それに対して、提案手法ではおおむね滑らかに変化するよう高画素化されている。しかし、この領域でも一部にはバイリニア補間と超解像処理が混在することによる不自然な高画素化が見られる。また、低解像度画像やバイリニア補間では、全く元画像の性質が再現されていない高周波領域では、提案手法でも安定した効果が得られていない。これは、入力低解像度画像に、高周波成分の折り返しが強く現れた場合、MBLS 法で安定した視差が得られないことが原因である。

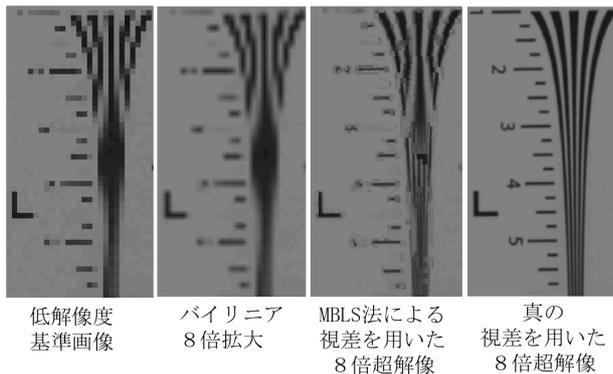


図 4.3 入力画像と高画素化画像の比較1

図 4.4 は、強いエッジが存在する、黒く塗りつぶされた領域の比較である。このようなエッジは、低周波成分から高周波成分まで混在する。この場合、MBLS 法による視差は精度良く計算されており、その結果、提案手法による高画素化は、真の視差を用いた高画素化と同程度の効果を示している。

図 4.5 の領域は、図 4.1 の視差画像を確認すると、視差が部分的に計算されている領域である。この領域でも提案手法がある程度の効果を示している。濃淡変化が存在する領域は、MBLS 法で正しく視差が計算されている可能性は高いが、今回用いた信頼性判定の閾値では、多くの画素の視差は信頼性が低いと判断されている。したがって、MBLS 法の設定を改良することで、高画素化の効果を改善できる可能性がある。

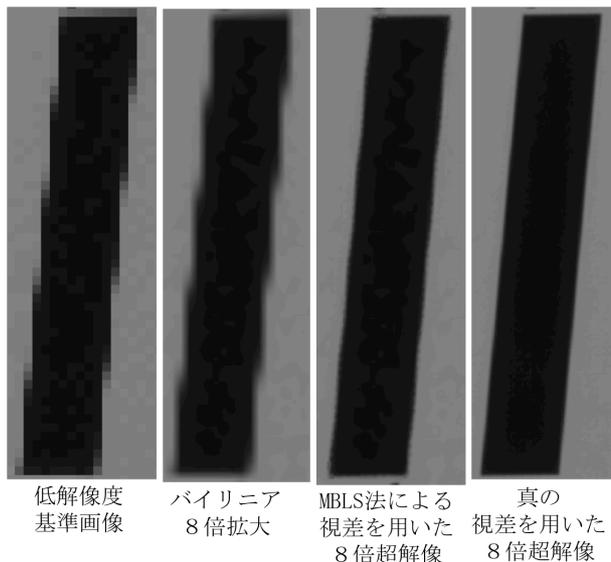


図 4.4 入力画像と高画素化画像の比較2

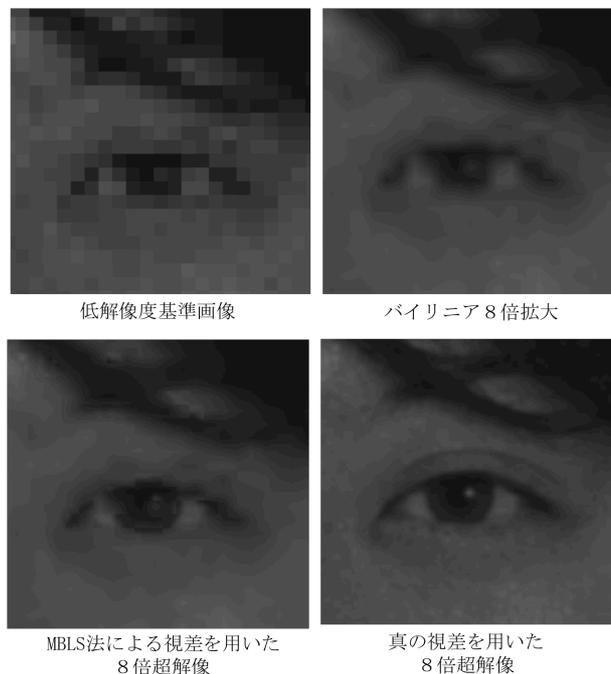


図 4.5 入力画像と高画素化画像の比較2

5 結論

複数画像からの超解像処理における位置合わせに、MBLS 法で計算する視差を用いる手法を提案した。提案手法は、低周波成分と高周波成分が混在する塗りつぶし領域のエッジ部で高い効果を示した。また、顔画像における目のように、ある程度の濃淡変化がある領域でも効果があることを確認した。一方、高周波成分が支配的で、入力低解像度画像に折り返しが強く現れる場合には、MBLS 法で正しく視差を計算することができないため、提案手法は有効ではなかった。

参考文献

- [1] B. Wilburn, N. Joshi, V. Vaish, E. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz and M. Levoy, "High Performance Imaging Using Camera Arrays," ACM Trans. on Graphics, Vol. 24, No.3, pp.765-776, July 2005.
- [2] R. Ng, M. Levoy, M. Bredif, G. Duval, M. Horowitz and P. Hanrahan, "Light Field Photography with a Hand-Held Plenoptic Camera," Stanford University Computer Science Tech Report CSTR 2005-02, April 2005.
- [3] R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," IEEE Trans. on Image Processing, Vol. 5, Issue 6, pp. 996-1011, June 1996.
- [4] 田中正行, 奥富正敏, "周波数領域最適化法による MAP 型超解像処理の高速化," 情報処理学会論文誌 CVIM, Vol. 47, No.SIG10, pp. 12-22, 2006 年 7 月.
- [5] 奥富正敏, 金出武雄, "複数の基線長を利用したステレオマッチング," 信学論, Vol.J75-D2, No.8, pp1317-1327, 1992 年.