

# MAP 型超解像処理における復元誤差を考慮した ライトフィールドカメラ画像の高画素化

蚊野 浩<sup>†</sup>

<sup>†</sup> 京都産業大学コンピュータ理工学部 〒603-8555 京都市北区上賀茂本山

E-mail: <sup>†</sup> kano@cse.kyoto-su.ac.jp

**あらまし** 撮像系に内蔵したマイクロレンズを用いて光線空間を取得するライトフィールドカメラは、一枚の撮像素子でライトフィールドを取得するため、最終画像の画素数が非常に小さくなる。この問題の解決に MAP 型超解像処理が有効である。そこでの課題の一つは、高周波成分の折り返しを強く含む複数画像を画素単位で位置合わせすることである。複数画像間の画素ごとの位置ずれ量に対する、超解像処理画像と入力画像の誤差を評価すると、正しい位置ずれ量で誤差が極小値をとる。これを利用して、勾配法などでは不可能な折り返し成分を強く含む画像間での位置合わせを可能とし、画素単位の超解像処理を実現した。

**キーワード** デジタルカメラ, ライトフィールドカメラ, 超解像処理, MAP 推定, 折り返し成分

## 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。そこでは、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、この画像群に一種の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ことである。

LFC の開発は[1]に遡る。これは大規模なカメラシステムで、実用性にはとぼしい。これに対して、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光線空間を取得可能な LFC を発明した（以下マイクロレンズ LFC と略記）[2]。この技術をベースにして、Lytro 社が安価でコンパクトなマイクロレンズ LFC を製品化している[11]。

Ng らのマイクロレンズ LFC は優れたものであるが、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  $256 \times 256$  画素になり、最終画像もこの大きさになる。

本論文は、マイクロレンズ LFC の最終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、超解像処理における復元誤差を用いることで、従来手法では不可能な高周波成分の折り返しを強く含む画像群を超解像処理することを可能にする。

## 2. 従来技術

### 2.1. ライトフィールドカメラ

ライトフィールドは 3 次元空間を満たす光線集合と定義される。図 2.1 は 3 次元空間に存在する  $xy$  平面上の点  $p$  における光線の分布を模式的に示している。 $xy$  平面上で、十分に密に点  $p$  をサンプリングし、かつ、各点  $p$  で光線分布を十分に密にサンプリングすることができれば、これらの光線集合を用いてカメラによる撮影現象をシミュレーションし、写真画像を生成することが可能である。このような原理に基づくカメラシステムをライトフィールドカメラと呼ぶ[10]。

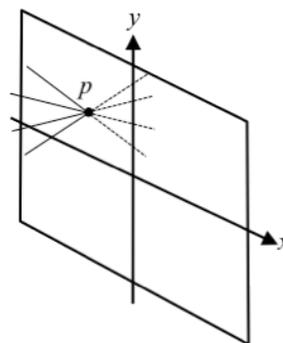


図 2.1  $xy$  平面上の点  $p$  における光線

多数の小型カメラをアレイ状に並べた、ライトフィールドカメラの機能試作が開発されている[1]。これは、上記の説明によるライトフィールドをそのまま取得するカメラシステムである。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 2.2 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  $15 \times 15$  画素程度をカバーする大きさで、撮像素子の  $0.5\text{mm}$  程度前に配置される。例えば、

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  $15 \times 15$  画素をカバーすれば、マイクロレンズアレイの構成は  $256 \times 256$  個になる。このように設計することで、カメラアレイ構成に換算したとき、 $256 \times 256$  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  $15 \times 15$  個配置するのと等価な LFC になる。

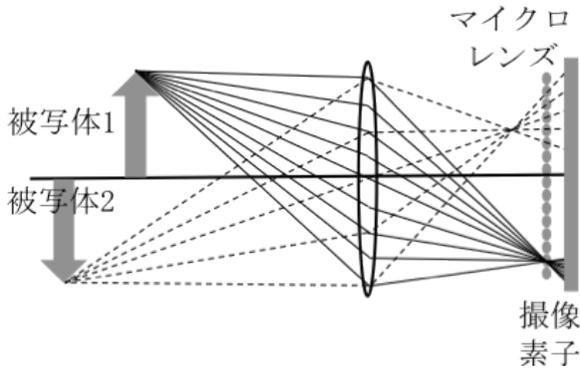


図 2.2 Ng らのマイクロレンズ LFC

Lumsdaine からもマイクロレンズを用いた LFC を開発している[10]。Lumsdaine らのものは、主レンズが結ぶ光像をマイクロレンズで再撮影する構成になっており、光線方向を直接記録する Ng らのものとレンズシステムの構成が異なる。本論文は、Ng らの LFC を前提した超解像処理を提案する。

## 2.2. ライトフィールドカメラ画像の超解像処理

マイクロレンズ LFC が取得する小画像群には、高周波成分の折り返しが強く生じる。折り返しを含む複数画像を合成処理することで高解像度画像を生成する超解像処理が、古くから研究されている[8]。MAP 型超解像処理[3][4]は、最も一般的な複数画像からの超解像処理であり、ライトフィールドカメラ画像の高画素化に有効である。

複数画像からの MAP 型超解像処理は、処理の一部として複数画像の位置合わせを行う。多くの場合、前段で位置合わせを行い、後段で超解像度処理が行われる。一方、折り返しが強く現れる複数画像を高精度に位置合わせすることは容易でない。これを解決するために、位置合わせと超解像処理を交互に繰り返すことで、位置合わせと高画素化を同時に行う手法が提案されている[5]。本論文の提案手法は、これと類似しているが、複数画像がステレオ画像を構成することを利用して、位置合わせの確実性を向上させることが可能である。

ライトフィールドカメラ画像を超解像処理する研究に[6][7]がある。[6]は提案手法と関係が深い従来手法である。ライトフィールドの超解像度処理に変分ベイズ推定を用いている。[7]は Lumsdaine らの LFC に用いる超解像処理を提案している。MAP 型超解像処理のようにデコンボリューション演算による超解像化ではなく、間引いてサンプリングされたライトフィールドをすき間無く整列させ

ることで折り返し成分を高周波成分として再現する手法である。

## 3. 提案手法

マイクロレンズ LFC 画像に MAP 型超解像処理を用いる場合の課題について再考し、提案手法の骨子である超解像処理における復元誤差を説明する。

### 3.1. 複数画像からの超解像処理における課題

マイクロレンズ LFC で取得する複数の小画像は、同じ被写体を異なる位置から撮影したステレオ画像である。それらの中には被写体までの距離に応じた位置ずれ（視差）が生じる。それらを超解像化するためには、複数画像を画素単位でかつサブピクセル精度で位置合わせする必要がある。

画像間の位置合わせ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。

- (1) 画素単位での位置合わせが難しい。
- (2) 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、

- (a) 処理対象領域を、位置ずれ量が変化しない比較的大きな領域とする、
- (b) 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 3.1 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縦線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。

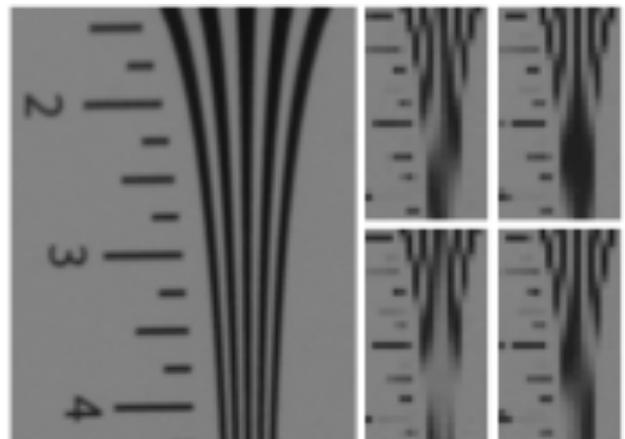


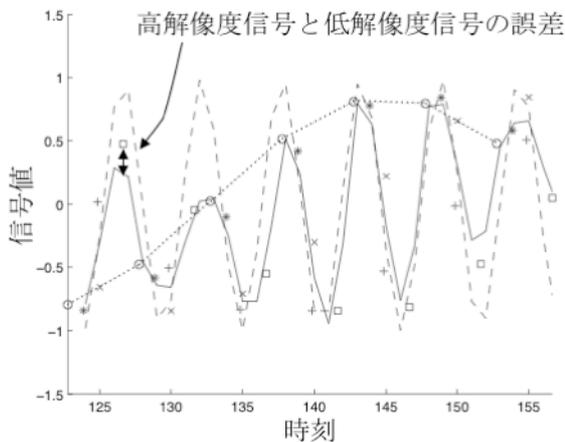
図 3.1 解像度チャートとそれを 1/8 に間引いた画像

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、比較的狭い領域（例えば3×3画素）間の位置合わせが可能である。しかし、この方法でも折り返しが非常に強く現れた部分では誤対応が発生する。このように、勾配法と領域ベースのマッチングは、折り返しを強く含む画像の位置合わせが難しい。

### 3.2. 超解像処理における復元誤差

MAP 型超解像処理は高精度な位置合わせが必要である。低解像度画像が誤差ゼロで位置合わせされている場合、サンプリング定理によれば、全画素の画素間隔の中で最も広い画素間隔の2倍を周期とする周波数まで、高解像度画像で再現することができる。ここでは、位置合わせ誤差と推定される高解像度画像の関係を検討する。以下では、理解を容易にするため1次元信号を用いたシミュレーションで説明する。

図3.2は0.18Hzの正弦波を1sec単位で標本化したものを元の低解像度信号とし、それにランダムな位置ずれ（位相ずれ）を加え、1/5に間引いた信号を○、□、\*、×、+で表す。正しい位置ずれ量を用いて5倍超解像処理を施すことで得る高解像度信号を破線で示す。破線は0.18Hzの正弦波を1sec間隔で標本化したものと一致した。正しい位置ずれ量にランダムな誤差を加え（平均的には20%程度の誤差）5倍超解像したものを実線で示す。推定された高解像度信号は元の低解像度信号と大きな誤差を持ち、また、低解像度信号とも誤差を有している。なお、点線は○の信号系列を結んだ折線であるが、低解像度信号に折り返しが強く現れていることがわかる。



○、□、\*、+、×：低解像度信号列  
 実線：位置合わせ誤差が大きい5倍超解像  
 破線：位置合わせ誤差のない5倍超解像  
 点線：○の信号系列を結んだ線分

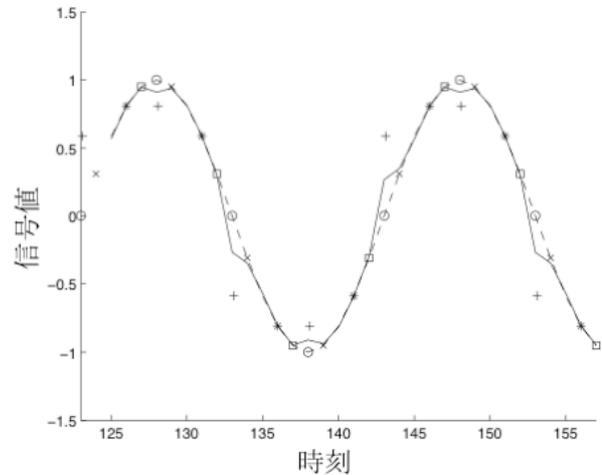
図3.2 1次元信号を超解像処理したときの  
 高解像度信号と低解像度信号の誤差

このように、低解像度信号の位置合わせに誤差を含む

と、推定される高解像度信号と低解像度信号の間に誤差を生じる。低解像度信号の位置関係を間違えて高解像度化するため不要な高周波成分を生成することや、逆に最小2乗的に最適解を求める手法であるため高解像度信号にぼけが発生するという現象の、両方が発生しており、これが誤差要因になる。

図3.2と同様に5系列の低解像度信号からの5倍超解像であって、一つの系列だけが位置ずれ誤差を有する時の効果を確認した。図3.3において、○、□、\*、×は-2sec、-1sec、1sec、2sec位置ずれした1/5低解像度信号である。+は1/5低解像度信号で、その位相に-2secから2secまで誤差を与える。図は-2secの誤差を与えた状態である。誤差を有する状態で超解像処理を行なうと、高解像度信号が誤差を有する低解像度信号に引っ張られる傾向があることを確認できる。

図3.4に、+系列の位置誤差量を横軸として、高解像度信号と低解像度信号の差の絶対値の平均を示す。位置誤差がゼロであれば、高解像度信号と低解像度信号は一致する。位置の誤差が増すにしたがって、信号値の差も増える傾向にある。なお、これらの傾向は、MAP法における正則化項の重みによって制御することができる。



○、□、+、\*、×：低解像度信号列（+の信号だけがずれている）  
 実線：+の信号が位置誤差を有する5倍超解像  
 破線：位置合わせ誤差のない5倍超解像

図3.3 複数の低解像度信号列の中で  
 一つの系列が誤差を有するときの超解像処理の例

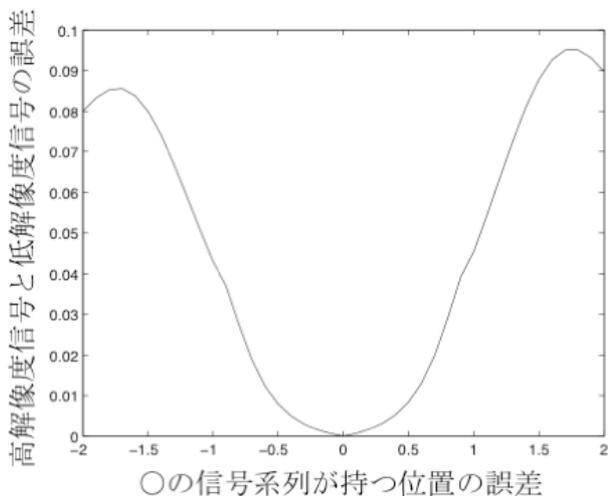


図 3.4 位置合わせ誤差と

高解像度信号と低解像度信号の誤差の関係

#### 4. LFC 画像の高画素化

マイクロレンズ LFC を模擬する装置を用いて、提案手法の検証を行なった。

##### 4.1. マイクロレンズ LFC の模擬装置

図 4.1 にマイクロレンズ LFC を模擬する装置を示す。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蓄積する。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に、メカニカルステージによる移動距離を 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は軽微である。また、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。

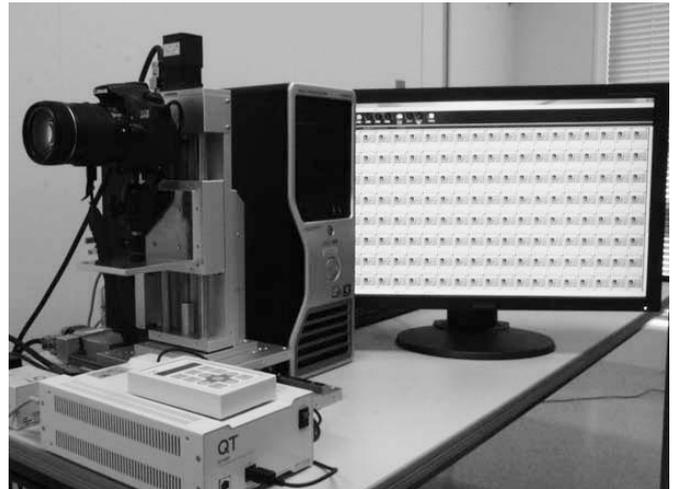


図 4.1 ライトフィールドカメラの模擬装置

上記の設定で 225 枚の  $1,728 \times 2,592$  画素の画像を取得する。これを 1/8 間引き処理によって 225 枚の  $216 \times 324$  画素の画像に低解像度化した。この処理は、 $216 \times 324$  素子のマイクロレンズアレイを用い、一つのマイクロレンズが  $15 \times 15$  画素をカバーする設定を模擬するものである。図 4.2 に、この模擬装置で撮影した画像の一枚を示す。 $1,728 \times 2,592$  画素の画像を単純に 1/8 に間引いたため、解像度チャートの部分に非常に強い折り返しが生じていることがわかる。

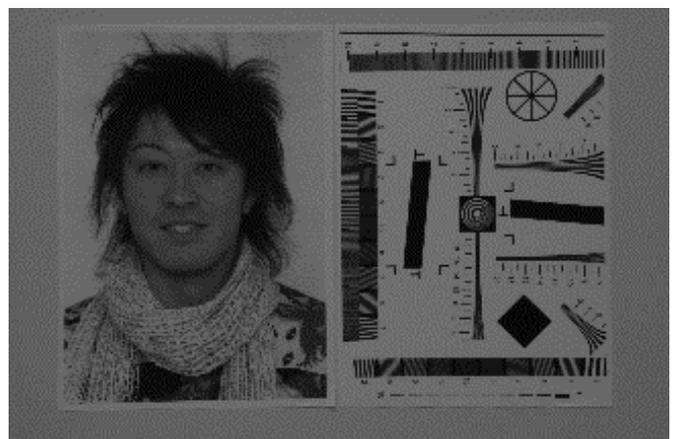


図 4.2 マイクロレンズ LFC の模擬装置で撮影した画像の 1 枚

#### 4.2. 高解像度画像の生成および高解像度画像と低解像度画像の誤差の評価手順

このように準備した 225 枚の低解像度画像を、高解像度画像との誤差評価を用いつつ MAP 型超解像処理で 8 倍に高画素化した。その手順と、推定された高解像度画像と低解像度画像の誤差を評価する手順を、図 4.3 を用いて説明する。

225 枚の画像は縦横  $15 \times 15$  に配列している。その中央画像をベース画像と呼ぶことにする。ベース画像に注目

画素を設定する。その他の 224 枚の画像は、ベース画像に対して被写体までの距離に応じた位置ずれ量（視差）を生じる。真の位置ずれ量は初期状態において不明であるが、225 枚の画像は均等間隔で移動したカメラで撮影したため、隣接画像間の真の位置ずれ量は同一である。隣接画像間の位置ずれ量  $d_1 < d_2 < d_3 < \dots$  を順次生成し、そのずれ量に応じた対応画素を 225 枚の画像から抽出する。位置のずれ量は小数点以下の精度を持つため、バイリニアによる画像補間を行なう。このようにして、ベース画像の注目画素を 8 倍に高解像度化するための 225 個の低解像度画素を用意した。

仮定した位置のずれ量  $d_1$  に対して、MAP 型超解像処理による低解像度画像から高解像度画像への変換式を求め

る。この変換式を、上記の 225 個の低解像度画素に適用することで、注目画素に対する  $8 \times 8 = 64$  個の高解像度画素を生成する。注目画素をベース画像においてラスタ走査することで、全画素を高解像度化する。

225 個の低解像度画素と対応する高解像度画素の値の差の絶対値を平均化し、これを高解像度画像と低解像度画像の誤差とした。ここでもバイリニア法による画像補間を用いた。  $d_1 < d_2 < d_3 < \dots$  のように生成した位置のずれ量を横軸に、高解像度画像と低解像度画像の誤差を縦軸にプロットすると、真の位置ずれを示す位置で明確な極小値（通常は最小値）を示す。この位置ずれ量における高解像度画像を最終結果とする。

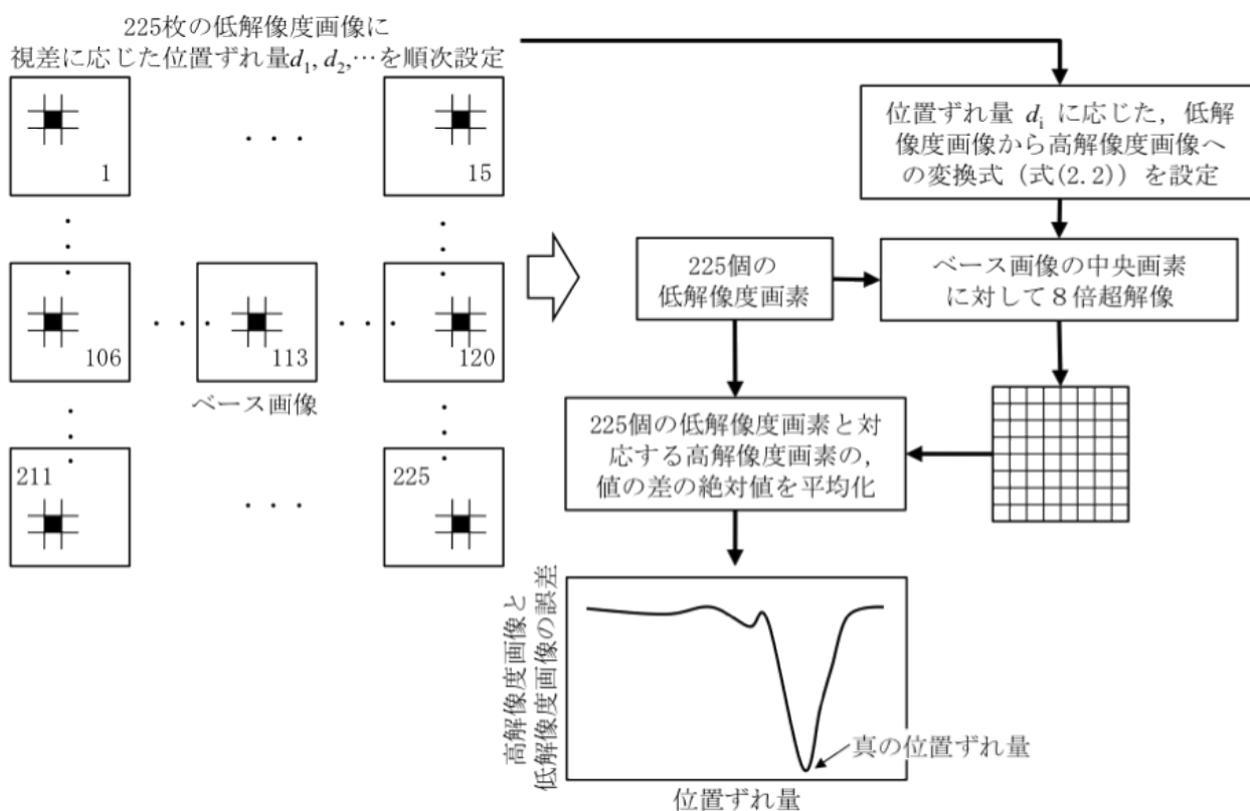


図4.3 提案手法における高解像度化および高解像度画像と低解像度画像の誤差評価の手順

### 4.3. 実画像を用いた提案手法の検証

マイクロレンズ LFC を模擬する装置を用いて縦横  $15 \times 15 = 225$  箇所画素を撮影し、図 4.2 と同様の  $216 \times 324$  画素の画像を 225 枚取得した。225 枚の画像を MAP 型超解像処理で 8 倍に高画素化した。図 4.4 に入力された低解像度画像の一部（中央画像と上下左右端の画像）と、それらを 8 倍に高画素化した画像の一例を示す。



図 4.4 低解像度入力画像と 8 倍に高画素化した画像

図 4.4 の MAP 型超解像処理において、隣接する低解像度画像間の真の位置ずれ量は約 0.9 画素であった。真の位置ずれ量を求めるために、位置ずれ量をパラメータとして、低解像度画像と推定した高解像度画像の誤差を評価した。図 4.5 は、横軸を隣接する低解像度画像間の推定位置ずれ量、縦軸をその位置ずれ量における高解像度画像と低解像度画像の誤差として、低解像度画像の中央付近の 100 個の画素についてプロットしたものである。真の位置ずれ量である 0.9 画素近辺で、全画素の評価値が最小になった。

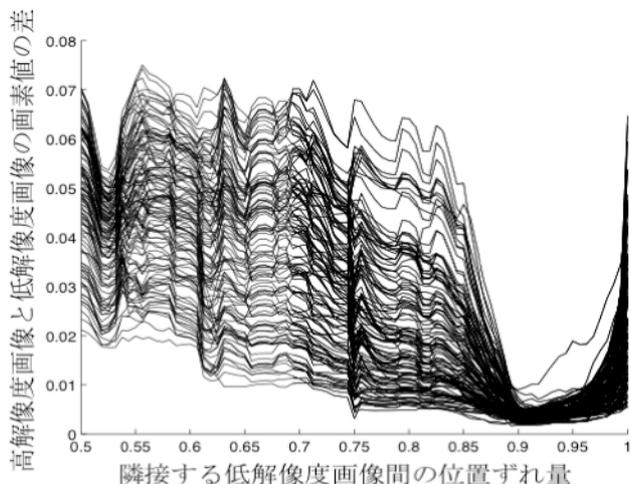


図 4.5 図 4.4 の例における隣接する低解像度画像間の位置ずれ量と高解像度画像と低解像度画像の誤差の関係

図 4.4 と図 4.5 で、提案手法が正しく動作することを確認した。次いで、高周波成分の折り返しを強く含む低解像度画像に提案手法を適用した。図 4.6 右は折り返し成分を強く含む 225 枚の低解像度画像の一部（中央画像と上下左右端の画像）と、それらを 8 倍に高画素化した画像である。折り返しを強く含んだ低解像度画像であっても、真の位置ずれ量を用いて MAP 型超解像処理を行なうことで、高解像度画像に高周波成分が正しく再現される。

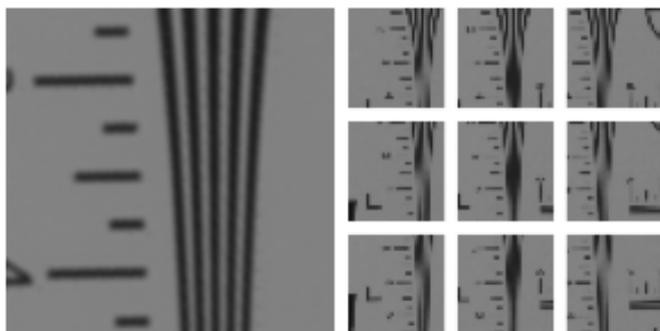


図 4.6 折り返し成分を強く含む低解像度画像とそれらを 8 倍に高画素化した例

図 4.6 の例についても、真の位置ずれ量を求めるために低解像度画像と高解像度画像の誤差を評価した。図 4.7 にその結果を示す。図 4.5 と同様に、全画素の評価値が位置ずれ量 0.9 画素近辺で最小になった。図 4.5 と比較すると、位置ずれ量 0.9 画素近傍での谷形状がより急峻になっている曲線と、ほぼ平になった曲線が混在している。これは、解像度チャートのエッジにあたる画素では位置ずれ量を精度良く求めることができ、濃淡変化がない領域では位置ずれ量の決定が難しいことを示している。

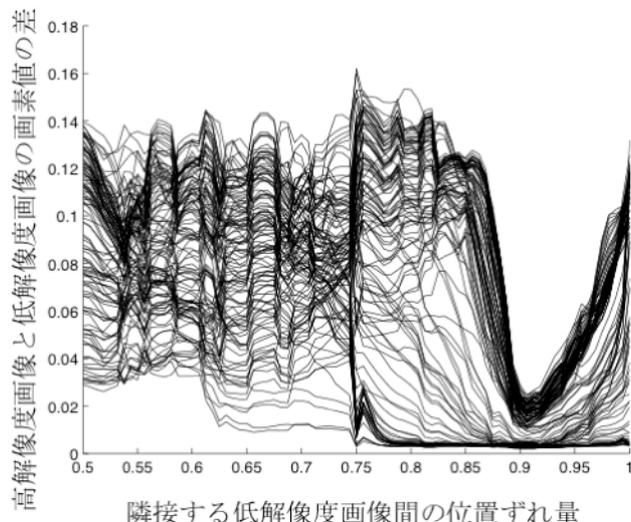


図 4.7 図 4.6 の例における隣接する低解像度画像間の位置ずれ量と高解像度画像と低解像度画像の誤差の関係

## 5. 結論

MAP 型超解像処理において、入力である複数の低解像度画像と推定した高解像度画像の誤差に基づいて、正しい高解像度画像を得ることが可能であることを示した。この手法は、高周波成分の折り返しを強く含む低解像度画像の位置合わせに特に有効である。また、提案手法をライトフィールドカメラ画像の高画素化に応用し、その有効性を検証した。

提案手法は、まだ初期的な動作確認を行なった段階である。今後、その性質をより詳細に研究し、適用可能な範囲を明らかにする予定である。また、今回は一定の距離にある被写体だけを対象にしたが、凹凸のある被写体やジャンプエッジが存在する被写体への応用可能性を検証する予定である。

## 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] 後藤知将, 奥富正敏, "画像復元とレジストレーションの同時最適化の実験的検証," 電子情報通信学会論文誌, Vol. J90-D, No. 6, pp. 1632-1635, 2007年
- [6] T.E. Bishop, S. Zanetti and P. Favaro, "Light Field Superresolution," IEEE Int. Conf. on Computational Photography, pp.1-9, 2009.
- [7] T. Georgiev, G. Chunev and A. Lumsdaine, "Superresolution with the Focused Plenoptic Camera," In Proc. SPIE 7873, 2011. doi:10.1117/12.872666.
- [8] S. C. Park, M. K. Park and M. G. Kang, "Super-resolution image reconstruction: a technical overview," IEEE Signal Processing Magazine, Vol. 20, Issue 3, pp. 21-36, May 2003.
- [9] M. Levoy and P. Hanrahan, "Light Field Rendering," ACM Proceedings of Siggraph'96, pp. 31-42, 1996.
- [10] A. Lumsdaine and T. Georgiev, "The Focused Plenoptic Camera," IEEE Int. Conf. on Computational Photography, pp. 1-8, 2009.
- [11] <http://www.lytro.com/>