

## 手ぶれビデオ映像からの超解像処理手法の提案

山口 拓真<sup>†</sup> 福田 悠人<sup>††</sup> 古川 亮<sup>†††</sup> 川崎 洋<sup>††††</sup> 久野 義徳<sup>††</sup>

PeterSturm<sup>††††</sup>

<sup>†</sup> 日本信号株式会社研究開発センター 〒 346-8524 埼玉県久喜市大字江面字大谷 1836-1

<sup>††</sup> 埼玉大学 工学部 〒 338-8570 埼玉県さいたま市桜区下大久保 255

<sup>†††</sup> 鹿児島大学 工学部 〒 890-0065 鹿児島県鹿児島市郡元 1-21-40

<sup>††††</sup> 広島市立大学 情報科学部 〒 731-3194 広島県広島市安佐南区大塚東 3-4-1

<sup>†††††</sup> INRIA Grenoble – Rhone-Alpes 655 Avenue de l'Europe, 38330 Montbonnot St Martin, France

あらまし ビデオは手持ちで撮影されることも多く、また、カメラに比べて解像度が低いため、動画から1枚の切り出した際には、モーションブラーや解像度不足が目立つ場合が多い。本論文では、ビデオで撮影された動画から、モーションブラーも考慮した高解像度画像の生成手法を提案する。

キーワード 超解像処理, モーションブラー, デフォーカスブラー, 3次元形状物体

## Super resolution technique for video data containing motion blur

Takuma YAMAGUCHI<sup>†</sup>, Hisato FUKUDA<sup>††</sup>, Ryo FURUKAWA<sup>†††</sup>, Hiroshi KAWASAKI<sup>††††</sup>,  
yoshinori KUNO<sup>††</sup>, and Peter STURM<sup>†††††</sup>

<sup>†</sup> Research & Development Center, The Nippon Signal Co., LTD

<sup>††</sup> Faculty of Engineering, Saitama University

<sup>†††</sup> Faculty of Engineering, Kagoshima University

<sup>††††</sup> Faculty of Information Sciences, Hiroshima City University

<sup>†††††</sup> INRIA Grenoble – Rhone-Alpes

**Abstract** Since resolution of video camera is usually lower than digital camera and video data usually contains a motion blur in the sequence, simple frame capture can produce only low quality image. In this paper, we propose a method to restore a sharp and high-resolution image from a video sequence by motion deblur for each frame followed by super-resolution technique. By using geometric information, we first apply motion deblur for each frame, and then, super-resolve the images from the deblurred image set.

**Key words** Super-Resolution, Motion Blur, Defocus Blur, 3D Object

### 1. ま え が き

ビデオで撮った映像、例えば、手持ちカメラの映像や、監視カメラの映像などから、一枚の高画質な静止画が必要とされることは良くある。しかし、一般にビデオの品質はカメラに比べて低いため、ビデオから単純に画像を切り出すだけでは解析に用いるには不十分な場合がある。各フレームの品質が低い原因は、ビデオの解像度が低いことや、手振れや撮影シーン内の動きによるモーションブラーが含まれていることが挙げられる。

このような問題に対して、過去のアプローチの一つとして、2台のカメラを使う方法[1],[2]が提案されているが、対象ビデオカメラ以外にデバイスが必要でありシステムが複雑化すると

いう問題がある。また、超解像においては、モーションブラーが含まれていても、それらをノイズ(特にL1ノイズとして扱う方が良くとされている)として扱われることが多く、画質改善は限られる[3]。

そこで、我々は、最初にフレーム毎にモーションブラーの除去を行い、それらフレームを用いて超解像処理を行うことで画質の改善を行う。モーションブラーの除去を行うためには、動きの情報が必要であるが、モーションブラーが含まれる入力画像に対しての動きの推定は難しい。また、複数のオブジェクトが独立して動いている場合、さらにこれは困難となる。本論文では、入力がビデオであることから、フレームレートが高く、前後フレームでブラーの大きさにも大きな変化がないことを仮

定できるため、ブロックマッチングによるオプティカルフローにより動きを推定する。

単純な超解像処理では、モーションブラーが含まれる画像に適用しても高い効果は得られにくいですが、本手法では、モーションブラーを事前に除去することによって超解像処理の効果を引き出すことができる。また、シーン中に複数の独立に移動する物体が存在していても、シーン全体を複数の平面で近似し、画素単位でレジストレーションを行うことによって、そういったシーンも扱うことが可能となる。さらに、その際に推定されるジオメトリ情報を用いて、奥行き毎にデフォーカスのスケールが異なることにも対応したカーネル推定も行う。

## 2. 関連研究

モーションブラーやデフォーカスブラーは、ブラーカーネルとのコンボリューションにより発生するため、これまでに既知のカーネルでのデコンボリューションによる復元方法が提案されてきた [4]。ブラーカーネルが未知の場合の単純な解決法としては、ブラーカーネルを推定するセンサーを用いることである [5]。近年では、自然画像の情報 (heavy tailed な勾配分布) を用いるブラインドデコンボリューション手法 [6], [7] が提案されており、本論文でも同様の手法によりモーションブラーカーネルの推定を行う。

デフォーカスブラーに関しては、対象とするシーンに無視できないほどの奥行きのある場合には、単一のブラーカーネルだけでは対応できないため、それぞれの奥行きに応じたカーネルを推定する手法 [8] が提案されている。我々もシーンに動きに応じた領域分割を行い、領域ごとにブラーカーネルの推定を行う。

超解像処理については、複数の低解像度画像から、高解像度画像を生成することが多く研究されてきた [9], [10]。それらの中では、奥行きを無視できるシーンが対象となっており、また、モーションブラーまでは考慮されていない。このような仮定の下では、フレーム間の位置合わせは、アフィン変換または平面射影変換により、十分な結果が得られる。より一般的なシーンに適用する場合、それらのシーンは 3D シーンであり動きのある非剛体な物体 (例えば、布) などから構成される。しかし物体の動きや視点変化のためにシーンの見えは変化するので、既存の超解像処理において、このようなシーンを扱うのは困難である。本論文では我々の提案した 3 次元形状に基づく超解像手法 [11] を用いる。

モーションデブラーリングと超解像処理を行う手法は Tai ら [2] により提案されている。基本的な考えは我々と似ているが、シーン内の動きを推定する方法は異なり、Tai らは追加のデバイスを用いているのに対し、我々はビデオデータのみを用いたものとなっている。

## 3. 提案手法の概要

本論文では、モーションブラーを除去した複数のフレームを利用して超解像処理を行う。図 1 に処理の流れを示す。まず、入力画像群からブロックマッチングによりオプティカルフロー

を計算する (図 1(a))。次に、オプティカルフローをグラフカットを用いて、それぞれの領域が一定な動きベクトルを持つように、領域分割を行う (図 1(b))。その後、領域ごとにブラーカーネルの推定を行い (図 1(c))、推定したモーションブラーカーネルを用いてモーションブラーの除去を行う (図 1(d))。それらのモーションブラーを除去した画像を入力とし、超解像処理によって高解像度画像を生成する (図 1 II)。

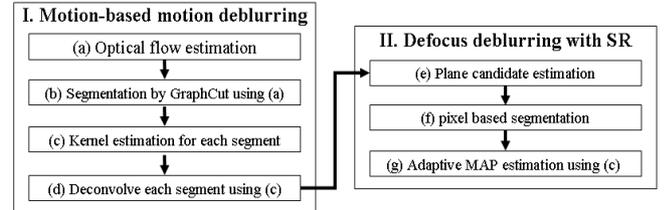


図 1 Flow of the deblurring process.

## 4. 複数の物体に対してのモーションブラー除去

本論文では、ブロックマッチングに基づくオプティカルフローを用いて領域分割し、これを元に各領域のモーションブラーカーネルを推定する。問題を単純にするため、本論文ではブラーカーネルを線状のモーションブラーカーネルとする。モーションブラーカーネルを線状としているため、カメラの動きが複雑である場合には、正しい復元ができないが、本研究では入力データはビデオ映像であり、通常シャッタースピードは十分速いため、フレームあたりのカメラの動きを、線状としたシンプルカーネルでも十分な復元が行える。

モーションブラーカーネルの方向についてはオプティカルフローの方向を用い、スケーリングの決定には、自然画像の勾配分布から推定を行う。このようなスケーリングパラメータの推定は、近年一般的に用いられている [6], [7]。実際のカーネル推定は以下ようになる。

- (1) ブロックマッチングによるオプティカルフロー推定
- (2) グラフカットによるオプティカルフローに基づく領域分割
- (3) 領域毎に動きベクトルの方向と、それに直交する方向の勾配分布から、モーションブラーカーネルを推定
- (4) 推定したカーネルを用いたモーションブラー除去

### 4.1 ブラー画像に対する領域分割

提案手法に用いる入力データは一連の画像群であり、撮影するカメラは、固定カメラまたは、手持ちカメラのように動いている。シーン中には、複数の物体があり、それらは動いている可能性があり、物体ごとの領域分割が必要となる。まず、オプティカルフローを推定し、得られたフロー領域をグラフカットによって分割を行う。モーションブラーが含まれている入力画像に対しては、特徴点ベースのオプティカルフローの推定はうまくいかない場合が多いため、領域ベースのピラミッド型のブロックマッチングを用いる。我々の実装では、動きベクトルの大きさより、方向に重みを置いて、多値グラフカットによる領域分割を行っている。また、計算の高速化のため、最大 5 つの

領域に分かれるものと仮定した。

#### 4.2 オプティカルフローを用いたブラーカーネル推定

自然画像においては、どの方向においても勾配分布が経験的に類似した分布となることを利用し、入力画像にブラー方向に直交する方向のモーションブラーを加えながら、ブラー方向とそれに直交する方向の勾配分布が最も近くなるようなブラーの大きさを求める。図 2(a)(b) は、ブラー方向の勾配分布とブラー方向に直交する方向の勾配分布を表している。図 2(b) を図 2(a) に近づけるようにして、ブラーの大きさを求めることとなる。図 2(c) は、図 2(b) が図 2(a) に最も近くなるようなモーションブラーを入力画像に加えた場合の勾配分布を示す。このブラーの大きさ推定処理を各領域に対して行うことで、それぞれの領域のブラーカーネルを自動で決定することができる。

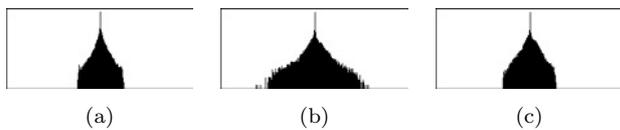


図 2 motion blur kernel estimation.

#### 4.3 領域ごとのモーションブラー除去

デコンボリューションアルゴリズムについては、幾つかの手法が存在するが、我々の手法では、Iterative Back Projection 法を用いる [12]。

### 5. 複数物体に対する超解像処理

複数の移動物体が含まれるシーンや 3 次元物体が含まれているなど奥行きが無視できないシーンにおける超解像処理は、サブピクセル精度での位置合わせが課題となる。本論文では、領域分割に基づく画素単位のレジストレーション方法によって、この課題を解決する。また、奥行きが異なる位置に存在する物体においては、デフォーカスブラーの度合いが異なることも考慮し、それぞれの物体に対して適応的にデフォーカスカーネルを決定する。物体ごとにブラーカーネルを推定することは簡単ではないため、デフォーカスブラーカーネルを 1 つのパラメータを持つ第一種一次ベッセル関数を用いることで単純化して扱う。シーン中の移動物体においては、同一の物体であってもデフォーカスブラーカーネルは変化する可能性があるが、我々が超解像処理に用いるのは 20 から 40 フレームであるため、シーケンス中では各々の物体のデフォーカスブラーは一定であるものとして扱う。

#### 5.1 特徴点トラッキングによる候補平面推定

3 次元復元のためのシーンからの平面抽出に関しては、これまで多くの報告がなされている [13], [14]。それらの報告の中で、平面の領域は特徴点のクラスタリングにより、パッチの集合として得られている。しかし、実際には、特徴点のトラッキングは、視点変化などにより見え方が大きく変わるなどのアウトライアにより、簡単に影響されるため、全体の動きを正確に推定できたとしても、正確な平面近似を行うことは困難である。加えて、特徴点は物体の境界によっては、しばしば得ることができないことがあるため、パッチを得る問題と別としても、困

難である。本論文では、パッチ単位の平面推定より適した、ピクセル単位の平面推定を提案する。具体的には、シーンを三角パッチに分割する代わりに、はじめに、候補平面群を推定し、各画素がどの平面に属するかを分類する。このため、平面候補群に正解平面が含まれるように、十分な数の候補を推定しておけば、陽に平面検出を行わずに 3 次元物体の超解像処理が実現できる。

平面候補群の推定アルゴリズムを Algorithm 1 に示す。はじめに、入力画像の特徴点を計算する。その後、特徴点の軌道により、最初の平面候補群を生成する。特徴点の軌道を用いて、基準フレームと他フレームとの平面射影変換行列を計算する。次に、平面候補群の更新を行う。特徴点において計算した評価値が閾値 (本論文では 0.2 ピクセル) より小さい場合、その点を平面を表す点群に追加する評価値には、すべての対応点での再投影誤差の平均の値を用いている。平面射影変換行列の計算、平面候補群の更新を、収束するまで繰り返し行うことで、シーン中の平面推定を行う。上記の処理の様子を図 3 で示す。ここでは、黒い点は既にいずれかの特徴点集合に含まれている点、白い点はまだどの特徴点集合にも属していない点を表している。最終的に、図 3 では、3 つの特徴点集合 (平面) が検出されたことになる。

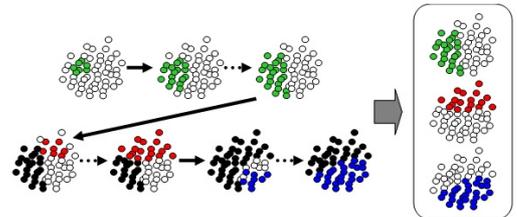


図 3 Candidate plane detection.

#### 5.2 最投影誤差最小化による画素単位レジストレーション

前述した手順によって得られた平面領域は陽に定義されたものではなく、特徴点集合から平面候補として推定されたものであるため、実際に各画素がどの平面に乗っているかは未だ分かっていない。このため各画素の対応関係も求まっていない。フレーム間の平面の射影変換行列はすでに求まっているので、画素の対応関係は各候補平面を構成する特徴点集合から得られる平面射影変換行列を用いて、基準フレームを各フレームへ射影した場合の輝度差の平均を画素毎に求め、最も小さい値を与える平面を各画素について選択することにより推定できる。

本論文では、フレーム間の平面の射影変換行列はすでに求まっているので、画素の対応関係は各候補平面を構成する特徴点集合から得られる平面射影変換行列を用いて、基準フレームを各フレームへ射影した場合の輝度差の平均を画素毎に求め、最も小さい値を与える平面を各画素について選択する。これは、入力フレーム数を  $N$ 、 $i$  番目の特徴点集合から得られる基準フレームから  $k$  番目のフレームへの射影変換行列を  $\mathcal{H}_i^{(k)}$ 、基準フレームおよび  $k$  番目のフレームの任意の点での輝度値をそれぞれ  $I(\cdot)$ 、 $I^{(k)}(\cdot)$  とした場合、

---

**Algorithm 1** Candidate plane estimation.

- 1:  $X$  is defined as the set of all corresponding feature point tracks across input frames.
  - 2:  $P(x)$  is defined as a predicate that is true if point track  $x$  is not selected and unlabeled.
  - 3: **while**  $\exists x \in X; P(x)$  **do**
  - 4:   Select a feature point track  $a(\subseteq \{x \in X; P(x)\})$  and the  $k$  nearest neighbors  $b(\subseteq X)$  (in this paper  $k := 7$ ).
  - 5:    $A^{(0)} := \phi$ ,  $A^{(1)} := a \cup b$ ,  $i := 1$
  - 6:   **while**  $A^{(i)} \neq A^{(i-1)}$  **do**
  - 7:     Compute the homography matrix  $\mathcal{H}$  of  $A^{(i)}$  for each frame.
  - 8:      $A^{(i+1)} := \phi$
  - 9:     **for**  $\forall y \in X$  **do**
  - 10:      **if** Adequateness of  $\mathcal{H}$  for  $y \geq \text{threshold}$  **then**
  - 11:        $A^{(i+1)} := A^{(i+1)} \cup y$
  - 12:      **end if**
  - 13:     **end for**
  - 14:      $i := i + 1$
  - 15:   **end while**
  - 16:    $A^{(i)}$  is a group of feature point tracks residing in the same plane.
  - 17: **end while**
- 

$$\hat{i}p = \arg \min_i \left[ \frac{\sum_{k=1}^M \left\{ I(p) - I^{(k)}(\mathcal{H}_i^{(k)} p) \right\}^2}{M} \right] \quad (1)$$

を基準フレームの各画素について求めることで計算される．ここで、 $p$  はある座標を表すベクトルであり、 $M (\leq N)$  は、射影先の画素が有効だった (画像内に存在していた) フレーム数を表す．画素  $\hat{i}p$  がどの平面に属するか決定される様子を図 4 に示す．

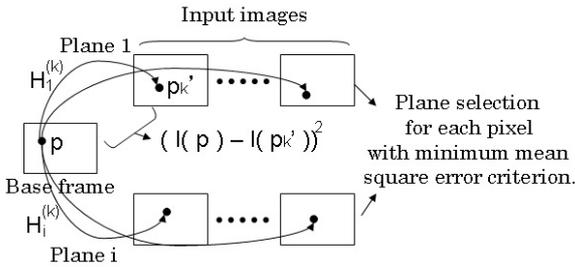


図 4 Plane selection for each pixel.

### 5.3 MAP に基づく超解像処理

本論文では、MAP(Maximum A Posterior) 法による定式化を用いた複数フレームを入力とする超解像処理によって画像復元を行う．これは以下の式によって表される．

$$\hat{X} = \arg \min_X \left[ \sum_{k=1}^N \|D_k H_k F_k X - Y_k\|_2^2 + \lambda \|\Gamma X\|_2^2 \right] \quad (2)$$

$F_k$  は、高解像画像  $X$  と  $k$  番目の低解像度画像  $Y_k$  との間の動きを表す． $H_k$  はデフォーカスブラーを表す行列であり、 $D_k$  はサブサンプリングを表す行列である． $\|\Gamma X\|_2^2$  は Tikhonov の

正則化項であり、 $\lambda$  は正則化項の寄与度合いを表すパラメータである．一般的に  $\Gamma$  にはハイパスフィルタが用いられ、本論文ではラプラシアンフィルタを用いた．

サブサンプリング演算が全て同じで (i.e.  $\forall k, D_k = D$ )、デフォーカスブラーが全て同じ場合 (i.e.  $\forall k, H_k = H$ )、式 (2) は以下のように表せる．

$$\hat{X} = \arg \min_X \left[ \sum_{k=1}^N \|DF_k HX - Y_k\|_2^2 + \lambda \|\Gamma X\|_2^2 \right] \quad (3)$$

この最小化問題を [3] と同様に以下の 2 つのステップに分けて解く．

(1) 低解像度画像からデフォーカスブラーが含まれる高解像度画像  $\hat{Z} (= H\hat{X})$  を計算する

(2) デフォーカスブラーが含まれる高解像度画像  $\hat{Z}$  から高解像度画像  $\hat{X}$  を推定する．

$\hat{Z}$  は以下の最小化問題を解くことで得られる．

$$\hat{Z} = \arg \min_Z \left[ \sum_{k=1}^N \|DF_k \hat{Z} - Y_k\|_2^2 \right] \quad (4)$$

デフォーカスブラーの除去された高解像度画像  $\hat{X}$  は以下の式により得られる．

$$\hat{X} = \arg \min_X \left[ \|W(HX - Z)\|_2^2 + \lambda \|\Gamma X\|_2^2 \right] \quad (5)$$

ここで、 $W$  は対角行列であり、対角成分の値は高解像度画像の各画素に対応する低解像度画像の画素の数となっている．式 (2) から分かるように、この定式化では、画素ごとにデフォーカスブラーカーネルを設定することが可能である．

## 6. 実験

### 6.1 実データを用いた実験

提案手法の効果を確認するために、モーターステージを用いた実験を行った．実験に用いたシーンは図 5(a) のようになっており、テクスチャのある平面 2 枚を対象とし実験を行った．実験では手前の平面にピントを合わせている、そのため奥の平面ではデフォーカスブラーが発生している．これら 2 つの平面を 2 台のモーターステージを用いて異なる方向に異なる速さで移動させながら、シーンの撮影を行った．提案手法を用いて超解像処理を行った結果は図 5(i) である．結果を見ると、エッジ周辺にはリングングが発生しているが、大きなデフォーカスブラーは除去されており、超解像処理がうまく行われている事が確認できる．

次に、複雑な曲面を持った対象を用いて実験を行った．結果を図 6 に示す．図 6(c) では、滑らかな曲面が幾つかの平面により近似できていることが分かる．図 6(h) では、平面でないものに対しても、モーションブラー除去が行われ、超解像処理がうまく行われていることが分かる．

### 6.2 手持ちカメラによる実験

実験では、図 7(a) に示すような手持ちカメラを用いた実験を行った．図 7(f) はモーションブラー除去を行った結果であり、モーションブラーが除去されているのが分かる．図 7(c) はモー

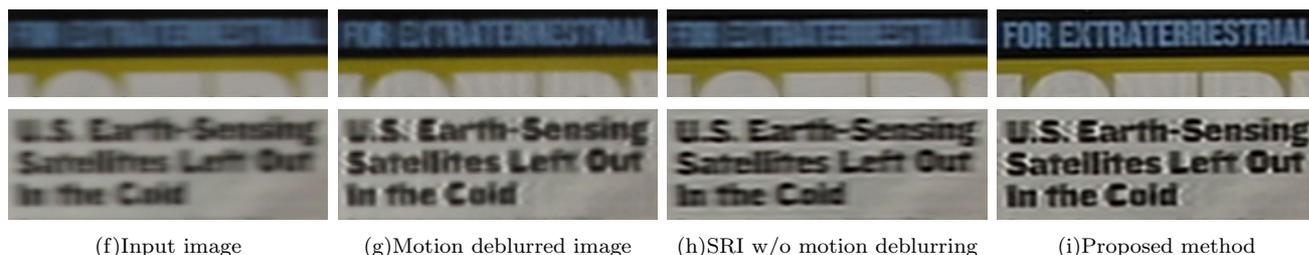
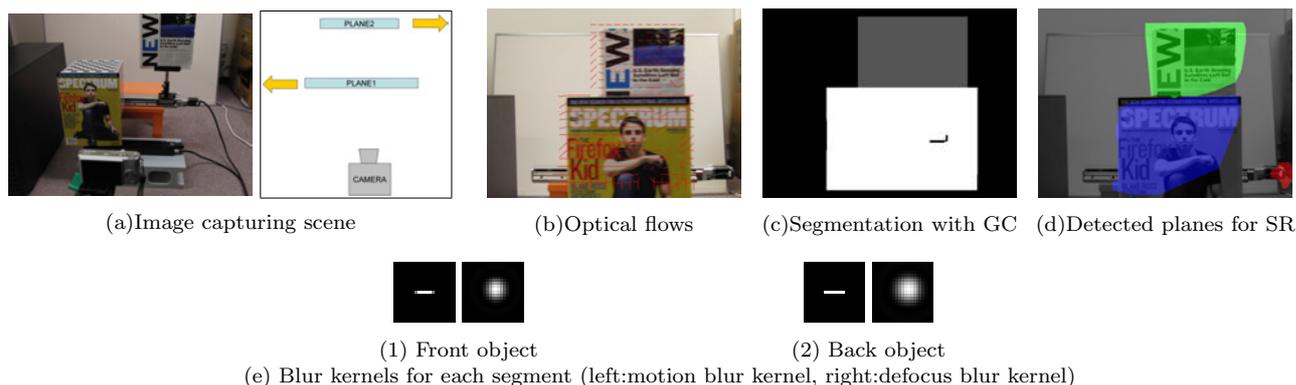


図 5 Multiple object motion by motorized stage.

ションブラー除去後の平面分割結果であり、最終的な超解像処理結果は図 7(h) である。このような手持ちカメラのデータにおいても、うまく領域分割でき、超解像処理が行われていることが分かる。モーションブラー除去をせずに超解像処理を行った結果は図 7(g) であり、提案手法を用いることにより、うまく画像復元できていることが確認できる。

### 6.3 移動する複数物体での実験

固定カメラを用い、複数の移動物体のあるシーンに対して実験を行った。図 8(a) は入力画像に対し、オプティカルフローを計算した結果である。図 8(f) はモーションブラー除去を行った結果であり、モーションブラー除去結果に、超解像処理を行った結果が図 8(h) である。図 8(g) は単純に超解像処理を行った結果であり、提案手法の有用性が確認できる。

## 7. む す び

本論文では、手持ちカメラの、独立した動きやデフォーカスブラーを複数含むようなシーンの画像からシャープで高解像度の画像を復元する方法を提案した。これはまず、フレーム内の移動物体ごとにモーションブラー除去を行い、次にデフォーカスブラーに対して適応的な超解像手法を用いる手法となっている。モーションブラーは、オプティカルフローと、自然画像の情報からカーネル推定を行い、デコンボリューションによりモーションブラー除去を行った。デフォーカスブラーは、MAP 法を用いたピクセル毎の平面分割を用いた超解像手法により除去を行った。実データを用いた実験により、提案手法の有効性を示した。

## 文 献

[1] Ben-Ezra, M., Nayar, S.: Motion Deblurring using Hybrid Imaging. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Volume I. (2003) 657–664  
 [2] Tai, Y.W., Du, H., Brown, M.S., Lin, S.: Image/video deblurring using a hybrid camera. In: CVPR. (2008)

[3] Farsiu, S., Robinson, M.D., Elad, M., Milanfar, P.: Fast and robust multiframe super resolution. IEEE Trans. on Image Processing **13** (2004) 1327–1344  
 [4] Lucy, L.: An iterative technique for the rectification of observed distributions. J. of Astronomy **79** (1974) 745–754  
 [5] Li, F., Yu, J., Chai, J.: A hybrid camera for motion deblurring and depth map super-resolution. In: CVPR. (2008)  
 [6] Sroubek, F., Gabriel, C., Flusser, J.: A unified approach to superresolution and multichannel blind deconvolution. IEEE Trans. on Image Processing **16** (2007) 2322–2332  
 [7] Fergus, R., Singh, B., Hertzmann, A., Roweis, S.T., Freeman, W.T.: Removing camera shake from a single photograph. In: SIGGRAPH. (2006) 787–794  
 [8] Cho, S., Matsushita, Y., Lee, S.: Removing non-uniform motion blur from images. In: ICCV. (2007) 1–8  
 [9] Farsiu, S., Elad, M., Milanfar, P.: A practical approach to superresolution. In: Visual Communications and Image Processing. Volume 6077. (2006)  
 [10] Park, S.C., Park, M.K., Kang, M.G.: Super-resolution image reconstruction: a technical overview. IEEE Signal Processing Magazine **20** (2003) 21–36  
 [11] Yamaguchi, T., Kawasaki, H., Furukawa, R., Nakayama, T.: Super-resolution of multiple moving 3d objects with pixel-based registration. In: ACCV. (2009)  
 [12] Irani, M.I., Peleg, S.: Improving resolution by image registration. CVGIP **53** (1991) 231–239  
 [13] Dick, A., Torr, P., Cipolla, R.: Automatic 3d modelling of architecture. In: BMVC. (2000) 372–381  
 [14] Bartoli, A.: Piecewise planar segmentation for automatic scene modeling. In: CVPR. (2001) 283–289

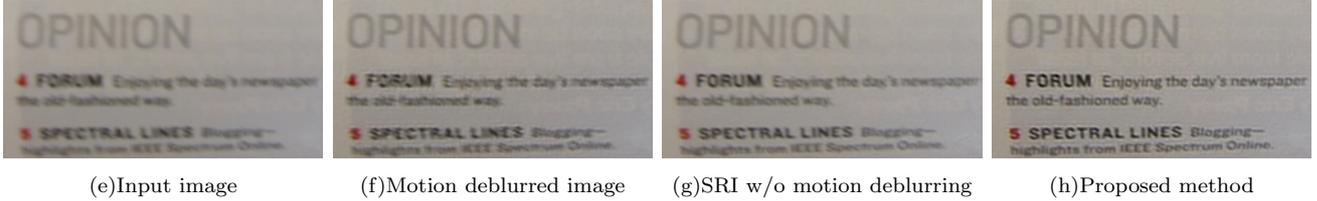
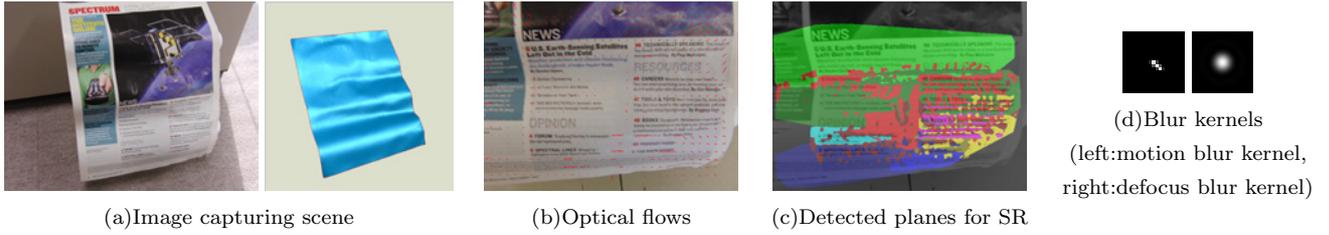


Figure 6: Motion blur and super-resolution for curved surface.

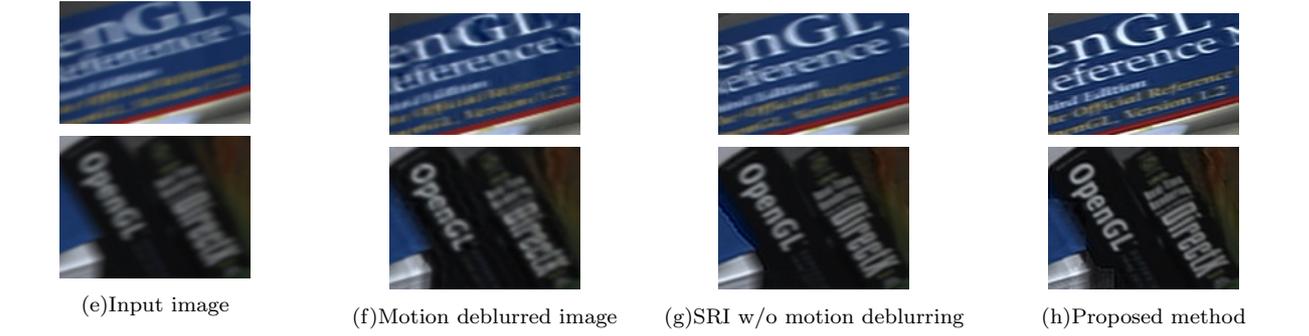
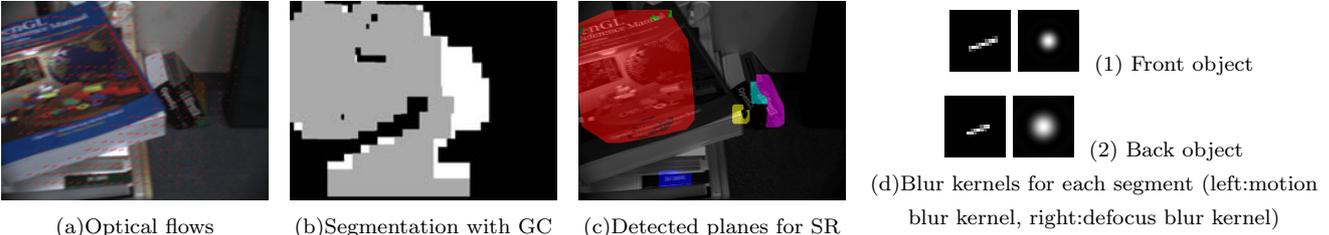


Figure 7: Video data capture by handheld camera.

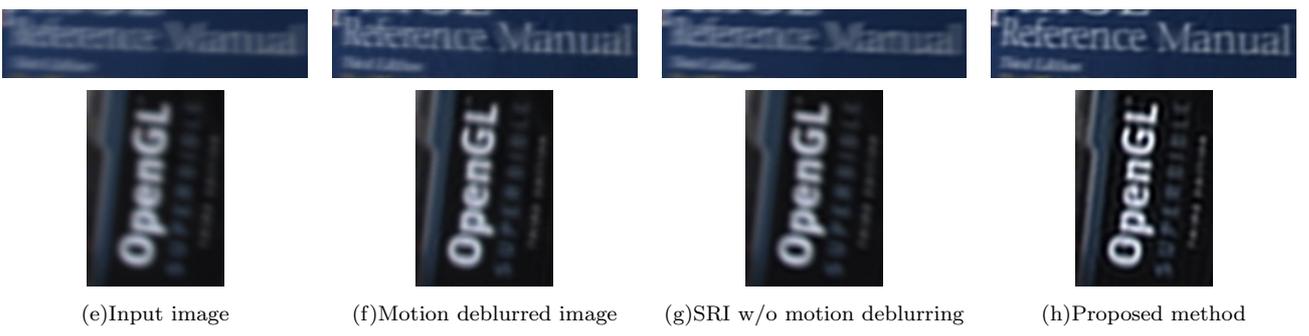
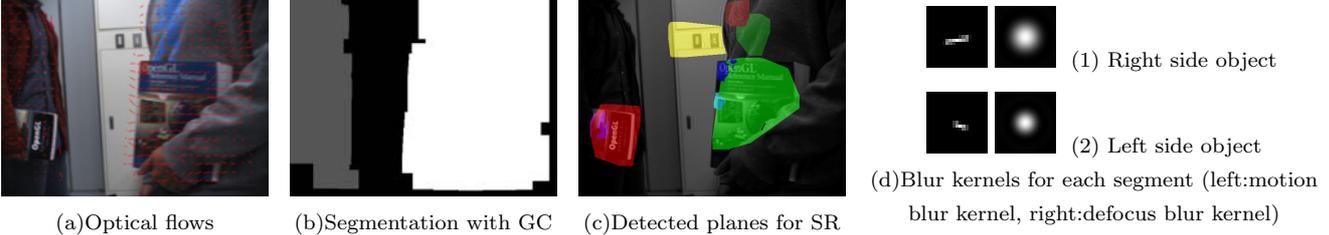


Figure 8: Video data captured by static camera.