

モーションブラーを含むシーンの
ビデオ映像からの
高精度な画像切り出し手法

2018年3月

松下 侑輝

概要

本研究では、撮影シーンに様々な方向のブレが混在している映像を対象とし、高解像度化する方法を提案する。これまで、撮影シーンにブレを含む映像を対象とした超解像が提案されているが、平行移動のみを考慮した領域分割に基づく手法のため回転などの動きに弱かった。そこで、本論文ではピクセルごとにモーションブラーカーネルを導出して、ブレ除去を行う。また、シーンのブレによって特徴点のトラッキングの誤差が発生する問題がある。シーケンスごとのトラッキング結果を統合し、さらに特徴点の位置を高精度化することでブレのあるシーンでのトラッキングを実現した。シーン中に動物体がある場合、動物体と背景のオクリュージョン部分で色の混成が発生しそれが画像生成時にノイズになるが、前景のアルファチャンネルと背景画像を別個求めることでオクリュージョンを考慮したブレ除去を実現した。これらの手法を用いることでブラーカーネルとフレームごとの位置ずれを計算し、超解像やブレ除去の処理を行うことでより精度の高い高解像度画像を生成することができる。実験では、各処理ごとに従来手法との比較を行い、提案手法の有効性を確認した。

目次

第 1 章	はじめに	1
1.1	背景	1
1.2	研究目的	1
1.3	本論文の構成	2
第 2 章	関連研究と技術課題	3
2.1	基礎的な超解像	3
2.2	基礎的なブレ除去	4
2.3	動物体シーンにおけるブレ除去	5
2.4	一様でないブレを対象としたブレ除去	6
2.5	特殊な露光方法によるブレ除去	6
2.5.1	導入	7
2.5.2	シーケンスの相補セットと符号化露出	8
2.5.3	符号化開口ビデオ	12
2.5.4	実験結果	17
2.5.5	まとめ	23
2.6	ピクセルごとにブレを計算したブレ除去法	23
2.6.1	導入	24
2.6.2	一般的な動画のブレ除去	26
2.6.3	最適化フレームワーク	31
2.6.4	実装の詳細	33
2.6.5	実験結果	35
2.6.6	まとめ	38
2.7	関連研究まとめ	39
第 3 章	処理手法の概要	40

第 4 章	ブレのある動画像からのブレ除去と超解像の同時処理	44
4.1	はじめに	44
4.2	関連研究	45
4.3	提案手法	46
4.4	詳しい方法	47
	4.4.1 ピクセルごとのブレ推定	47
	4.4.2 ブレのある画像のトラッキング	47
	4.4.3 マップ推定を用いた画像復元	48
4.5	実験結果	50
4.6	まとめ	53
第 5 章	オクリュージョンを考慮したブレ除去	57
5.1	はじめに	57
5.2	関連研究	58
5.3	提案手法	59
	5.3.1 アルゴリズム概要	59
	5.3.2 前景と背景の分離	60
	5.3.3 アルファマップとテクスチャのブレ除去	62
5.4	実験結果	63
	5.4.1 シミュレーション画像のブレ除去	63
	5.4.2 実測画像のブレ除去	64
5.5	まとめ	65
第 6 章	考察	71
6.1	入力枚数と結果の相関	71
6.2	ブレ除去と超解像の同時処理	72
6.3	前景部分推定と結果の関係	74
第 7 章	おわりに	81
	参考文献	83
	付録	97

第 1 章

はじめに

1.1 背景

監視カメラやビデオカメラの動画から静止画を取り出したいことがある。しかし、動画から単純にフレームを切り出しただけの静止画では、品質が低い場合が多い。これは一般に、動画は静止画に比べて低解像度であることや、撮影されている物体の動きや、カメラのブレなどによるモーションブラーの発生などが要因である。これらの問題を解決するためブレ除去や超解像処理が研究されている。一般に、これらの手法では、フレーム全体のカメラの動きを推定し、位置合わせすることでブレ除去や画像の超解像度化を実現している。しかし、シーン内に異なる奥行き of 物体や、独立した動きがある場合、フレーム内に方向や大きさの異なるモーションブラーが発生し、一つの動きだけを仮定する過去の一般的な手法を適用することが困難になる。さらに、超解像処理についてはモーションブラーがあることで画像の位置合わせに誤差が生じるだけでなく、モーションブラーがノイズとして扱われることで超解像処理にも悪影響を及ぼす。

上記の問題を解決するために、画像を小さなパッチに分割して処理をしたり、シーンを構成する形状を平面ごとに分割し、その分割した領域ごとに処理をする手法が提案されている。しかし、複数枚の連続するフレーム間において、モーションブラーがある状態で、トラッキングを行うことは依然として困難である。また、ブレ除去処理については動物体のモーションブラーがあると動物体そのものと背景の境界部分周辺で色の混成が起これ、生成画像に悪影響を与えている。物体がモーションブラーを起こしているシーンに対応したブレ除去の手法は提案されているがその問題については平滑化項によるノイズ除去によって対応している手法が多い。

1.2 研究目的

そこで本研究では、動物体を含むシーンの同物体のテクスチャを復元する方法を目的として以下のような方法を提案する。

1. ピクセルごとにブレンダーカーネルを推定する手法の提案
2. ブレのある連続画像での高精度トラッキングの実現
3. 複数フレーム画像を入力とする, ブレ除去と超解像処理の同時実現
4. 前景背景のオクリュージョンを考慮した画像処理

これにより, 低解像度でブレのあるビデオを入力として高精度な静止画を作成できる.

1.3 本論文の構成

本論文は, 全 7 章で構成される. 第 1 章は序論とし, 第 7 章において結論を述べる. 第 2 章では従来のブレ除去, 超解像, 技術を述べる. 第 3 章では, 本手法のフレーム間のトラッキング及び画像の高精度化処理の概要を述べる. 第 4 章では, ブラーがある動画画像からブレ除去と超解像を同時に行う方法について述べる. 第 5 章では, 前景物体が動くシーンでオクリュージョンを考慮したブレ除去を行う方法を述べる. 第 6 章では, シミュレーション画像及び実測画像での画像の高精度化処理の結果について考察する.

第 2 章

関連研究と技術課題

2.1 基礎的な超解像

超解像は、単画像のみを使う手法 [1] と複数画像を使う手法 [2] があるが、画質や安定性の点からは、複数画像の方が優れている。複数画像を用いる超解像手法では、各フレームを基準フレームにサブピクセル精度で位置合わせする手法が多い。

一般的な画像復元の論文は観測画像を復元すべき潜在画像に何らかの劣化処理が施されたものであると仮定している。このとき観測画像に劣化処理の逆の処理を施すことで復元できる。特に劣化処理が行列の畳み込みでモデル化される場合は、劣化処理に値する行列の逆行列で観測画像を畳み込むことで潜在画像を取得できることになる。しかし、実際は逆行列の推定が困難である場合や逆行列を畳み込むことによってノイズ等の成分が増幅される場合が多いためデコンボリューション(逆畳み込み処理) やそれに準ずる手法が提案されている。

超解像はサブサンプリング等の処理が劣化処理として仮定されている。超解像の方法は低解像度画像とそれに対応する高解像度画像の辞書を用いる辞書ベースの方法と、複数の連続した低解像度画像の動画を用いて高解像度画像を直接導出する方法の二つに分かれる。

辞書ベースの方法として観測画像とそれを何倍かに拡大した画像群から類似するパッチを検出することで対応関係を得る自己相似性ベースの方法 [3, 4] や、あらかじめ低解像度のパッチと高解像度のパッチを学習させる方法 [5, 6, 7] が提案されている。これらの手法は処理結果が実際の潜在画像か保証されていない問題点がある。

低解像度の動画から高解像度画像を導出する方法は、動画の各フレーム間の位置合わせを行い、ピクセルの対応関係をサブピクセル精度で計算することで行われる [2]。この時、シーンが一つの平面で仮定される場合は位置合わせを一つのフレームのペアにつき一つの射影変換行列で表現するこ

とができるが、シーン中に三次元形状や布などの非剛体の物体がある場合や複数の動く物体がある場合に適用することは難しい。そのようなシーンではシーンの形状を推定し、形状ごとに三角形のパッチを設けパッチごとに処理を行う手法が提案されている [8, 9]。しかし、シーンによってパッチを正確に与えることが困難という問題がある。また、動画中に動物体がある場合は物体の動きによってブレが発生することが多い。超解像において物体の動きブレはノイズとして処理され、超解像の結果を劣化される要因となる。さらに、位置合わせの際に行う画像の特徴点トラッキングが困難になり超解像そのものに行えない可能性も高まる。

2.2 基礎的なブレ除去

画像のモーションブラーは、ブラーのない画像に、ブラーの原因となる動きやボケを表すカーネルを畳み込むことでモデル化される。従って、ブレ除去は、一般に、このカーネルを推定し、LR 法 [10, 11] や IBP 法 [12] などの逆畳み込み処理によって実現される。

ブレ除去はブラーカーネルと呼ばれるピクセルごとのブレを表す行列の畳み込みが劣化処理として仮定されている。伝統的なブレ除去の処理は、画像全体が同じカーネルでブレているという前提のもとで、単一のブラーカーネルを別途推定して、逆畳み込み処理を行う方法 [10, 11, 12] であった。このような単一ブラーカーネルの前提は、対象の奥行変化が少ないシーンで、カメラの手ブレが主なブレが原因の場合にはよく合致した。しかし、ブラーカーネルを別途推定することは事実上難しく実用性に乏しかった。

そこでここ十年近く、ブラーカーネルを観測画像のみから推定する手法 (ブラインドデコンゴリューション) が盛んに行われてきており、大きな進展があった。Shan はベイズ推定を用いた観測画像生成のモデルを用い、MAP 推定によって潜在画像とブラーカーネルを交互に更新する最適化を行った [13]。この時カーネルはシーン全体をブレ除去できるように推定されるため、ある程度奥行きに差があるようなシーンにも対応できる。Hirschi は一般的な平行移動を用いた畳み込みではなくホモグラフィ変換を用いた畳み込みを利用して画像中のブラーが一様でないブレ除去を実現した [14]。この方法によって光軸を軸とした回転等のカメラの動きにも対応できるようになった。また、Li らはシーンを分割してセグメントごとにカーネルを求めブレを除去する方法を提案した [15]。この方法によってシーン中に大きく奥行きが異なるオブジェクトがある場合に対応できるようになった。また、セグメンテーションベースの方法はシーン中にオブジェクトが独立して動く場合にも対応できる [16, 17]。

このようにブレ除去の方法はカメラが平行移動する場合から回転運動する場合へ、シーンが一つの平面と仮定できる場合から複数の奥行がある場合や動物体を含む場合へ、より複雑で実用的な場合に対応できるように進歩した。

一方で、シーン中の文字の領域が基本的に二色からなると仮定して局所領域ごとに2色の画像を出力させるような特定のシーンに特化したプライヤーを設ける手法 [18] や、カメラにジャイロセンサを取り付けてより正確にブラーカーネルを推定する手法 [19] など、ある特別なシーンやカメラを対象にしたブレ除去の手法も提案されるようになった。また局所領域ごとのブラーをエッジの角度の辞書の重みによって推定するスパースコーディングライクな方法や [20], CNN を用いてパッチごとのブラーの角度と長さを推定する方法 [21] など、ブレ除去の方法は多岐にわたる。以上の手法は、ブレ除去の目的が、カメラで撮影したボケ画像の修復にあったため、基本的には単画像を対象としたものであった。

さらに、ここ数年の動きとして、複数枚を入力とする手法が多く提案されるようになった。複数枚を入力する超解像と同様に複数枚使用するブレ除去も複数の観測画像から一枚の潜在画像を計算するものである [22]。同じシーンを異なるブラーカーネルによってブレを与えた複数の画像を入力することで一枚のみを使用してブレ除去処理を行うより結果を向上させることができる。また、ブラーがある画像とノイズがある画像、つまり露光時間とゲインを変化させた画像の対からブラーとノイズを共に低減させた画像を生成させる手法 [23] も提案された。

2.3 動物体シーンにおけるブレ除去

動物体シーンがある動画のブレ除去については、単画像の場合と比べて連続した複数のフレームがある点でセグメンテーションの計算が有利である。例えば、固定カメラで背景が静的である場合にはフレーム間の差分からシーンの動物体部分と背景部分の推測をある程度行うことができる。Armit の手法 [24] はその条件に加えて物体が等速直線運動していると仮定して動物体シーンのブレ除去を行っている。このときフレームごとに露光時間が変化する撮影方法を用いることでフレーム間のブラーカーネルを変化し、よりブレ除去の精度が向上している。この手法の応用として各フレームを複数のシャッターの開閉パターンによって取得する方法がある [25]。また、シーケンス中のブラーの影響度合いを判別し、パッチごとに最もブレが少ないフレームの画像を他フレームの対応部分に当てはめることでブラーの少ない動画を作る方法 [26] も提案されている。

近年、モーションブラーが存在する通常の動画から直接計算によってブレ除去を行う手法が提案されるようになった。Jonas はシーンを前景と背景に分けたモデルを考え、潜在前景画像・潜在背景画像・前景と背景の位置合わせ情報・セグメンテーション情報を順番に交互に最適化することでブレ除去を行った [27]。Kyoung はオプティカルフローの空間的滑らかさと空間的滑らかさを考慮した最適化関数を設け、ブレ除去とシーンの動き推定を交互に行うことでシーンのブレ除去を行った [28]。両者の手法はシーンのオプティカルフローを用いて動物体部分の識別を行うがオクルージョン部分周辺の動きの推定が困難であるためブレ除去の結果もオクルージョン部分周辺で動物体

の欠損やリングングなどのアーティファクトが発生する問題がある。

2.4 一様でないブレを対象としたブレ除去

複数の入力画像を用いた超解像, ブレ除去いずれにおいても, フレーム間のピクセルの対応付けが必須であり, そのために特徴点をトラッキングする必要がある. 代表的な特徴点抽出法 [29, 30, 31] やトラッキング手法 [32] では, 画像の勾配を計算に利用していることが多く, モーションブラーによって高周波成分が消失する場合や, フレーム間で周波数スペクトルが変化する場合に精度の低下や, 特徴点の検出や追跡に失敗する. 回転などのカメラの手振れに対応するために, ブロックごとにカーネルを推定する手法もある [33]. 具体的には開口中のカメラの動きを図 2.1 左のようにカメラを原点とした三方向の回転の集まりと考え, カメラの回転に応じて画像を変換したものの線形和をブレのモデルとした. このとき, 画像変換をホモグラフィ行列とすることで, 図 2.1 右のように領域ごとにブラーカーネルを変更してブレの除去を行った.

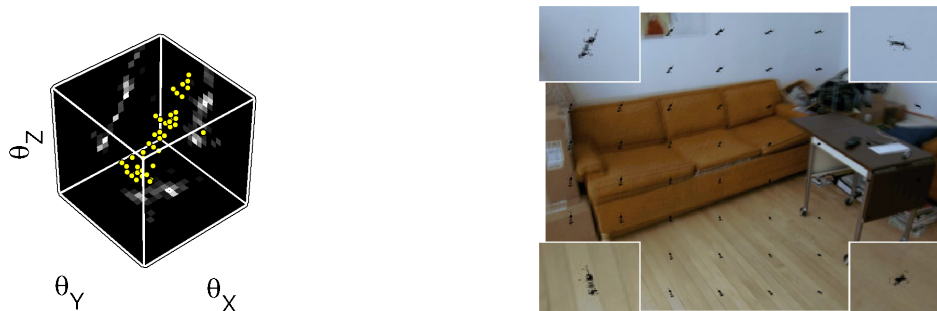


図 2.1 領域ごとに一様でないブレの推定の様子

2.5 特殊な露光方法によるブレ除去

本項目では, 符号化露光技術を用いた複数画像を用いた動きブレ除去法を述べる. ここでの研究の重要なアイデアは, 空間周波数を詳細に保持するために, 一連の相補的なシャッター開閉パターンでビデオフレームをキャプチャしたことである. 現代の通信理論に基づいて 2 進シーケンスの相補的なセットを生成するためのアルゴリズムを導入し, 既製のマシンビジョンカメラを用いて符号化された露光ビデオシステムを実装した. ここでの方法の有効性は, 画像のブラー除去のために使用される他のコンピューショナルフォトグラフィー手法との定量的および定性的な比較を用いて, 様々な困難な例で実証される.

2.5.1 導入

画像ぼけは、撮影中に高周波情報が失われるために本質的に不良条件問題となる困難な課題である。過去 10 年間に、リチャードソン・ルーシー (Richardson-Lucy) [10, 11] やウィナー・フィルタ (Wiener filter) [34] などの従来のぼけ補正手法よりも性能を向上させる画像ぼけ補正研究の重要な進展があった。

ブレ除去処理に複数のブレがある画像を使用したことは関心を集めた研究方向の一つである。補完情報が提供されるため、単一の画像を使用したブレ除去方法よりも優れた性能を示す。Yuan らはブラーカーネルを推定するためにノイズは無いがブレがある画像とブレは無いがノイズがある画像のペアを使用した [35]、Cai らは複数の重度にモーションブラーが生じた画像を用いた手法を提案している [36]、また、Chen らは反復ブラーカーネル推定とブレ画像を二つ用いたブレ除去を実行した [37]。Cho ら [26] は、フレーム内の鮮明な領域を使用して、近くのフレーム内の同じコンテンツのブレがある領域を復元した動画のブレ除去アプローチを提示した。Shahar ら [38] や、Shechtman ら [39] は、時間超解像を用いて動画のフレームレートを増加させることで、ビデオのモーションブラーを低減した。

画像ブレ除去研究のもう一つの発展は、ブレ除去問題をより実現可能にするために、画像を撮影した方法を変更したことである。例えば、各画素を 2 値パターンで独立に変調したように映像を取り込むことで、高い時間分解能の復元が可能となる [40, 41, 42]。

一方、この研究は、特に画像キャプチャ中に、ピクセルごとではなく画像全体の露出を制御したイメージングシステムを使用した作品に関連している。Raskar ら [43] は、空間周波数の詳細を保存するために、露光時間内に特別な仕方でもカメラのシャッターを開閉して符号化された露出撮影を行い、デコンボリューションが良設定問題になるようにした (図 2.2 (a))。Jeon らは [44] シャッター開閉パターンを最適化したことによってデコンボリューション性能を改善した。

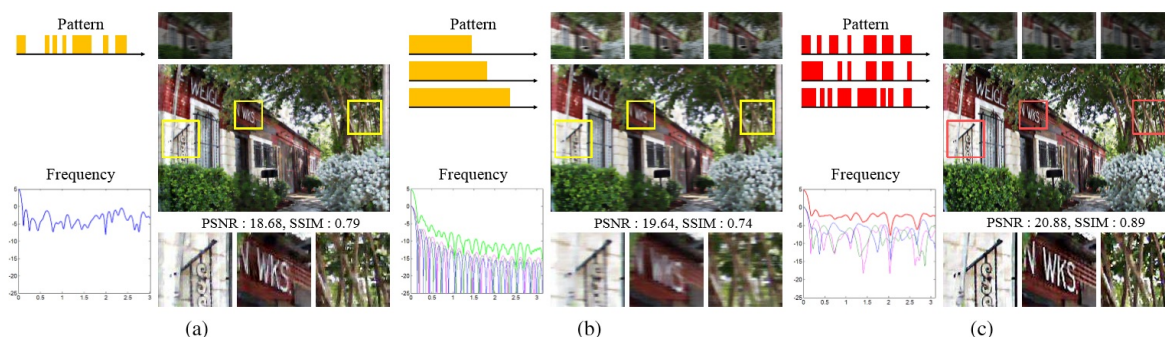


図 2.2 (a) 符号化開口によるブレ除去 (b) 複数露光ビデオによるブレ除去 (c) 複数の符号化開口ビデオによるブレ除去

Agrawal ら [45] は, 1 回の露光でシャッターを開閉させるのではなく, 連続したフレームの露光時間を変化させて様々な露出ビデオフレームワークを提案した (図 2.2 (b)). 主なアイデアは, 様々な PSF (点広がり関数) で同じオブジェクトを撮影して, あるフレームで NULL となる周波数成分を他のフレームで埋めることができたように露出を変化させることで, ブレ除去を良設定問題にしたことである. Holloway ら [46] は, 符号化絞りの概念をビデオカメラに適用した. このアプローチは, 連続したフレームにおいて異なるシャッター開閉パターンを有した一連の符号化された露光画像を撮影し, 圧縮センシングを介して時間分解能のアップサンプリングを実行した. このアプローチは時空間連続性を持たないシーンに対応することができず, オブジェクトの動きを推定するために多くの枚数を必要とした.

この手法では, コード化露光イメージング [44, 43] と様々な露出ビデオ [45] の利点を組み合わせたコード化露出ビデオ方式を提案した. フレーム中の露出を変えるのではなく, 一定の露出時間でビデオをキャプチャし, 各フレームにコード化された露出を適用した (図 2.2 (c)). 画像取り込み中の周波数情報損失を最小にするために, 連続したフレームの損失を補償した補完的な一組の開閉パターンを生成した方法を紹介した. 補完的な開閉パターンを使用したことにより, 柔軟なフレームレートで撮影できた多種類露光時間シーケンスを生成した. 以前の方法と比較して改善された SN 比で高品質のブレ除去結果を達成したことが可能になる. さらに, ここでのフレームワークが, 意図的なブラーを追加したことによって Li らの手法のようなビデオ監視のためのプライバシー保護などの他のアプリケーションに適用できたことを示す [47].

2.5.2 シーケンスの相補セットと符号化露出

この手法の重要なアイデアは, 撮影された画像が空間周波数を保持したように, 各フレームの周波数損失を補償した一連の開閉パターンで動画のフレームを撮影したことである. このような開閉パターンを生成するために, 多入力多出力 (MIMO) レーダや符号分割多元接続 (CDMA) 技術などの多くの工学アプリケーションで広く使用されている二進シーケンス [48, 49] を紹介した. このサブセクションでは, 符号化された露出映像と, 変化した露出映像 [45] との相補的な一組の開閉パターンを用いた符号化された露出映像の利点を理論的に示す.

符号化開口と複数露出の比較

Raskar らの手法 [43] では, 平坦なスペクトルを有したシャッター開閉パターンは, 符号化された露光イメージングにおけるブレ除去の品質を改善することが示されている. スペクトルの平坦さを測定するには, 開閉パターンの自己共分散関数の合計を使用した. また, バイナリ系列の自己相関関数を自己共分散関数で近似できることも Jeon らによって示されている [44]. 長さ n の二

進シーケンス $U = [u_1, \dots, u_n]$ を用いて, 自己相関とシーケンスのフーリエ変換による変調伝達関数 (MTF : 二進シーケンスの周波数応答の大きさ) との間の関係が導出された.

$$\sum_{k=1}^n \Psi_k^2 = \frac{1}{\pi} \int_0^1 \left[|F(U)|^2 - n \right]^2 d\theta \quad (2.1)$$

ここで, $F(U)$ はシーケンス U のフーリエ変換を表す. k は以下のように示すシーケンスの自己相関関数の k 番目の要素を示す,

$$\Psi_k = \sum_{j=1}^{n-k} u_j u_{j+k} \quad (2.2)$$

Jeon ら [44] は, より小さな値の Eq. (2.1) の値が小さいほどメリット係数が高いことを示していることを示した. Ukil らは [50], Eq. (2.1) の最小値が $n/2$ によって制限されることを証明している.

この符号化露光ビデオのフレームワークでは, 相補的なセットは, セット内のシーケンスの自己相関関数の合計がゼロである二進シーケンスのセットとして定義される. 長さ n の p (≥ 2) 個の列 U_1, \dots, U_p からなる相補的な集合 Δ を持つ場合, その関係は以下のように示される.

$$\sum_{i=1}^p \Psi_k^i = 0 \quad s.t. k \neq 0 \quad (2.3)$$

ここで, Ψ_k^i は, Δ 内の i 番目のシーケンスの自己相関関数の k 番目の要素を表す.

Naghsh ら [51] は, 相補的な集合 Δ が以下の式の最小化によって計算されることが示している.

$$\sum_{k=1}^{n-1} \left| \sum_{i=1}^p \Psi_k^i \right|^2 = \frac{1}{2} \int_0^1 \left[\sum_{i=1}^p |F(U)|^2 - pn \right]^2 d\theta \quad (2.4)$$

最適な場合には, 式 (2.4) の最小値は, 式 (2.3) から 0 になる. これは, 上記のように, 式 (2.1) が単一の二進シーケンスに対して $n/2$ に限定されているので, 相補的なセットの結合スペクトルが単一の二進シーケンス (符号化された露光イメージング) のスペクトルよりも平坦なスペクトルを持つことを意味する. ここでは, Tendero ら [52] によって提案された数値的尺度を用いてシーケンスの 1 つの補集合の一例のスペクトル利得を推定した. シーケンス集合のゲインは 0.85 であり, 単一符号化パターンの最適なスナップショットのゲインは 0.56 である.

物体の速度の不変性

ショット中に物体が n 個の画素の範囲を移動した場合, 開閉パターン $U = [u_1, \dots, u_n]$ の最適な長さは n である. McCloskey [53] が示すように, 物体が 2 倍速く動くと, 有効な PSF は $\frac{1}{2n} [u_1, u_1, \dots, u_n, u_n]$ に伸び, PSF の可逆性は保証されない.

シーケンスの補数セットは有効な PSF だが物体の速度によってスーパーサンプルされたり伸張された場合であってもスペクトル情報の損失を最小化したことを示す。一例として、伸ばされたシーケンスに起因した MTF の変化を係数 2 で導出した。

$$\begin{aligned}
b_i &= \sum_{j=0}^{2n-i-1} U_{2n} U_{2n}(j+i) \\
&= \sum_{p=0}^{n-q-1} [U_{2n}(2p)U_{2n}(2p+2q) + U_{2n}(2p+1)U_{2n}(2p+2q+1)] \\
&= \sum_{p=0}^{n-q-1} [U_n(p)U_n(p+q) + U_n(p)U_n(p+q)] = 2\Psi_k \\
&\quad \text{whare } j = 2p \text{ and } i = 2q
\end{aligned} \tag{2.5}$$

延伸された PSF の MTF の分散は、以下のような PSF よりも一定倍大きい。

$$\sum_{i=1}^{n-1} b_i = 4 \sum_{k=1}^{n-1} \Psi_k^2 = 2 \int_0^1 [|F(U_n)|^2 - n]^2 d\theta \tag{2.6}$$

相補的な配列の組における $\frac{1}{2} \int_0^1 [F(U_n) - n] d\theta$ の最適な境界は、理論的にはゼロである。したがって、削減された PSF の最適な境界もゼロになる。実際には、ここでの一連のシーケンスは最適な境界に近いので、提案された相補的なセットは速度依存性の問題を扱うことができ、どのような要因の場合でも示すことができた。オブジェクトの速度に対するここでのシステムのロバスト性は、実験セクションでも実証されている。

複数露出ビデオとコード化露出ビデオの比較

ここでのコード化された露出ビデオフレームワークを様々な複数のビデオ手法と比較するために、ここではこれら 2 つの方法について MTF の上限を分析した。比較の主な基準は、MTF の分散、最小、および平均である。符号化された露光撮影における二進シーケンスの MTF は、画像ブレ除去の性能に直接的な影響を及ぼす [43, 54]。MTF の分散と平均はデコンボリューションノイズに関連しており、スペクトルのピークは、ブレがある画像の空間周波数を破壊するため、ブレ除去に悪影響を及ぼす。

複数の露光手法 [45] の MTF は、時間領域における長方形関数の異なる持続時間の結合スペクトルである (図 2.3 (b))。1 から p まで露出が変動した場合 MTF の上限 X_{vary} は以下で示される、

$$|X_{vary}(\omega)| = \sum_{i=1}^p |F(\Pi(l_i))| = \sum_{i=1}^p \left| \frac{\sin \frac{l_i}{2} \omega}{\sin \frac{\omega}{2}} \right| \leq \left| \frac{p}{\sin \frac{\omega}{2}} \right| \tag{2.7}$$

ここで、 $\Pi(l)$ は長さ l の矩形関数を表し、 ω は $[-\pi, \pi]$ での円形の周波数を示す。

符号化された露出ビデオの上限を計算するために、 Φ を $+1$'s と -1 's の 2 進シーケンスと定義した。Parker ら [55] は、長さ n の p 個の系列 $\Phi_i^i = [1, \dots, p]$ の補集合のフーリエ変換成分の総和が最大 \sqrt{pn} であることを示した。コード化された露光イメージングにおける開閉パターンはその物理的性質のために $+1$'s と 0 's で作られるので、相補的な組の結合 MTF X_{comp} の上限は以下のようなになる

$$\begin{aligned}
 |X_{comp}(\omega)| &= \frac{1}{2} \sum_{i=1}^p |F(\Phi_i + \Pi(n))| \\
 &\leq \frac{1}{2} \sum_{i=1}^p |F(\Phi_i)| + \frac{1}{2} \sum_{i=1}^p |F(\Pi(n))| \leq \frac{1}{2} \sqrt{pn} + \left| \frac{p}{2 \sin \frac{\omega}{2}} \right|.
 \end{aligned} \tag{2.8}$$

図 2.3 は、変動した露出ビデオと符号化された露出ビデオのジョイント MTF を、式 (2.7) および式 (2.8) の理論上の上限と比較した。

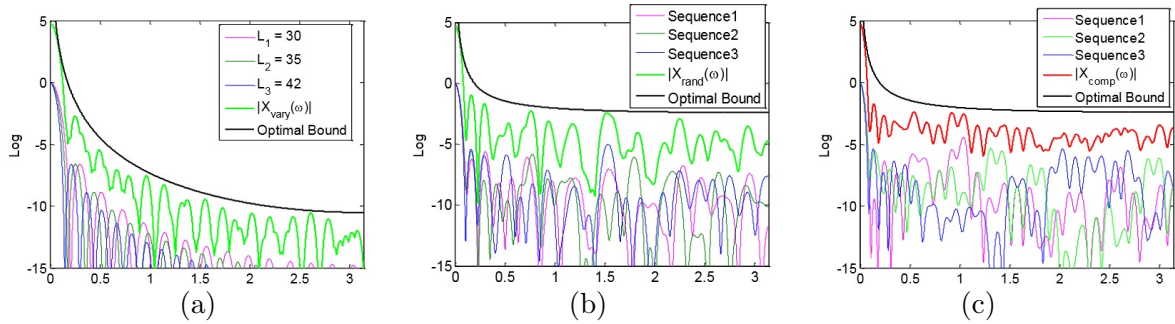


図 2.3 各メソッドの結合 MTF とその理論上の上限 (a) 符号化開口によるブレ除去 (b) 複数露光ビデオによるブレ除去 (c) 複数の符号化開口ビデオによるブレ除去

図 2.3 (a) における複数露出ビデオに対して、露出時間は Wiener[34] や Spasojevic ら [49] の手法で求められ、Agrawal ら [45] の手法で処理される。符号化された露出ビデオの場合、図 2.3 (b) の Raskar ら [43] の手法にある 3 つのランダムシーケンスと図 2.3 (c) の 2 進シーケンスの補集合を使用した。

図 2.3 からわかるように、相補的な組の各コード化されたパターンの MTF にはヌル周波数は観測されない。これは、従来のコード化された露光イメージングのように各単一フレームが可逆的になることを意味する [43, 44]。図 2.3 (b) では、ランダム二進シーケンスのピークスペクトルは緩和されるが、結合 MTF の分散は依然として大きい。相補セット内の各シーケンスのピークスペクトルは、図 2.3(c) の結合 MTF によって十分に補償され、変化した露出方法 (a) および結合セットの両方の結合 MTF よりもはるかに平坦で MTF が高いランダムサンプルシーケンス (b) を生成する。複数の露出ビデオおよび符号化された露出ビデオの MTF 特性の理論的な上限および実際の性能測定値を図 2.4 にプロットする。補完的な開閉パターンの有効性を検証するために、Raskar ら

[43] のランダムサンプル探索法を用いて、単一画像および 3 つの画像の場合の二進シーケンスを生成した。コード化された露出と変動した露出の両方の MTF は、下限（分散）や上限（平均および最小）に達しないが符号化された露光パターンは、変化した露光方法およびランダムシーケンスのセットよりも良好な MTF 特性を示す。具体的には、相補的なセットは、変化した露光方法の理論的境界よりもさらに良好な、より高い平均値および最小の MTF 値を持つ平坦なスペクトルを有した。これは、相補的なセットが、各フレームの周波数損失を補償したことによって、空間周波数を良好に保つことを示している。相補的なセット内のすべてのシーケンスを利用したことは理論的には理想的であるが、図 2.3 および図 2.4 に示すように、相補的なセットの部分的なセットを利用したことも有効であることは注目に値する。

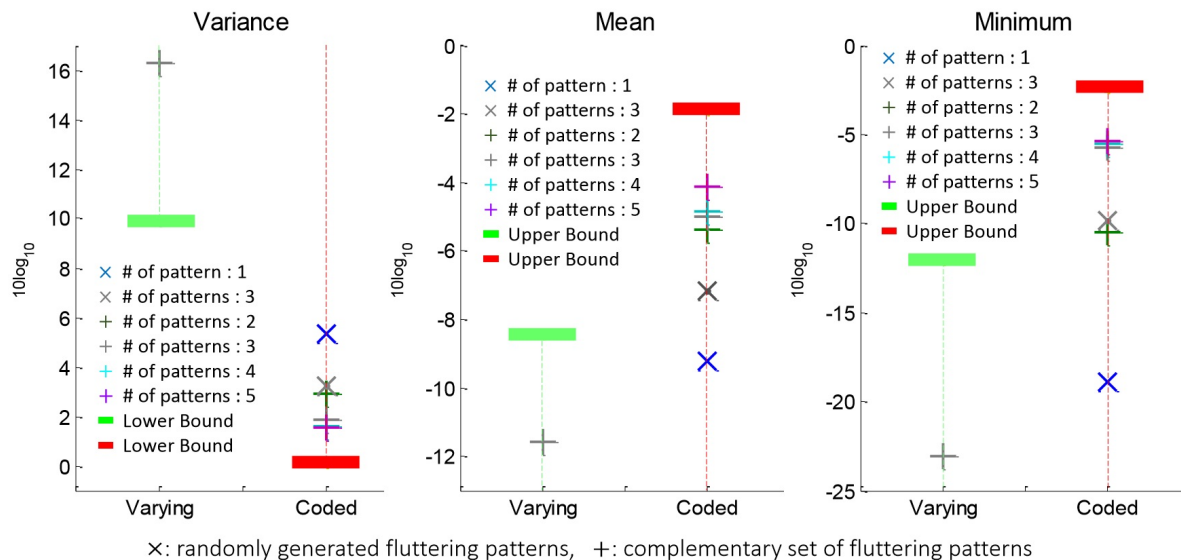


図 2.4 MTF 特性の観点からの複数露光ビデオと符号化開口ビデオの理論上の上限および実際の性能測定値

2.5.3 符号化開口ビデオ

ハードウェアのセットアップ

コード化された露出ビデオのためのハードウェアシステムを構築したことは自明ではない。ここでは多重露光パルス幅モード（トリガモード 5）と外部トリガパルスを生成するための ATmega128 マイクロコントローラ（図 2.5）をサポートする Point Grey Flea3 GigE カメラを使用して、コード化された露出ビデオシステムを実装した。マイクロコントローラは、シリアル通信を介して二進シーケンス信号を外部トリガとしてカメラに送信した（図 2.5 のステップ 1）。シーケンス内の c 個のピーク（図中の灰色の領域）の後にカメラが写真の撮影を終了し、シーケンスの終わりを示す（図 2.5 のステップ 2）。次にカメラは画像と信号をコンピュータに送信し（図 2.5 のステップ 3）、

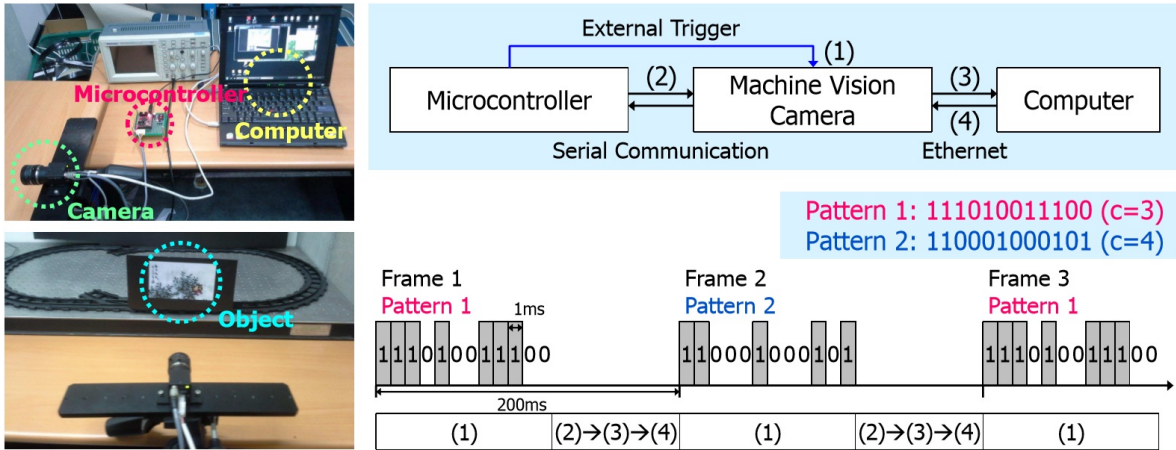


図 2.5 符号化開口ビデオのハードウェア設計

コンピュータは次のシーケンスのパラメータ c をカメラに渡す (図 2.5 のステップ 4) . このプロセスは, 新たなシーケンス (図 2.5 のステップ 1) の下で新たな画像セットを取り出すために再び繰り返される. このシステムを使用して, 複数露出ビデオと符号化された露出ビデオの両方をキャプチャした. 各シャッターは 1ms の長さで, フレームレートはハードウェアの制限によりシーケンスの長さに関係なく 5 フレーム/秒に固定されている. 実装マニュアルとフレームワークのソースコードは, Web サイト (https://sites.google.com/site/hgjeoncv/complementary_sets) で公開されている.

シーケンス生成

符号化された露光イメージングでは, カメラシャッターの開閉パターンは, 一般に 20 ビットよりも長いシーケンス, または 100 ビットよりも長いシーケンスからなる. 開閉パターンの長さは, 照明条件または物体の動きによって変化しうるので, パターンの長さに柔軟性を有することは有益である. ここでは, 柔軟な長さの羽ばたきパターンの補集合を生成した方法を紹介する.

シーケンス長の柔軟性を得るためのここでの戦略は, 相補的集合として知られている小さなサイズの初期集合を拡張することによって相補集合を生成することである. 相補的集合構築の研究は長い歴史を有するので, 以下のような多くの既知の相補的集合が存在する.

$$\Delta = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix} \quad (2.9)$$

ここで, Δ は行列形式の補集合を示す. Δ では, 各行ベクトルは 1 つのシーケンスを表し, すべての

行ベクトルのセットは補集合である.

長さ n の p 個のシーケンスからなる初期補集合 $\Delta_{(p,n)}$ から, より大きな補集合を反復生成したことができた [48]. 相補的な集合 Δ を用いて, より長い長さのシーケンスを有する新しい補集合 Δ^1 が以下の式によって得られる

$$\Delta^1 = \begin{bmatrix} \Delta & \Delta & \bar{\Delta} & \Delta \\ \bar{\Delta} & \Delta & \Delta & \Delta \end{bmatrix} \quad (2.10)$$

ここで, $\bar{\Delta}$ は, Δ のすべての要素 δ_s がフリップされた行列を示す. 拡張 t 回の適用後, 長さ $4^t n$ の $2^t p$ シーケンスを含む相補的行列 $\Delta^t \in \mathbb{R}^{2^t p \times 4^t n}$ を得る.

可変長シーケンスを生成した別の選択肢は, Δ を以下の式と同じ長さの 2 つの行列に分割したことである

$$\Delta^1 = [\Delta_L \Delta_R] \quad (2.11)$$

この場合, 行列 Δ_L と Δ_R は共に補集合となる [56].

式 (2.10) と式 (2.11) の 2 つの行列演算により, $2^t p \times 2^{2t} n$ または $2^t p \times 2^{2t-1} n$ の補集合を生成したことができた. さまざまなサイズがよく知られた最初の相補的なセットがたくさんあるので, 2 つの方法を使用して, シーケンス長の巨大な柔軟性を持つ相補的なセットを生成することが可能になる.

動画のブレ除去では, 必要なシーケンス (または画像) の数は通常 2~5 に制限される. それは周波数ロスを補償するのに十分であり, 多くの画像を撮ると, アライメントや視野の問題などの追加の問題が生じる可能性があるからである. したがって, まず必要な配列長に適合した相補的なセットを生成し, 次にそのセット内の多数の候補配列の中から必要な配列数を選択する.

利用可能な一連の配列から配列を選択した基準に関しては, 開口フレームの数を考慮した. 一般に, 生成された配列は, 同じ数の $n/2$ 等の開口フレームを有したが, 特に短い配列の場合には若干異なる可能性がある. この場合, 同じ数の開口フレームでシーケンスを選択したことは, フレーム間のちらつきを避けるための重要な基準になる.

ブレのあるオブジェクトの抽出とブレの除去

コード化された露光イメージングの 1 つの実践的な問題は, 動く物体のブレ除去のための正確なマット画像を抽出したことである. ブレのプロファイルは, 露光の開閉のために局所的には滑らかにならないため, 困難である. Agrawal と Xu [57] は, 遷移を最小限に抑え, 連続開口フレームを最大化した開閉パターン設計ルールを提案し, PSF 推定 [58] と可逆性の両方の基準を達成できたことを示した. Tai ら [59] は, モーションパスとマグニチュードを推定するために, ユーザーのストロークで静止した背景からブレのあるオブジェクトを抽出した. McCloskey ら [54] は, 画像がモーショ

ンブラーを含むオブジェクトのみがあると仮定して、符号化された露出のための PSF 推定アルゴリズムを提示した。この手法では、PSF、オブジェクトマッティング、複数画像ブレ除去を共同で評価することで、このマッチング問題を解決した。ここでは、静止したカメラから画像を取得し、一定の速度の 1 次元運動によって動く物体にブラーが発生したと仮定した。

■初期化 ブレのある物体を正確に抽出するために (図 2.6 (a)) , ここではまず背景画像を捕捉し、ガウス混合モデル (GMM) を用いて背景の各画素をモデル化した。物体が背景を通過したとき、各画像の画素と背景モデルとの間のマハラノビス距離を計算することによって、最初的前景層を推定した。推定されたレイヤーには多少のノイズがある、モルフォロジー演算を適用して三色マップを作成してレイヤーを整形した。ここでは三色マップを用いて (図 2.6 (b)) , 閉形式マッティング [60] を介して各画像でブレのあるオブジェクトを抽出した。物体マッティングの後、一定速度、一定加速度、および調和運動の場合に扱うことができた McCloskey ら [54] の方法に基づいて、各画像の PSF を推定した。

具体的には、まずラドン変換を行い、最大分散を持つ方向をブレの方向として選択した。これは、ブレのある画像の高い空間周波数がブレの方向に応じて崩壊するためである。次に、ブレ画像のパワースペクトルの密度と様々なブレのサイズに対する開閉パターンの MTF との間のマッチングスコアを計算した。最高のマッチングスコアを選択して、ブレのサイズを決定した。この方法では、各符号化ブレ画像のブラーカーネルを独立に推定した (図 2.6 (c)) 。これは、実際に頻繁に起こるフレーム間の一定の動きの仮定の侵害の影響を受けていないため有用である。

捕捉された画像は推定された PSF を用いて独立してブレ除去処理される。次に、すべての画像を SIFT 特徴点マッティング [29] を用いてアフィン行列で整列させ、すべての撮影画像をアルファマップとともに統合した (図 2.6 (d)) 。

■イテレーションごとの更新 初期化の後、ブレ除去画像とセグメンテーションのマスクを繰り返して最適化した。PSF との統合された画像に基づいて、以下のエネルギー項を最小化することによって、ノンブラインドの複数画像ブレ除去を実行した。

$$\arg \min_Y \sum_{j=i}^m \|B_j - K_j Y\|^2 + \lambda_d \|\nabla Y\|^p \quad (2.12)$$

Y は潜在的なぼけのない画像、 ∇Y は潜像の勾配、 B_j は PSF 行列の集合 K_j によって捕捉された線状のブラー画像の集合である。 λ_d は平滑度の重みであり、 m は画像の数である。画像のブレ除去のために $\rho = 0.8$ を設定した [61] Tai ら [59] の手法に従って、マージされたアルファマップのブレ $\bar{\alpha}$ に対して $\rho = 0.5$ を設定した。ブレ除去されたアルファマップ $\bar{\alpha}$ に再びブレを与えてガイダンスアルファマップ $\hat{\alpha}$ を得る。これは、移動物体のアルファマップ α を精緻化するために、近似マップに

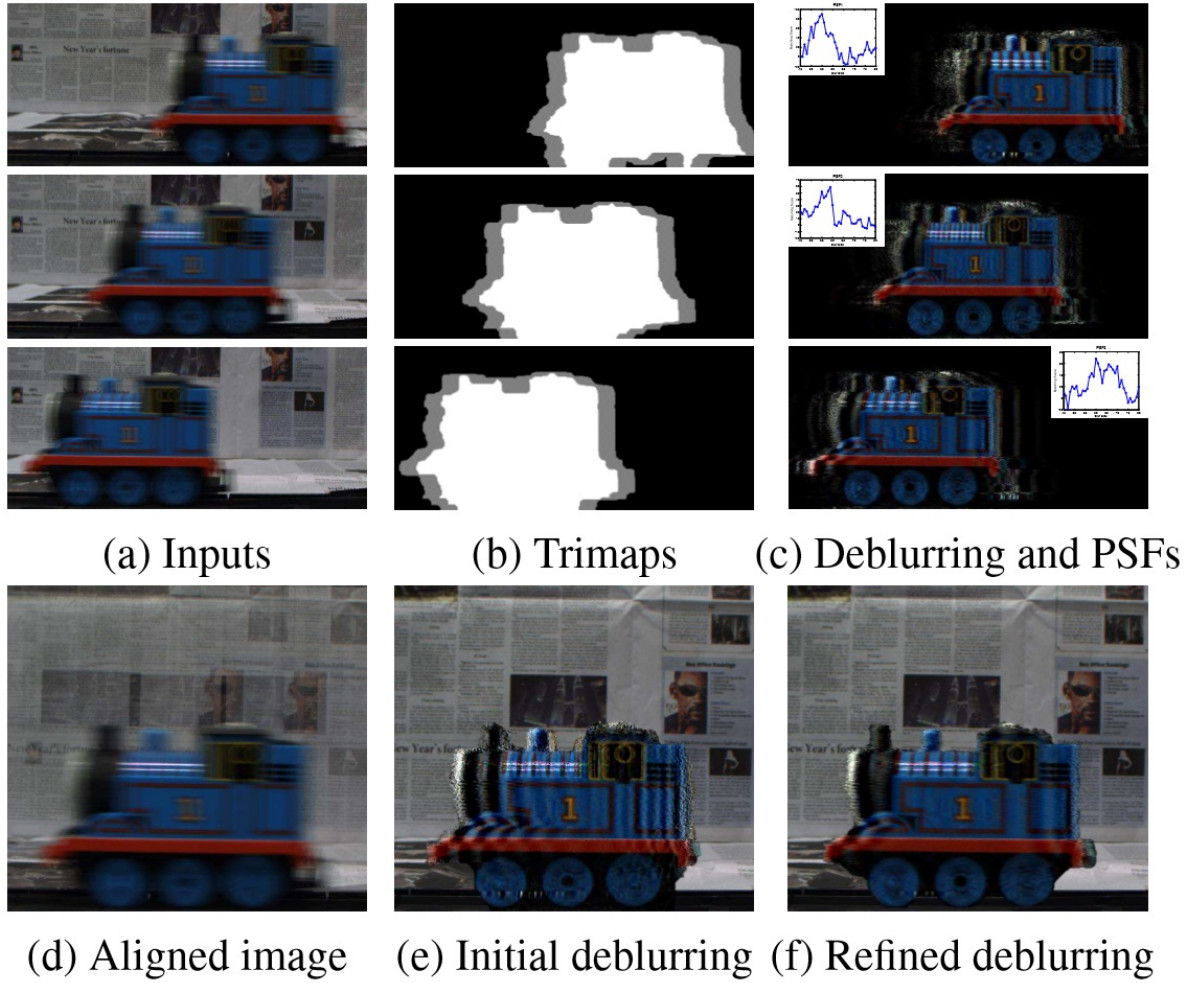


図 2.6 複数枚使用ブレ除去の中間出力

ソフトな制約として組み込まれる [62].

$$\arg \min_{\alpha} \alpha^T L \alpha + \lambda_m (\alpha - \hat{\alpha})^T D (\alpha - \hat{\alpha}) \quad (2.13)$$

ここで, L は閉形式マッピングのラプラシアン行列, D は対角行列, λ_m はソフトな制約の重みを示す.

リファインされたアルファマップを使用して, アフィン行列のセットを最適化した. これは, 次のような, ステレオマッチングと同様のエネルギー関数になる.

$$\arg \min_H \sum_{j=1}^{m-1} \lambda_{\alpha} \min(|X_{ref} - H_j X_j|), \tau_{color} \quad (2.14)$$

$$+ (1 - \lambda_{\alpha}) \min(|\nabla X_{ref} - \nabla(H_j X_j)|), \tau_{grad}$$

ここで, X は独立したブレ画像であり, X_{ref} は参照画像である. λ_α は色と勾配項とのバランスをとり, σ_{color} は外れ値の対応を示す切り捨て値である.

図 2.6 (f) に示すように, 本発明者らのアルゴリズムは, 複雑なバックグラウンドで同物体のブレ除去の有望な結果を示している. 高精度化は最終結果に対して 2~3 回反復され, MATLAB 実装では 800×600 解像度の画像で 5 分を要した. ここでは経験的に $\lambda_d, \lambda_m, \lambda_a, \tau_{color}, \tau_{grad} = 0.01, 0.1, 0.5, 0.3, 0.5$ を設定した.

2.5.4 実験結果

提案手法の有効性を検証するために, 符号化された露出撮像 [43] および複数露出ビデオ [45] のような計算イメージング手法との定量的および定性的比較を行うコード化された露光方法 [43] では, Agrawal のコードによって生成された長さ 48 の開閉パターンを使用した. [45] に記載された露光シーケンス [30,35,42ms] は, 変化した露光方法に使用される. 提案された方法の長さ 48 の開閉パターンは, [000010100100; 001001111101; 101000100011; 001110010111] を初期セットとして式 (2.10) を 1 回適用することによって生成される.

シミュレーション画像での実験

定量的評価のために, ここではシミュレーション画像での実験を行う. 合成データとして, Kodak Lossless True Color Image Suite からダウンロードした 29 枚の画像を使用. 画像ブラーは, 各方法によって生成された異なる露光シーケンスを用いた 1D フィルタリングによってシミュレートされる. 実際の写真をシミュレートするために, 強度偏差ガウス雑音を標準偏差 $\sigma = 0.01 \sqrt{i}$ で追加した. ここで, i は [0,1] [63] のブレ画像のノイズフリー強度を示す. ピーク信号対雑音比 (PSNR) およびグレースケール構造類似性 (SSIM) [64] が評価値として使用される. 公正な比較のために, 画像ブレ除去のためのパラメータチューニングを行い, 各画像/方法の最高 PSNR 値および SSIM 値を記録した.

図 2.7 は, 提案手法の平均 PSNR と SSIM を使用画像数を表す. より多くの画像でより良い性能が達成されることがわかるが, 画像の数が増えるにつれてパフォーマンスの向上率は低下した. この実験は, 3 つの開閉パターンを利用することは, 提案された方法の性能向上と複数画像のブレ除去の負担との間の良好なトレードオフであることを示している. したがって, 残りの実験には 3 つの開閉パターンを使用した.

異なる方法の定量的比較を図に示す. 完全な検証のために, ランダムサンプル探索 [43] によって生成された 2 組の符号化された露光シーケンスをさらに考慮した. 3 つのシーケンスの各セットは, 同じ開閉パターンと 3 つの異なるパターンからそれぞれ構成されている. コード化された露出ビデオ

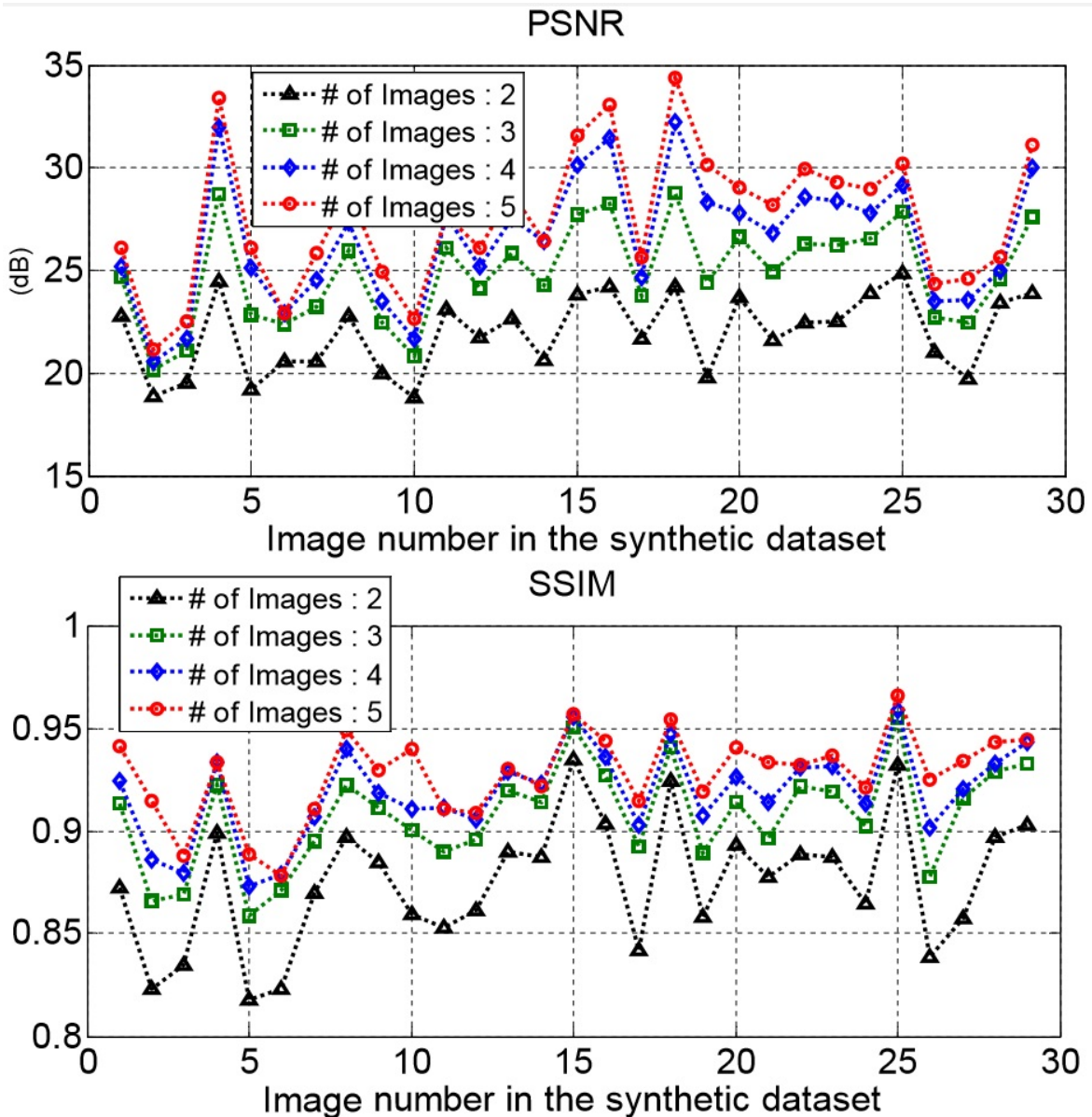


図 2.7 使用したシャッター開閉パターンの数に応じて提案された方法の性能変化

オに対する単一のコード化された露出のベースライン拡張として2つのシーケンスのセットを含める。提案された方法は、特に SSIM において大きなマージンを有したすべてのデータセットについて、以前の方法より優れていた。これは、提案された方法が、高い空間周波数の損失のためにテクスチャ領域を回復できない一方で、高品質のぼけ除去結果をもたらすためである。図 2.2 に合成結果の例を示す。ここでの方法は、他の方法よりも質的にも定量的にも優れている。

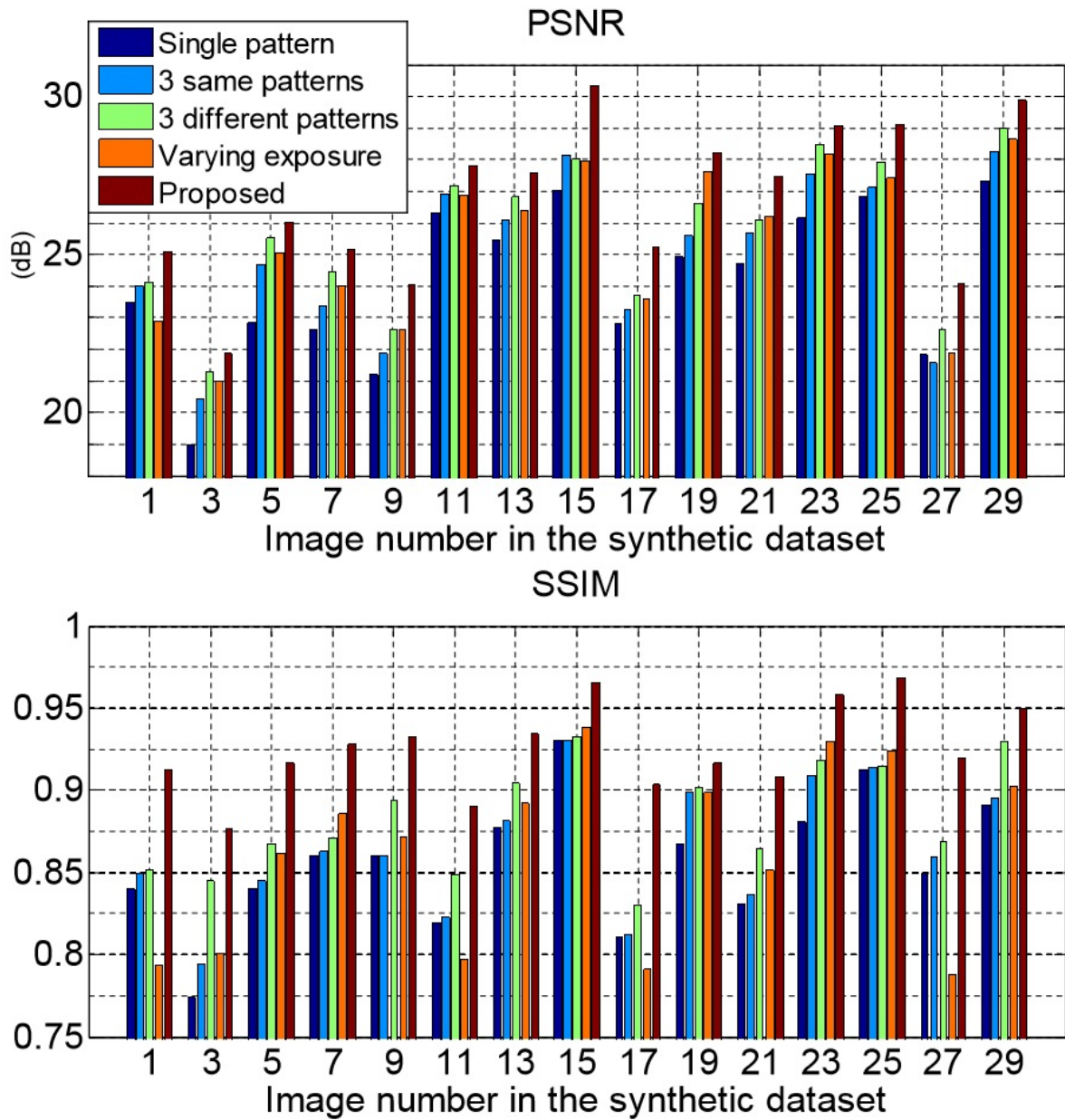
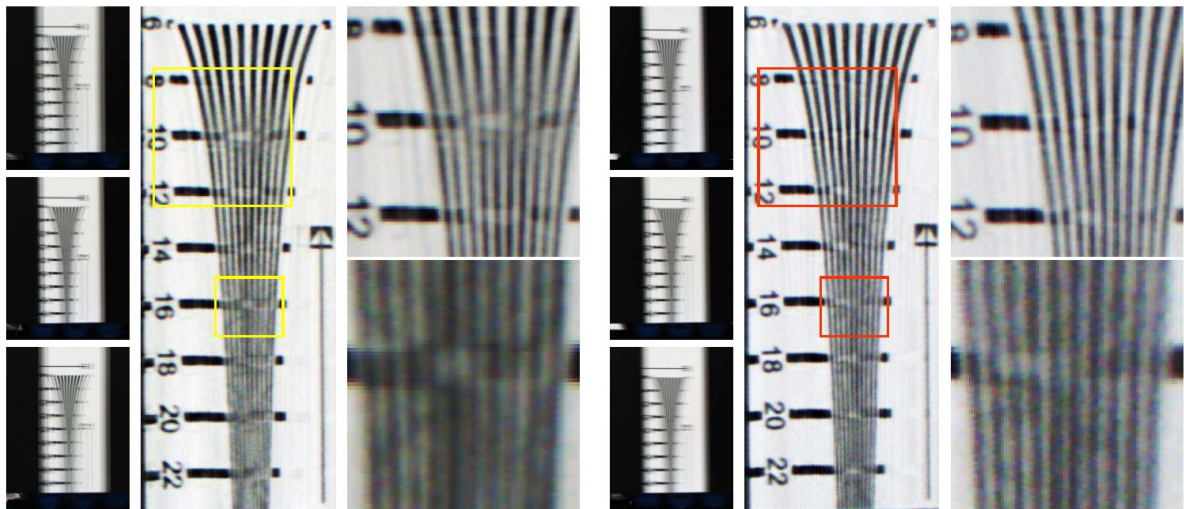


図 2.8 ブレ除去に関する異なる方法の定量的比較

実測画像での実験

図 2.9 では、性能問題の実証的検証を示す。慎重に管理された環境で解像度チャートを取得し、1 回の露光で 2 ピクセル移動した。ここでは、相補的なセットを、ランダムに生成された開閉パターン のセットと比較した。図 2.9 に示すように、PSF が引き伸ばされると、ランダムシーケンスセットによっ て生成されたブレ除去画像は、エッジ周辺に顕著なアーティファクトを有した。一方、提案された 相補的なセットは、解像度チャートの詳細をよく保存している。図 2.10 は、速度の異なる複数の



(a) A random sequences set

(b) Proposed

図 2.9 PSF を伸ばしたときのランダムセットと提案された補集合を用いた結果の比較

オブジェクトのブレ除去の結果を示している. 各オブジェクト・三色マップを別々にセグメンテーションするために, グラフカットによるマルチラベル最適化を行った [65].



図 2.10 複数物体があるシーンでのブレ除去. 左から符号化開口, 複数露光ビデオ, 複数パターン符号化開口ビデオによるブレ除去結果

次に, それぞれのオブジェクトをブレ除去して背景に独立して貼り付けた. 図 2.10 では, 一方のオブジェクトは高度にテクスチャ付けされ, 速く動く一方, 他方のオブジェクトはテクスチャが低く, 動きが遅い. 提案された方法は, 他のブレ除去方法と比較して最良の結果を示す.

次に, 図 2.11 に示すように, 高速で移動したオブジェクトを撮影して屋外で実測画像の別の実験を行った. 動きの方角とブラーカーネルは, 移動した車と同じ色を持つ複雑な背景で自動的に推定される. ここでも, コード化された露出ビデオ法が他の 2 つの方法よりも優れている.

ビデオ監視のためのプライバシー保護への応用

ビデオ監視のプライバシー保護は最近, ビデオ監視が一般的になってきたことから重要な問題となっている. コンピュータビジョンの問題に対処するための様々な試みがなされており [66, 67], プ



符号化開口



複数露光ビデオ



複数パターン符号化開口ビデオ

図 2.11 実測画像による実験

ライバシー保護のために監視ビデオを戦略的にブレを与えるような、曖昧なブラスキームを使用した興味深い研究もある [47].

共起ブレは、共起を満たす 2 つの異なるブローカーネルを適用してビデオストリームを暗号化し、公開ストリームとプライベートストリームを形成した。プライベート・ストリームとパブリック・ストリームの両方にアクセスできたときは、ブローのないストリームを矛盾のないブレ除去アルゴリズムを用いて回復できた。ブラインドデコンボリューションをパブリックストリームのみで適用することは非常に難しかったため、ビデオストリームのライバシーは保護され、フレームごとに異なるブローカーネルを選択したことでより高度なセキュリティを実現できた。Li ら [47] は 2 つの二進シーケンスからコモンブローカーネルを合成し、効率的なブレ除去アルゴリズムを提示した。彼らはまた、セキュリティレベルと回復されたビデオの品質に直接影響するため、平坦なスペクトルを持つブローカーネルのバンクを構築したことの重要性を強調した。

ここでの相補的な開閉パターンのセットは、同じ方法で虚像カーブを設計するために直接適用することができた。さまざまな長さの平坦なスペクトルを持つ多様な開閉パターンを生成することができたため、高レベルのセキュリティと高品質の回復の両方を実現することに適している。

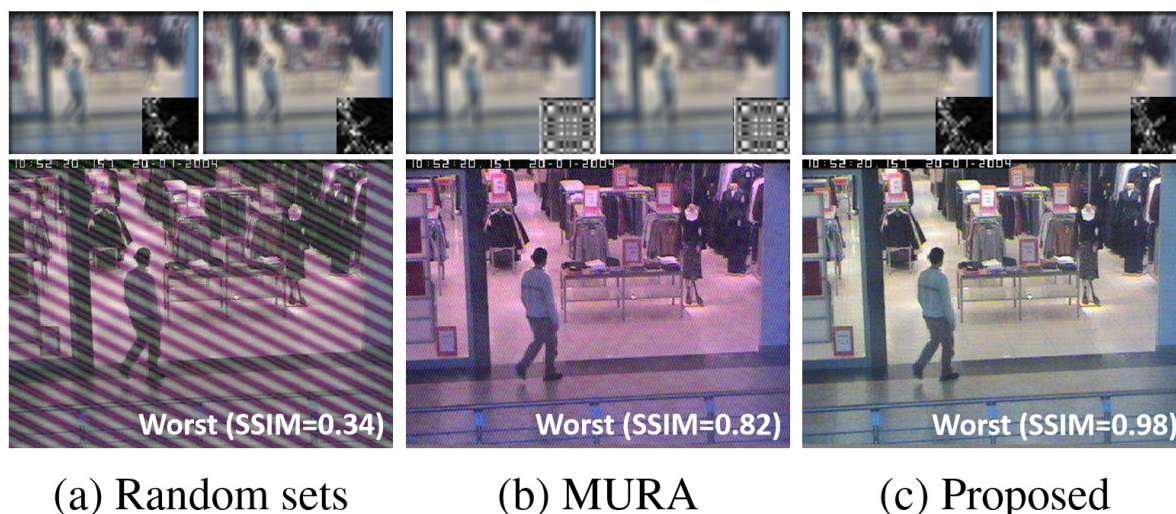


図 2.12 監視ビデオのライバシー保護

ここでは、共起ブラスキームに適用されたフレームワークの有効性を示すために実験を行った。まず、修正された一様冗長アレイ (MURA) [68] とランダムサンプル検索 [43] と補完シーケンスを用いて 6 対の共起ブローカーネルを用いて、1 組のコモンブローカーネルを生成した。次に、各フレームを合成的にブレを与えることにより動画を暗号化し、[47] の互いに素なブローを与えた画像を用いる方法 (CBP 法) に従って解読した。

共起法による暗号化結果と復号化結果の両方を図 2.12 に示す。図 2.12 第 1 の行は、虚偽のブローカーネルを有した暗号化されたフレームを表し、第 2 の行は、ブレ除去結果を示す。CBP 法

は、相補的な配列で高品質の再構成結果を一貫して生成したが、他の配列が使用される場合には重大なアーティファクトが発生することがある。これは、[44] で説明したように、大きなサンプルスペースで長さの長いシーケンスを生成することができず、MURA に深いディップが含まれ、Raskar ら [43] の手法で見られたスペクトルリークが発生するためである。一方、ここでの相補的配列は、様々な長さやセットを有した良好な配列対を生成することができた。

2.5.5 まとめ

複数画像のブレ除去のための新しい符号化露出ビデオフレームワークを導入した。提案された方法は、本質的に符号化された露光イメージングと複数の露光ビデオとを組み合わせ、2つの方法を利用して優れたブレ除去結果をもたらす。現在の仕事の限界は、等速オブジェクトの1D線形ブラーに対してのみ解くことである。歩行者や移動車のような多くの現実世界の物体の動きは Raskar ら [43] が述べたように1次元モーションブラーを生じることには注意したさらに、斜めのブラーシーンからのアフインブラーは、1Dモーションブラー問題と解釈することができた [43]。将来的には、既知の背景や線形運動の仮定などの現状の限界を克服し、マルチ画像超解像とロボット工学の視覚オドメトリの枠組みを適用するために、ここでの方法を拡張したいものである。

2.6 ピクセルごとにブレを計算したブレ除去法

いくつかの最先端のビデオブレ除去リング方法は、キャプチャされたシーンが静的であるという強い仮定に基づいている。これらの方法では、動的なシーンのブレのある動画をブレ除去することができない。ここでは、他の方法とは逆に、動的場面に固有の一般的なブラーに対処したビデオのブレ除去方法を提案した。手ブレのような様々な原因によって引き起こされる局所的に変化する一般的なブレを処理するために、ここでは、二方向のオプティカルフローを用いてピクセルごとにカーネルを近似した。そこで、ブレ除去の問題を解決するために、オプティカルフローとブレの無いフレームを同時に推定する単一のエネルギーモデルを提案した。また、エネルギーモデルを最適化するためのフレームワークと効率的なソルバーを提供した。提案されたエネルギー関数を最小限に抑えることにより、ブレを取り除き、ブレのあるフレームの正確なオプティカルフローを推定する際に重要な改善を達成した。広範な実験結果は、現実のビデオおよび挑戦的なビデオにおける提案された方法の優位性を示し、比較用の最先端の方法は、ブレ除去またはオプティカルフロー推定のいずれにおいても失敗した。

2.6.1 導入

モーションブラーは、手持ちカメラを使用して記録されたビデオの中で最も一般的なアーティファクトである。何十年もの間、いくつかの研究者がモーションブラーを対象としたブレ除去アルゴリズムを研究してきた。それらの方法は、撮影されたシーンが静的であるか否かに依存した。初期の単一画像のブレ除去に関する研究では、通常、シーンが一定の奥行きで静止していると仮定していた [69, 70, 71, 72, 13, 33]。ここでは成功したアプローチを自然な形でビデオのブレ除去に拡張した。Cai らの研究では [36]、ブラーカーネルの疎性と鮮明な領域を用いて位置合わせの誤差に対応した複数画像デコンボリューション法が提案された。しかし、この方法は、シーンに均一なブラーを与えるような 2 次元並進カメラの動きのビデオのみで可能である。したがって、提案された手法は、大きなモーションブラーの主な原因である回転した手ブレを扱うことができない [33]。この制限を克服するために、Li ら [73] は、 3×3 ホモグラフィで空間的に変化した動きをパラメータ化する方法を使用し、カメラの回転によって空間的に変化したブラーを処理することができた。Cho らの研究では [74]、3 次元空間におけるカメラの動きを特殊なハードウェアを利用せずに推定できた。さらに、投影カメラの動きによる不均一なブレを除去することができた。静的な場面における深さ変動による空間的に変化したブレは、Lee ら [75] の研究で最近扱われた [76]。しかしながら、シーンが静的であると仮定した以前のアプローチは、カメラの手ブレだけでなく、シーン内の動く物体や深度の変化からによる一般的なブレに悩まされていた。動的なシーンでの空間で変化したブラーカーネルをパラメータ化することは、単純なホモグラフィでは困難であるため、動的シーンに対応したカーネル推定はより困難になる。したがって、いくつかの研究者は、主に 2 つのアプローチ、セグメンテーションベースのアプローチと標本ベースのアプローチに分類される動的シーンの復元に焦点を当ててきた。

セグメンテーションベースのブレ除去のアプローチは、複数の動き、複数のカーネル、および関連した画像セグメントを同時に推定した。Cho ら [77] は、画像を均質な動きの複数の領域に分割し、対応したブラーカーネルを 1 次元ガウスクーネルとして推定した方法を提案した。したがって、この方法は、局所的に変化したブラーを生成するようなカメラの回転運動および物体の複雑な動きを処理することができない。Bar [78] は階層化されたモデルを提案し、2 つのレイヤー（前景と背景）に画像を分割した。さらに、彼らは前景層に対応した線形のブラーカーネルを推定した。この方法では、階層モデルを使用して遮蔽領域を明示的に処理できるが、カーネルは 1 次元のボックスフィルタのみに限定され、静的なカメラのみが使用できる。Wulff と Black [79] は Bar [78] らの以前の研究を拡張した。彼らは、前景と背景の両方の動きのパラメータを推定することに焦点を合わせた。しかし、各セグメント内の動きは、アフィンモデルを用いてのみパラメータ化され、多層シーンに拡

張することは困難である。なぜなら、このようなタスクは、レイヤの深度の順序の推定を必要としたからである。要約すると、セグメンテーションベースのアプローチは、動的なシーンで動いている物体によって生じるブラーを処理する点で有利である。しかし、各セグメントのモーションをパラメータ化することは依然として課題である [80]。つまり、Bar の手法 [78] で使われている単純なモデルを使ってそのようなことは難しいため、人の動きなど複雑な動きをパラメトリックに変化させることはできない。

Matsushita ら [81] および Cho ら [26] の研究は、典型的な標本ベースのアプローチである。これは、一般に長い画像シーケンスに存在した鮮明なパッチを補間することによってブラーの無いフレームを推定した。したがって、これらの方法は、リングングアーティファクトを出現させる要因となるセグメンテーションの精度およびデコンボリューションの工程を無視することができた。しかし、前者の方法では、オブジェクトが動くことによるブラーを処理することはできない。さらに、後者は、ホモグラフィからなるグローバルにパラメータ化されたカーネルを使用して、ブレのあるパッチに対応したブレの無いパッチを検索するため、シーン中でわずかに動く物体のみ処理が可能である。したがって、背景とは異なる動きを持つ動きの速いオブジェクトを扱うのは難しい。さらに、芝生や木々のような中間周波数のテクスチャを劣化させ、これは、空間的なプライヤーを用いたデコンボリューションを使用せず、補間を使用して潜在フレームをスムーズな結果が得られるように復元するためである。

ここでは従来の問題点を改善するために、グローバルモーションのパラメータ化とセグメンテーションを使わずに潜在フレームを推定する新しい一般的なビデオのブラー除去法を説明する。具体的には、双方向のオプティカルフローを推定し、それをを用いて画素単位で変化したカーネルを推定する。そのため、手ブレや複雑な動きのある動きのある被写体、奥行きの変化などで、混在したブレを自然に処理できる。しかしながら、正確なオプティカルフローを得るためには鮮明なフレームが必要である。なぜなら、領域ごとのフローを推定することは、ぼやけた画像間では難しいからである。さらに、鮮明なフレームを復元するためには、正確なオプティカルフローが必要である。このケースは典型的なチキンエッグ問題であり、したがってここでは同時に両方の変数を推定する。そこで、ここでは共同問題を解決するための新しい単一エネルギーモデルを提案する。また、モデルを最適化するためのフレームワークと効率的な手法を提供した。正確なオプティカルフローが連帯的に推定されるため、移動車が正常に復元された結果が図 2.13 に示されている

提案されたエネルギー関数を最小限に抑えることによって、図 2.14 に示すように、ここでは他の方法ではできない多数の実際の難しいビデオの大幅な改善を達成した。さらに、ブラーのある画像を処理した最先端のフロー推定法と比較して、より正確なオプティカルフローを推定した。パフォーマンスはここでの広範な実験で実証されている。

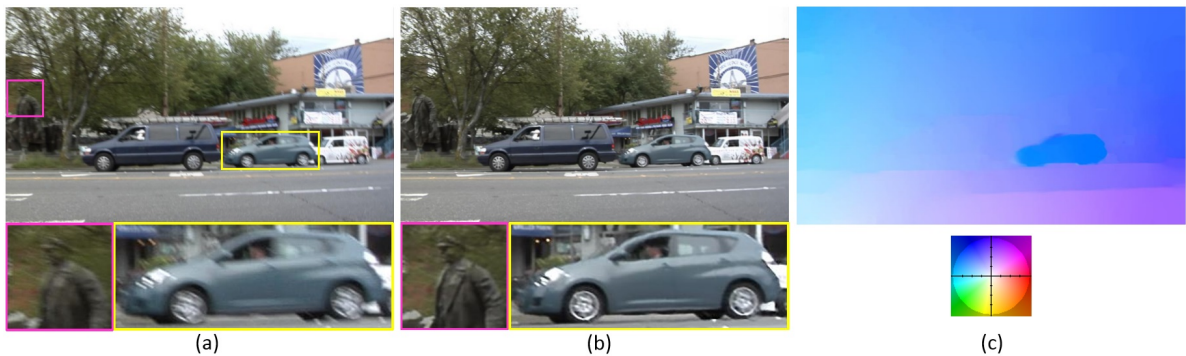


図 2.13 (a) 動いている車を含むブレがあるシーン.(b) この手法でのブレ除去結果.(c) オプティカルフロー

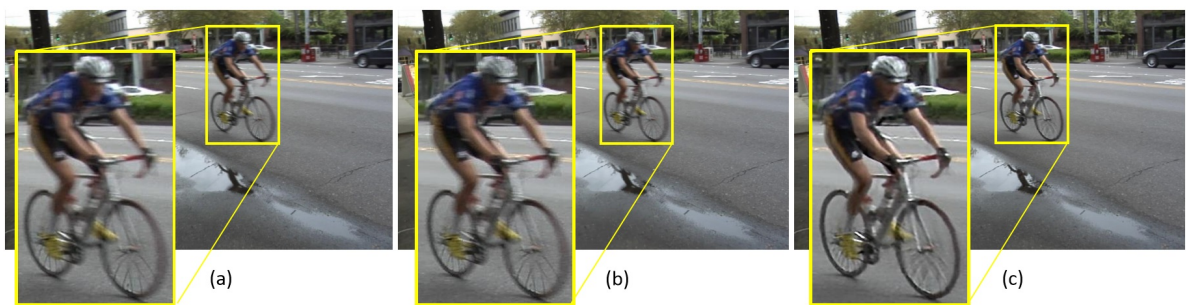


図 2.14 (a) ブレのある自転車シーン (b)Cho らのブレ除去結果 (c) この手法でのブレ除去結果

2.6.2 一般的な動画のブレ除去

ほとんどの従来のビデオブレ除去方法は、シーン全体またはセグメントごとにパラメータ化を推定しても動きを表現できないので、動的シーンから発生した様々な動きブレが共存することに苦しんでいる。一般的なブレを扱うために、シーン全体またはセグメントごとのパラメータ化ではなく、ピクセル単位のカーネル推定を使用した新しいエネルギーモデルを提案した。ブラインドブレ除去はよく知られている不良条件問題であるため、ここでのエネルギーモデルはデータと空間的正則化用語だけでなく、一時的な項も含んでいる。モデルは以下のように表される。

$$E = E_{data} + E_{temporal} + E_{spatial} \quad (2.15)$$

式 2.15 の各項の詳細は以下に述べる。

ブレ推定ベースのデータモデル

従来の作品では、各フレームのモーションブラーは、ホモグラフィやアフィンなどのパラメトリックモデルを用いて近似されている [78, 26, 73]。しかし、これらのカーネル近似は、モーションブラーがフレームまたはセグメント全体内でパラメータ化可能である場合に有効である。したがって、

一般的なブレに対処するには、ピクセル単位の動きとカーネル推定が必要である。ここでは、モーションブラー除去の研究 [58, 80, 82] に従い、双方向のオプティカルフローを用いてピクセル単位のブラーカーネルを近似した。

具体的には、動きの速度が隣接したフレーム間で一定であるという仮定の下で、ここでのブレモデルは以下のように表される。

$$B_i = \frac{1}{2\tau_i} \int_0^{\tau_i} H(L_i, t \cdot U_{i \rightarrow i+1}) + H(L_i, t \cdot U_{i \rightarrow i-1}) dt \quad (2.16)$$

ここで、 $U_{i \rightarrow i+1} = (u_{i \rightarrow i+1}, v_{i \rightarrow i+1})$ であり、 $U_{i \rightarrow i-1} = (u_{i \rightarrow i-1}, v_{i \rightarrow i-1})$ はフレーム i における双方向のオプティカルフローを示す。

ブレのあるフレームと潜在フレームはそれぞれ B_i と L_i である。フレームのカメラデューティサイクルは τ_i であり、相対露出時間 [73] を示す。画像の射影を定義する。 $H(L_i, t \cdot U_{i \rightarrow i+1})$ を $0 \leq t \leq 1$ のとき L_i から L_{i+t} への射影とし、 $H(L_i, t \cdot U_{i \rightarrow i-1})$ は L_i から L_{i-t} への射影とした。ここでのブラーモデルで使用されている双方向のオプティカルフロー、デューティサイクル、および対応した区分線形カーネルを図 2.15 に示す。

ブラーカーネルモデルは単純であるが、露出時間の短いビデオをある程度処理するため、モデルを正当化できる。したがって、ここでは、カーネルを双方向のオプティカルフローを用いて区分線形として近似した：

$$k_{i,x}(u, v) = \begin{cases} \frac{\delta(uv_{i \rightarrow i+1} - uv_{i \rightarrow i+1})}{2\tau_i \|u_{i \rightarrow i+1}\|}, & \text{if } u \in [0, \tau_i u_{i \rightarrow i+1}], v \in [0, \tau_i v_{i \rightarrow i+1}] \\ \frac{\delta(uv_{i \rightarrow i-1} - uv_{i \rightarrow i-1})}{2\tau_i \|u_{i \rightarrow i-1}\|}, & \text{if } u \in [0, \tau_i u_{i \rightarrow i-1}], v \in [0, \tau_i v_{i \rightarrow i-1}] \\ 0, & \text{otherwise} \end{cases} \quad (2.17)$$

ここで、 $k_{i,x}(u, v)$ は、画素 x での双方向オプティカルフローを用いたブラーカーネルであり、 δ は、クロネッカーデルタである。

このピクセル別カーネル近似を使用すると、従来の方法とは異なり、フレーム内の複数の異なるブラーを簡単に管理できた。ここでのカーネルモデルの優位性は、図 2.16 に示されている。ここでのカーネルモデルは、異なる動きのオブジェクトとカメラの動きからブレが発生する、従来のホモグラフィベースのモデルよりもはるかに優れている。

したがって、ここでは、オプティカルフロー推定の問題として、ピクセル単位のカーネル推定問題をキャストした。制約条件 (2.16) を離散化すると、以下のデータ項が得られる。

$$E_{data}(L, u, B) = \lambda \sum_i \sum_{\partial_*} \|\partial_* K_i(\tau_i, u_{i \rightarrow i+1}, u_{i \rightarrow i-1}) L_i - \partial_* B_i\|^2 \quad (2.18)$$

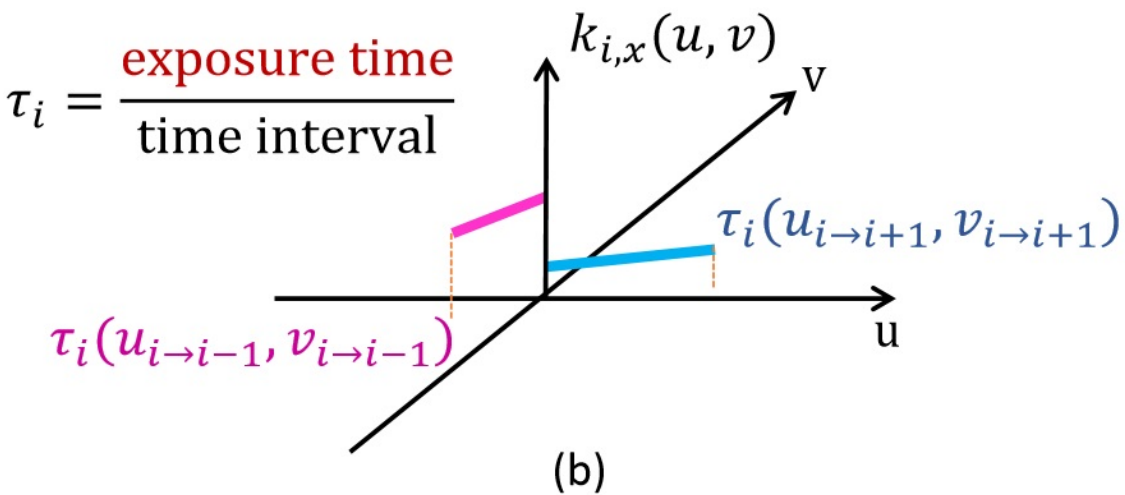
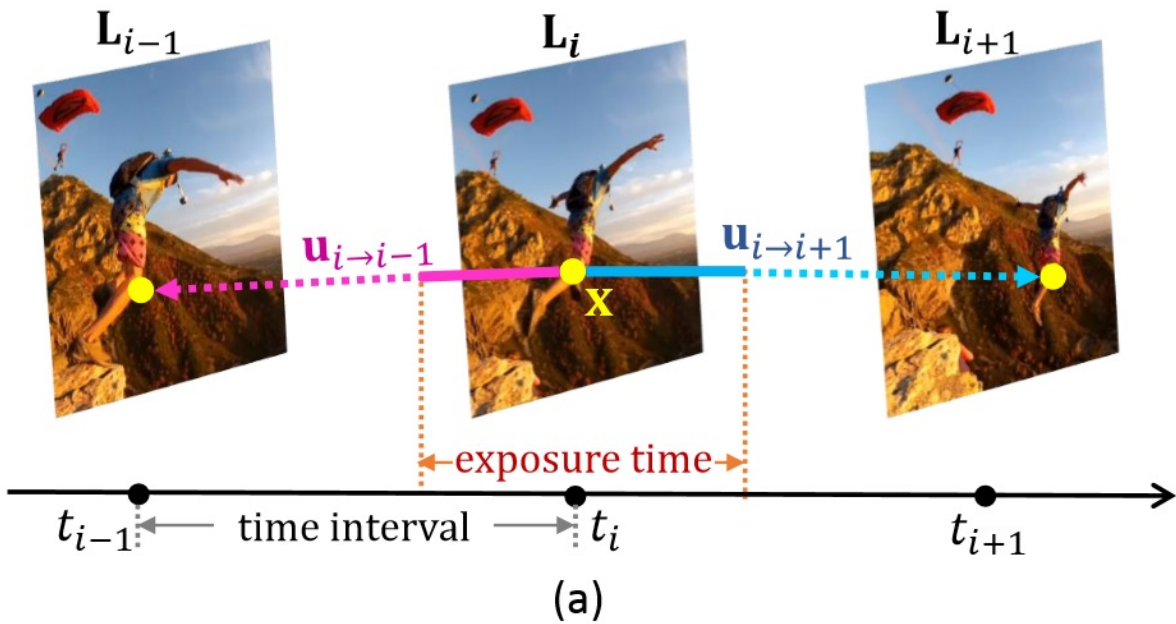


図 2.15 (a) 双方向オプティカルフロー (b) ピクセル位置 x における部分ごとの線形ブラーカーネル

ここで、画素 x のブラーカーネルに対応したブラーカーネル行列 K_i の行ベクトルは $k_{i,x}(\cdot)$ のベクトル形式であり、その要素は非負であり、それらの和は 1 に等しい。線形演算子 ∂_* は、部分（例えば、水平および垂直）微分フィルタに対応したテプリッツ行列を示す。パラメータ λ はデータ項の重みを制御し、 L, u, B はそれぞれ潜在フレーム、オプティカルフローおよびブレたフレームの集合を表す。

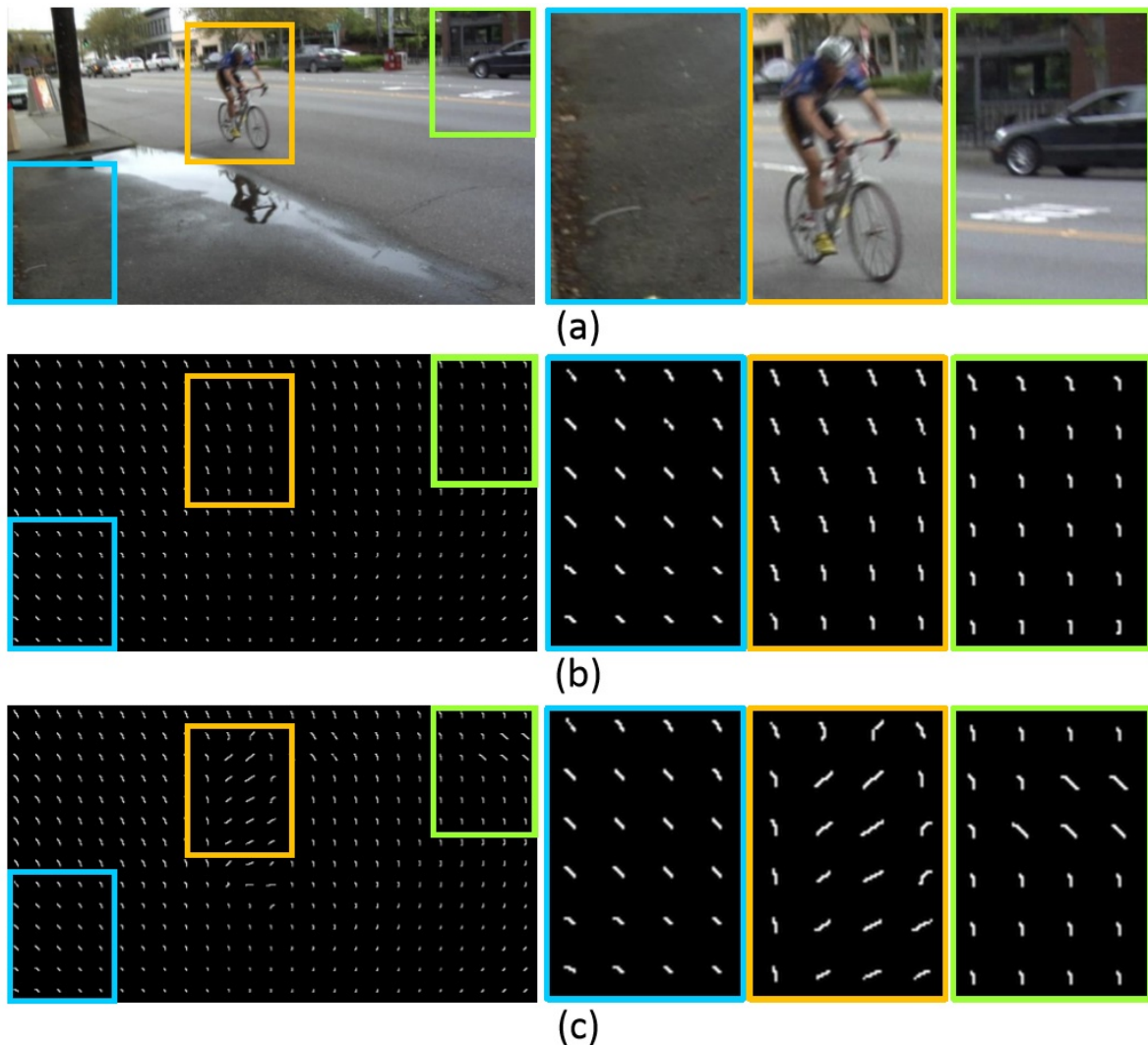


図 2.16 (a) 動的なシーンビデオのブレのあるフレーム (b) ホモグラフィーを用いた局所的に変化するブラーカーネル推定 (c) 双方向オプティカルフローを用いたピクセルごとのブラーカーネル推定

オプティカルフロー制約付き時系列コヒーレンス

ここでは、ピクセル単位のブラーカーネルを推定するためにオプティカルフローが必要であると判断した。しかしながら、提案されたデータ項は、(2.18)において輝度不変性または勾配不変性のような従来のオプティカルフロー制約を持たない。一般に、このような制約は、2つのブレのあるフレームの間には存在しない。したがって、Portz ら [82] は、ブレのある画像の間にフローの制約を適用する方法を提案した。Kim ら [83] のカーネルのシフト不変性の可換則に基づいて Zhang ら [82] は、各観測画像の近似されたブラーを他の画像に畳み込み、それらの間の一定の明るさを一致点で仮定した。しかし、カーネルが変換不変でないときには、可換プロパティは理論上保持されない。し

たがって、このアプローチは、モーションが十分に滑らかである場合にのみ機能する。

この問題を解決するために、ここでは、2つの潜在的な鮮明な画像間の対応関係を突き止め、動きおよび対応したカーネルの急な変化を可能にする新しいモデルを提案した。このモデルを使用する際には、ブローカーネルの変更は無効にする必要は無い。ここでのモデルは、潜像間の従来のオプティカルフローの制約、すなわち輝度一定性に基づいている。式は以下のように表される：

$$E_{temporal}(L, u) = \sum_i \sum_{n=-N}^N \mu_n |L_i(x) - L_{i+n}(x + u_{i \rightarrow i+n})| \quad (2.19)$$

ここで、 n は、 i における隣接したフレームのインデックスを示す。定数パラメータ μ_n は、総和の各項の重みを制御する。ロバストな L1 ノルムを適用して、外れ値とオクルージョンに対して頑健性を提供する。

特に、提案モデルと従来のオプティカルフロー推定法との主な違いは、ここでの問題は共同問題であるということである。すなわち、潜在フレームと光線の輝度を同時に推定する必要がある。したがって、ここでのモデルは潜在的フレームの時間的コヒーレンスを同時に強制し、対応関係を推定した。

空間的一貫性

高度に不良条件なブレ除去およびオプティカルフロー推定の問題の難点を緩和するために、いくつかの研究者が空間的正則化の重要性を強調してきた。したがって、空間的な一貫性を強制し、空間的な変動にペナルティを課しながら、潜在的なフレームとオプティカルフローの両方で不連続性を許容した。潜在フレームとオプティカルフローの空間事前分布は独立していると仮定する。それらは次のように表される。

$$E_{spatial}(L, u) = \sum_i |\nabla L_i| + \sum_{n=-N}^N g_i(x) |\nabla u_{i \rightarrow i+n}| \quad (2.20)$$

(2.20) の第 1 項は、潜在フレームの空間的正則化項を示す。総変量 (TV) が計算機的に安価であるため、より疎な L_p ノルム (例えば、 $p = 0.8$) が自然な鮮明な画像の勾配統計を良くしたが、ここでは従来の TV に基づく正規化を用いる。第 2 項は、オプティカルフローの空間的な滑らかさの項を表す。ここでは、端部のフローフィールドにおける不連続性を保持するために、エッジマップ結合 TV ベースの正則化 [84] を採用した。Kim ら [80] と同様に、エッジマップは以下のように表される。

$$g_i(x) = \nu \exp\left(-\left(\frac{|\nabla \bar{L}_i|}{\tau_I}\right)^2\right) \quad (2.21)$$

∇ はエッジマップのスケールを制御し、パラメータ τ_I は重みを制御し、 \bar{L}_i は反復最適化フレームワークにおける初期潜像である。

2.6.3 最適化フレームワーク

前節では、 E_{data} , $E_{temporal}$, $E_{spatial}$ の各項について説明した。カメラのデューティサイクル τ_I が分かれば、最終目的関数は次のようになる。

$$\begin{aligned} \min_{L, u} \lambda \sum_i \sum_{\partial_*} \|\partial_* K_i(\tau_i, u_{i \rightarrow i+1}, u_{i \rightarrow i-1}) L_i - \partial_* B_i\|^2 + \\ \sum_i \sum_{n=-N}^N \mu_n \cdot |L_i(x) - L_{i+n}(x + u_{i \rightarrow i+n})| + \\ \sum_i |\nabla L_i| + \sum_{n=-N}^N g_i(x) |\nabla u_{i \rightarrow i+n}| \end{aligned} \quad (2.22)$$

Cho ら [69] の研究とは異なり、ここでのモデルは、単一の目的関数を最小化することによって解を得る。しかし、その非凸性のために、ここでのモデルは近似解を得るための実用的な最適化法を採用した必要がある。したがって、元の問題を2つの部分問題に分割し、従来の反復最適化手法 [69, 79] を使用して、非凸の目的関数を最小化した。以下のセクションでは、効率的なソルバーを紹介し、未知数 L と u を推定する方法を説明した。

鮮明なビデオ復元

オプティカルフロー u が固定されている間、対応するブローカーネルもまた固定され、式 (2.22) の目的関数は L に関して凸面となり、以下のように表される。

$$\begin{aligned} \min_L \lambda \sum_i \sum_{\partial_*} \|\partial_* K_i L_i - \partial_* B_i\|^2 + \\ \sum_i \sum_{n=-N}^N \mu_n \cdot |L_i(x) - L_{i+n}(x + u_{i \rightarrow i+n})| + |\nabla L_i| \end{aligned} \quad (2.23)$$

L を得るために、Chambolle ら [85] の従来の凸最適化法を採用し、以下のように更新スキームを導出した。

$$\begin{cases} s_i^{m+1} = \frac{s_i^m + \eta_L AL_i^m}{\max(1, \text{abs}(s_i^m + \eta_L AL_i^m))} \\ q_{i,n}^{m+1} = \frac{q_{i,n}^m + \eta_L \mu_n D_{i,n} \begin{bmatrix} L_i^m \\ L_{i+n}^m \end{bmatrix}}{\max(1, \text{abs}(\begin{bmatrix} L_i^m \\ L_{i+n}^m \end{bmatrix}))} \\ L_i^{m+1} = \arg \min_{L_i} \lambda \sum \partial_* (\partial_* K_i L_i - \partial_* B_i)^2 + \\ \frac{(L_i - (L_i^m - \epsilon_L (A^T s_i^{m+1} + \sum_{n=-N}^N \mu_n D_{i,n}^T q_{i,n}^{m+1})))^2}{2\epsilon_L} \end{cases} \quad (2.24)$$

ここで, m_0 は反復回数を示し, s_i と $q_{i,n}$ は二重変数を示す. パラメータ η_L および ϵ_L は, 更新ステップを示す. 線形演算子 A は隣接画素間の空間差を計算し, 別の演算子 $D_{i,n}$ は $L_i(x)$ と $L_{i+n}(x + u_{i \rightarrow i+n})$ の間の時間差を計算する. 式 (2.24) において, 主変数を更新して L_i^{m+1} を得るために, 共役勾配法を適用して二次関数を最適化した

オプティカルフロー推定

潜在フレーム L は固定されているが, 時間コヒーレンス項 $E_{temporal}$ は凸になるが, E_{data} は非凸のままである. したがって, 非凸面忠実度関数 $\rho(\cdot)$ は次のように定義される.

$$\begin{aligned} \rho(x, u) = \lambda \sum_i \sum_{\partial_*} \|\partial_* K_i(\tau_i, u_{i \rightarrow i+1}, u_{i \rightarrow i-1}) L_i - \partial_* B_i\|^2 + \\ \sum_i \sum_{n=-N}^N \mu_n \cdot |L_i(x) - L_{i+n}(x + u_{i \rightarrow i+n})| \end{aligned} \quad (2.25)$$

最適化されたオプティカルフローの値を求めるために, 最初にテイラー展開を適用して非凸関数 $\rho(\cdot)$ を凸にした. Kim ら [80] と同様に, 反復処理の初期 u_0 付近で次のように関数を線形化した.

$$\rho(x, u) \approx \rho(x, u_0) + \nabla \rho(x, u_0)^T (u - u_0) \quad (2.26)$$

したがって, ここで近似するオプティカルフロー推定のための凸関数は以下のように表される.

$$\min_u \rho(x, u_0) + \nabla \rho(x, u_0)^T (u - u_0) + \sum_i \sum_{n=-N}^N g_i(x) |\nabla u_{i \rightarrow i+n}| \quad (2.27)$$

次に, 近似凸関数 (2.27) に Chambolle ら [85] の凸最適化手法を適用し, primal-dual 更新プロセスを以下のように表現した.

$$\begin{cases} p_{i,n}^{m+1} = \frac{p_{i,n}^m + \eta_u (G_i A) u_{i \rightarrow i+n}^m}{\max(1, \text{abs}(p_{i,n}^m + \eta_u (G_i A) u_{i \rightarrow i+n}^m))} \\ u_{i \rightarrow i+n}^{m+1} = (u_{i \rightarrow i+n}^m - \epsilon_u (G_i A)^T p_{i,n}^{m+1}) - \epsilon_u \nabla_{i,n} \rho(x, u_0) \end{cases} \quad (2.28)$$

$P_{i,n}$ はベクトル空間上の $u_{i \rightarrow i+n}$ の二重変数を表し, 対角行列 G_i は $G_i = \text{diag}(g_i(x))$ と表される重み行列である. パラメータ η_u と ϵ_u は更新ステップを表し, $\nabla_{i,n}\rho(x, u_0)$ は $\frac{\partial \rho(x, u)}{\partial u_{i \rightarrow i+n}}|_{u_0}$ を意味する.

2.6.4 実装の詳細

大きなブレを処理し, 高速な収束を導くために, ここでは経験的に決定されたパラメータを用いて従来のコーストアップファインのフレームワークにアルゴリズムを実装した. $\lambda = 250, \sigma_I = \frac{25}{255}, N = 2$ として, 他のパラメータを決定した. コーストアップファインのフレームワークでは, 高精細 (1280x720) ビデオの場合 17 レベルの画像ピラミッドを構築し, 倍率は 0.9 で, 次のピラミッドレベルにオプティカルフローと潜在フレームの両方を伝播するバイキュービック補間を使用した.



図 2.17 ある 3 フレームにおけるオプティカルフローの時間的一貫性

さらに, オプティカルフローにおける未知数の数を減らすために, $u_{i \rightarrow i+1}$ と $u_{i \rightarrow i-1}$ を推定した. $u_{i \rightarrow i+1}$ と $u_{i+1 \rightarrow i+2}$ を用いて $u_{i \rightarrow i+2}$ を近似した. 例えば, 図 2.17 に示すように, $u_{i \rightarrow i+2} = u_{i \rightarrow i+1} + u_{i+1 \rightarrow i+2}$ となり, これを $n \neq 1$ に対して簡単に適用することができた.

ここでのアルゴリズムの全体的なプロセスは, アルゴリズム 1 にある. デューティ比 τ_i の推定とアーティファクトを低減した後処理ステップの詳細については, 以下で説明する.

デューティ比推定

この研究では, カメラのデューティ比 τ_i はすべてのフレームで既知であると仮定した. Kinect を使用して RGB ビデオを撮影すると, 公開 SDK からデューティ比を得ることができた. しかし, 露出情報を提供しない従来のデータセットを用いてブレ除去を行う場合, Cho ら [26] の方法で提案された技術を適用してデューティ比を推定した. ここでは Cho ら [26] の方法とは対照的に, ホモグラフィの代わりにオプティカルフローを使用して, 最初に近似されたブラーカーネルを得る. し

Algorithm 1 手法の概要

Require: ブレのある動画のフレーム B

Ensure: ブレの無い動画のフレーム L とオプティカルフロー u

- 1: デューティ比 τ_i とオプティカルフロー u の初期化.
 - 2: 画像ピラミッドを構築した.
 - 3: 更新された u で鮮明な映像を復元した.
 - 4: 更新された L でオプティカルフローを推定した.
 - 5: オクルージョンを検出し, 後処理を行いる.
 - 6: 変数が存在した場合, 次のピラミッドレベルに伝播した.
 - 7: 粗いピラミッドレベルから 3-6 のステップを繰り返した
-

たがって, ここではほぼリアルタイムで実行される Wedel らの手法を [86] を用いて, ブレのある画像からオプティカルフローを最初に推定した. 次にそれらを初期のオプティカルとして使用し, ブラカーネルを近似してデューティ比を推定した.

オクルージョン検出と訂正

ここでの区分的線形カーネルは当然近似誤差をもたらし, リンギングアーティファクトなどの問題を引き起こした. また, この手法のデータモデル (2.18), (2.19) の時系列コヒーレンスモデルは閉塞領域では無効である.

このようなカーネルの誤推定やオクルージョンからなるアーティファクトを減らすために, 後処理として時空間フィルタリングを使用した.

$$L_i^{m+1}(x) = \frac{1}{z(x)} \sum_{n=-N}^N \sum_y w_{i,n}(x,y) \cdot L_{i+n}^m(y) \quad (2.29)$$

ここで, y は位置 $(x + u_{i \rightarrow i+n})$ の 3×3 隣接パッチ内のピクセルを示し, Z は $Z(x) = \sum_{n=-N}^N \sum_y w_{i,n}(x,y)$ のような正規化係数を示す. 注目すべきことに, ここでは, 空間フィルタリングのために (2.29) において $n = 0$ を満たした. ここでのオクルージョン認識重み $w_{i,n}$ は, 以下のように定義される.

$$w_{i,n}(x,y) = o_{i,n} \cdot \exp\left(-\frac{\|P_i(x) - P_{i+n}(y)\|^2}{2\sigma_w^2}\right) \quad (2.30)$$

オクルージョン状態 $o_{i,n}(x,y) \in 0, 0.5, 1$ は Kim らの方法 [84] で提案された方法を用いて決定される. 5×5 パッチ $P_i(x)$ は, フレーム i の中心 x に置かれる. 類似度制御パラメータ σ_w は, $\sigma_w = 25/255$ として固定される.

2.6.5 実験結果

以下では、提案された方法の優位性を実証した。

まず、ブレ除去の結果を、最新の標本ベースの方法 [26] と比較して、Cho らの方法 [26] で使用されたビデオと比較した。図 2.18 に示すように、撮影されたシーンは動的であり、複数の移動オブジェクトを含む。Cho らの方法 [26] は、オブジェクトの動きが大きく、背景とは異なるため、移動したオブジェクトの復元に失敗した。対照的に、ここでの結果は、動く物体や背景をぼかすことなく、より良いパフォーマンスを示している。図 2.19 に示すように、この標本的なアプローチはまた、大きなブレの画像の最初に推定されたホモグラフィが不正確であるため、大きなブレを扱うことに失敗した。さらに、この手法は、樹木のような中間周波数のテクスチャについては、ブレの無いフレームのための空間的なプライヤーのない補間に基づくため、滑らかな結果を生成した。

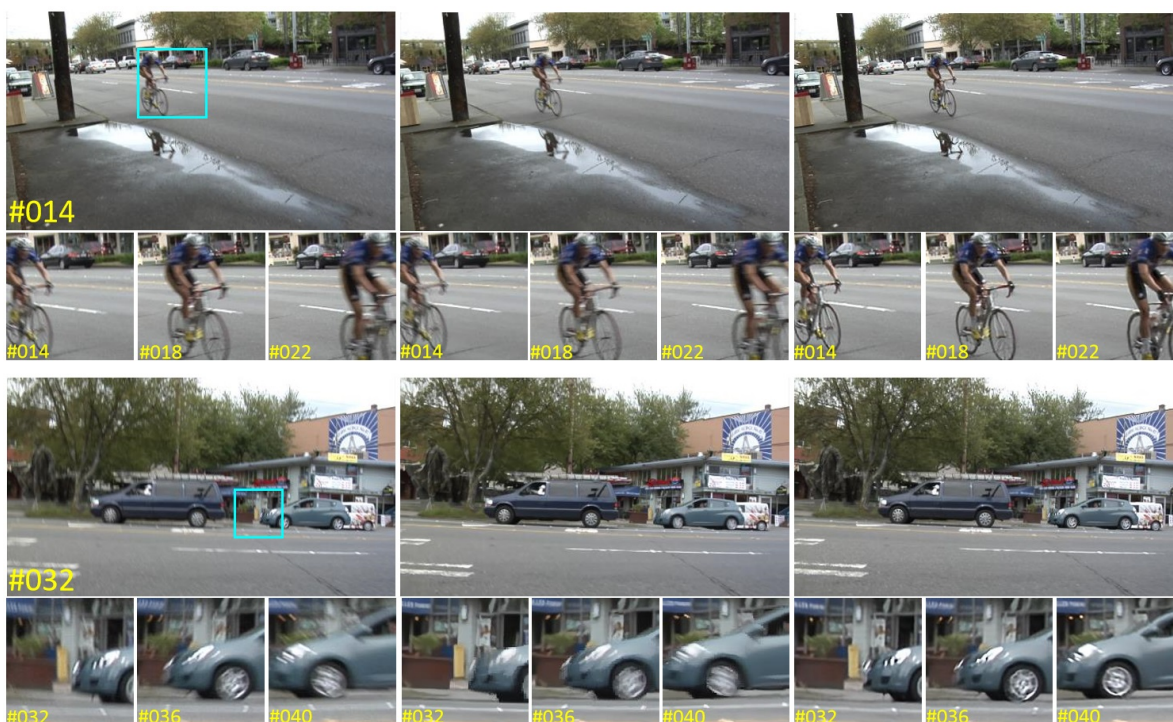


図 2.18 左から：動的シーンのブレのあるフレーム、 Cho らのブレ除去結果、 この手法でのブレ除去結果

次に、ここでの方法を最先端のセグメンテーションに基づくアプローチ [79] と比較した。図 2.20 において、撮影されたシーンは 2 層であり [79] で使用される。2 層シーンは、レイヤーモデルのパフォーマンスを検証するための良い例であるが、境界付近の不正確なセグメンテーションは、復元されたフレームに深刻なアーティファクトを引き起こす。対照的に、本発明の方法は、正確なセグメンテーションに依存せず、したがって、階層化されたモデルよりもはるかに良好に境界を復元した。

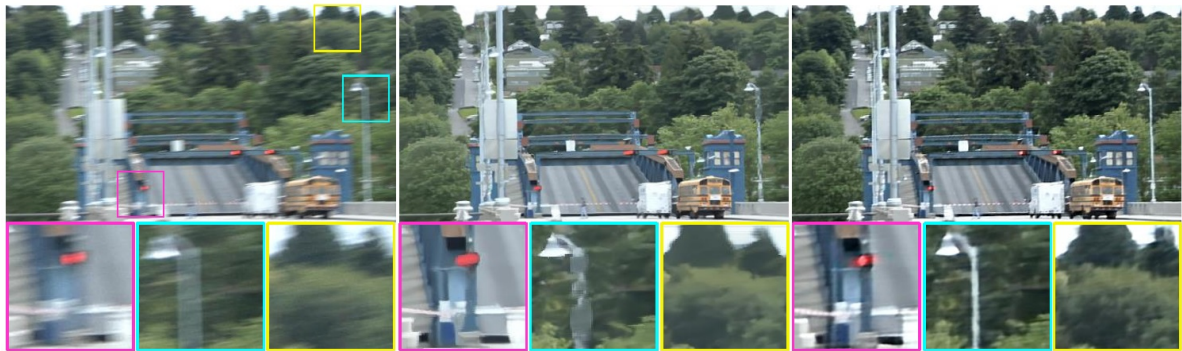


図 2.19 左から：ブレのあるフレーム， Wulff らのブレ除去結果， この手法でのブレ除去結果



図 2.20 セグメンテーションベースの手法との比較. 左から：ブレのあるフレーム， Cho らのブレ除去結果， この手法でのブレ除去結果

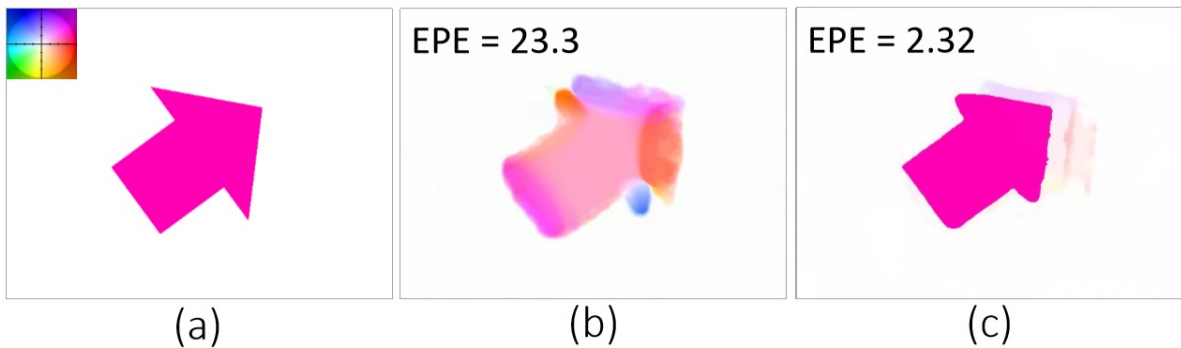
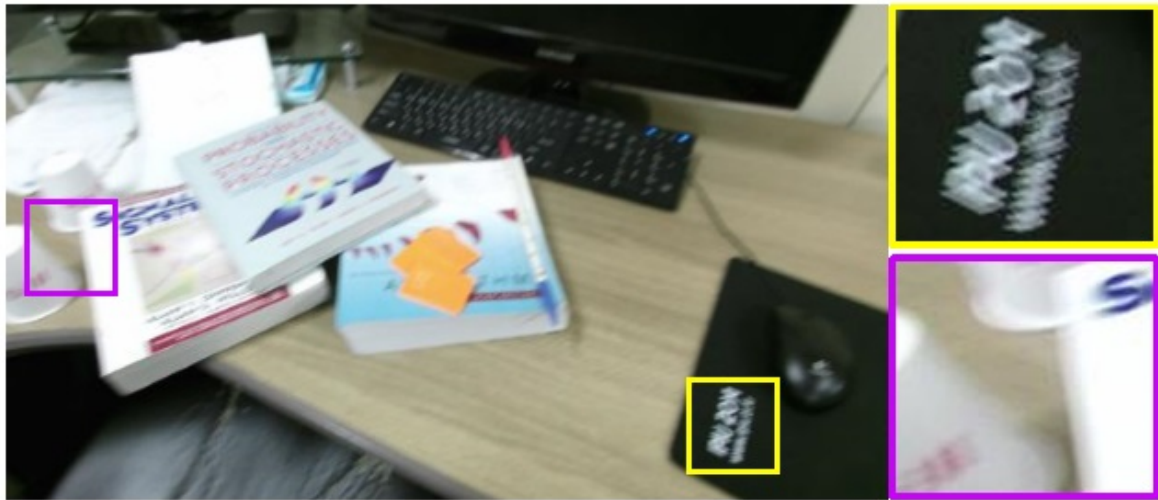


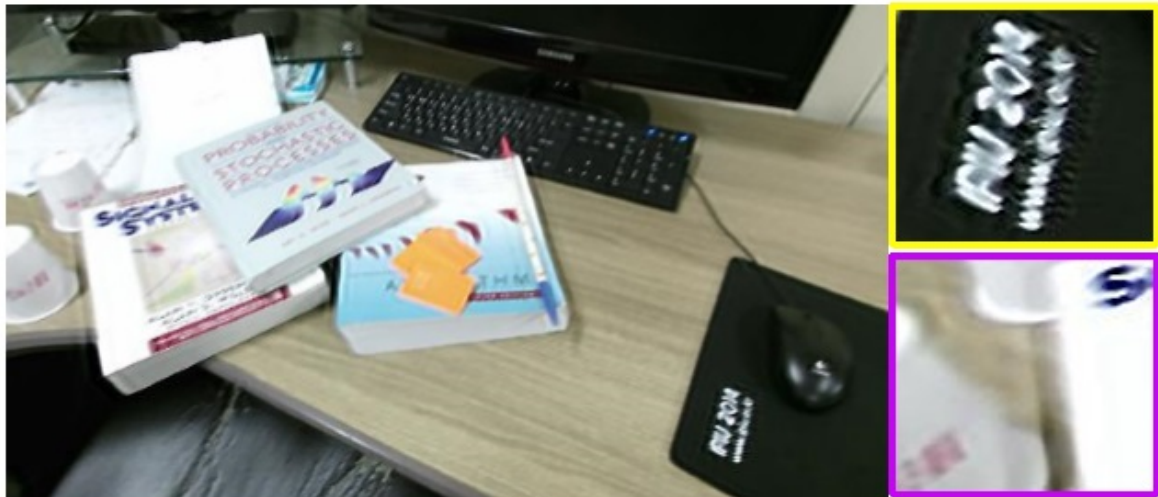
図 2.21 EPE はエラーの平均を表す. (a) オプティカルフローの GroundTruth (b)Portz らのオプティカルフロー推定結果 (c) この手法でのオプティカルフロー推定結果

図 2.21 において，ここでは定量的に，オプティカルフローの精度を合成したブラー画像で Zhang ら [82] の方法と比較した. Zhang ら [82] の手法は，オプティカルフロー推定においてブレのある画像を扱うことが提案されているが，ブレ除去には非常に重要な動き境界が仮定にはない. したがって，それらのオプティカルフローは，動く物体の運動境界において不正確である. しかし，ここでのモデルは，動きの急な変化に対応し，したがって，以前のモデルよりも優れた性能を発揮した.

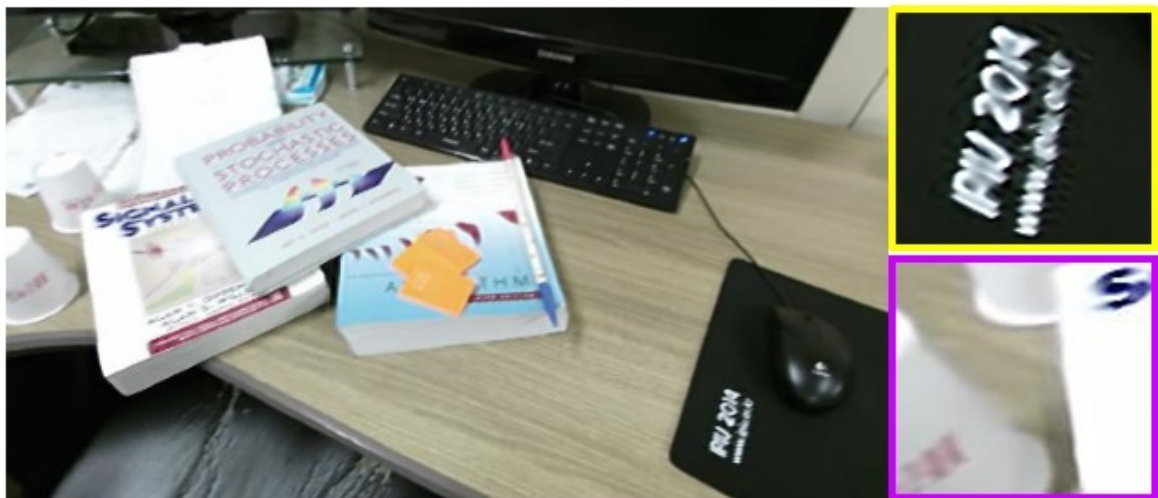
さらに，式 (2.19) の時間コヒーレンス項を用いずに，ブレを除去し，時間的コヒーレンスモデルが図 2.22 の辺縁付近のリングイングアーチファクトを著しく減少させることを検証した. 多数の実



(a)



(b)



(c)

図 2.22 (a) 実測ビデオのブレのあるフレーム (b) $E_{temporal}$ 項を用いないブレ除去結果 (c) $E_{temporal}$ 項を用いるブレ除去結果

ビデオからの他のブレ除去の結果が図 2.23 に示されている。特に、ここでのモデルは、人が回転移動するために非常に不均一なブラーを有した顔を正常に復元した（図 2.23 (e)）。

2.6.6 まとめ

ここでは、従来の手法では実現できなかった動的シーンの一般的なブレを除去した新しい手法を説明した。オプティカルフローを使用してピクセル単位のカーネルを推定することによって、一般的なブレを処理した。したがって、ここでは、オプティカルフローとブレの無いフレームを推定した新しいエネルギーモデルを提案した。

また、エネルギー関数を最小限に抑えるためのフレームワークと効率的なソルバーを提供し、動的シーンの一般的なブレを除去する際に大幅な改善を達成した。

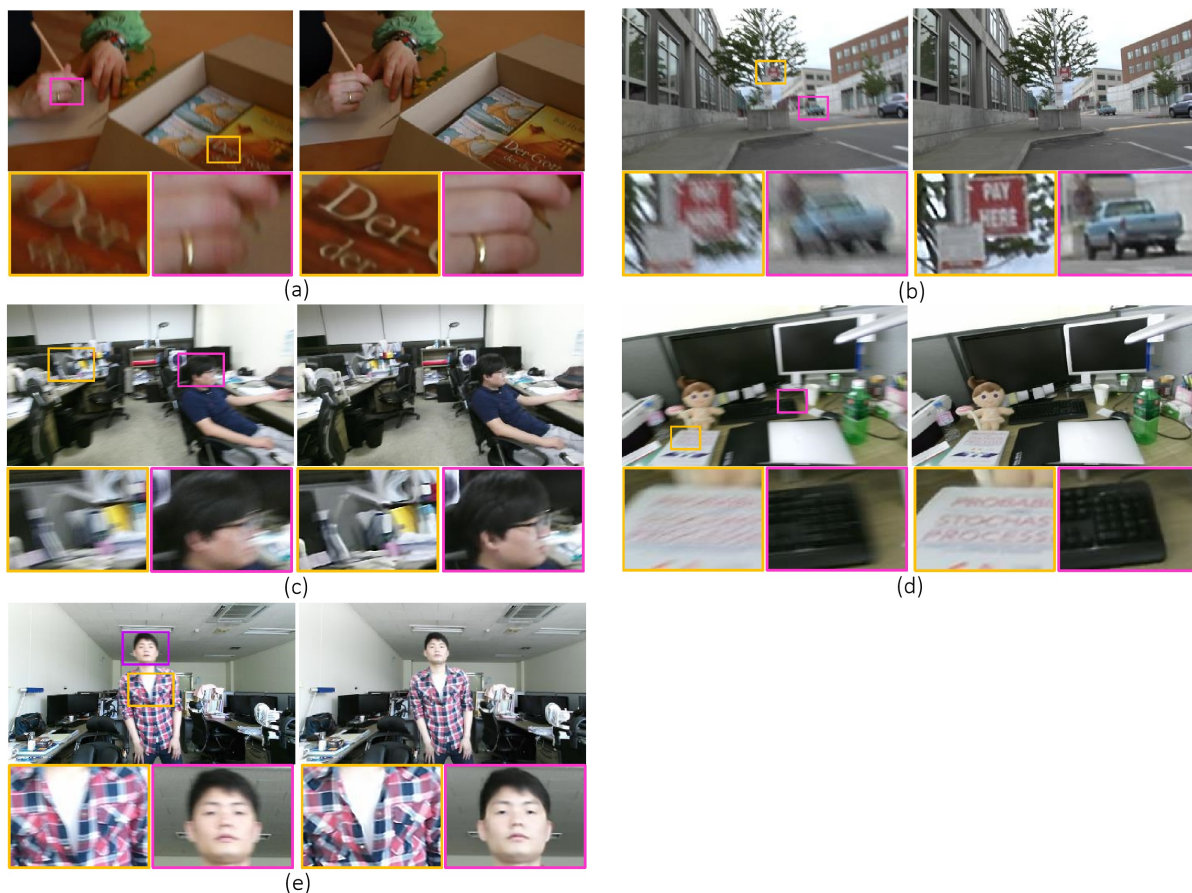


図 2.23 左から：ブレのあるフレームブレ除去結果。 (a)-(b)Cho らの手法で使われたデータ (c)-(e) キネクトで撮影された RGB 動画

2.7 関連研究まとめ

本章では本論文の関連研究について説明した。超解像処理によって低解像度の観測画像から高解像度の画像を生成できる。超解像処理ではシーン中の物体の動きブレが生成画像を劣化させる原因及び処理そのものを不可能にする原因になる。ブレ除去はシーン中のブラーカーネルを推定し、観測画像をブラーカーネルで逆畳み込み処理またはそれに準ずる処理を行うことでブレの無い画像を生成する。撮影に使用するカメラが回転するような動きをする場合や撮影するシーンが奥行の違いや独立して動く物体によって複数の平面を持つ場合は領域ごとにブラーカーネルを変化させる必要がある。

近年、動画像を対象とすることで動物体がブレているシーンに対応しようとする動きがある。InSoら [25] は固定カメラシーンにおいてフレームごとに開口パターンを変化させる撮影方法を用いることで動物体シーンのブレ除去を実現している。この手法ではフレーム間の差分によって背景と動物体部分のセグメンテーションを行う。フレームごとに開口パターンを変化させる撮影方法は、ブレ除去のノイズを小さくする他にフレームごとのブラーカーネルを変化させる要因にもなるため、この手法を一般的なカメラに置き換えた場合ブレ除去の精度は低下する。Kyongら [28] は一般的なカメラでの動物体シーンのブレを除去するために領域ごとのオプティカルフローによって前景領域と背景領域を分けた。オクリュージョン部分ではオプティカルフローの推定が困難であるため、オクリュージョン部分を検出しオクリュージョン部分らしさによる重みづけを行っている。そのため、物体の境界部分の平滑化が大きくなり、複雑な形状を持つシーンでのブレ除去も困難である。

このように一般的なカメラで撮影されたシーン中に動きブレがある物体シーンの動画の画像処理は依然として困難である。本論文はブレを含む動物体があるシーンの画像処理をより高品質にするために観測動画からブレの無い状態の前景の形状を正確に推測し、オクリュージョン部分周辺のノイズを低減する画像処理手法を提案する。

第 3 章

処理手法の概要

複数の移動物体が含まれるシーンや、奥行き (depth) の異なる複数の 3 次元物体が含まるシーンにおいては、ブラー除去や超解像処理の前に、シーンを動きや奥行きごとに分割を行う必要がある。分割した領域ごとにブラー除去と超解像処理を行い、最後に結果を結合することで、画像全体の画像処理が実現される。

まず、本手法で用いるシーンのモデルについて述べる。一般的に画像復元の論文では以下のように劣化画像を潜在画像 L と劣化行列 K の畳み込みによってモデル化する。

$$\text{deterioratedImage} = K * L \quad (3.1)$$

劣化行列 K は超解像の問題の場合はサブサンプリング、ブラー除去の場合はブラーを表す点広がり関数 (ブラーカーネル) 等を示す。また、前景と背景からなるシーンは以下のようにモデル化される。

$$\text{twoLayerImage} = F(\alpha) \circ F(c) + (1 - F(\alpha)) \circ B \quad (3.2)$$

ここでは前景のテクスチャを $F(c)$ 、アルファマップを $F(\alpha)$ 、背景を B としている。 \circ は同じサイズの行列に対して成分ごとに積を取る処理 (アダマール積) を表す。この式は $F(c)$ と B を $F(\alpha)$ による係数で重みづけした和である。また、アルファマップとはその物体の透明度を表し、非透明部分を 1、透明部分を 0 とする。アルファマップの値は物体の素材やブラーによる残像によって少数の値を持つ場合もある。式 3.1 と式 3.2 から劣化した前景と背景からなるシーンのモデルは

$$I = K_F * (F(\alpha) \circ F(c)) + (1 - K_F * F(\alpha)) \circ B \quad (r, g, b \in c) \quad (3.3)$$

のようになる。ここで、 I は観測画像 K_F は前景の劣化行列を表す。このときこのモデルを複数フレームの場合に拡張させると以下ようになる。

$$I_i = K_{F_i} * (F(\alpha) \circ F(c)) + (1 - K_{F_i} * F(\alpha)) \circ B_i \quad (r, g, b \in c) \quad (3.4)$$

i はフレームの番号を示す。このとき、潜在画像のテクスチャ $F(c)$ 及びアルファマップ $F(\alpha)$ はフレーム間で不変であると仮定する。また、劣化行列 K_{F_i} は単数フレームでの劣化行列と異なりサ

ブサンプリングやブローカーネルの他に各フレームにおける位置の情報を含むものとしている。図 3.1 は式 3.4 の概略図である。本研究ではシーンのブレの無い潜在的な前景のテクスチャ $F(c)$ を得ることが最終目的である。また、式 3.4 において劣化行列 K_{F_i} 背景画像 B_i 潜在的な前景のアルファマップ $F(\alpha)$ はシーン取得段階では未知である。

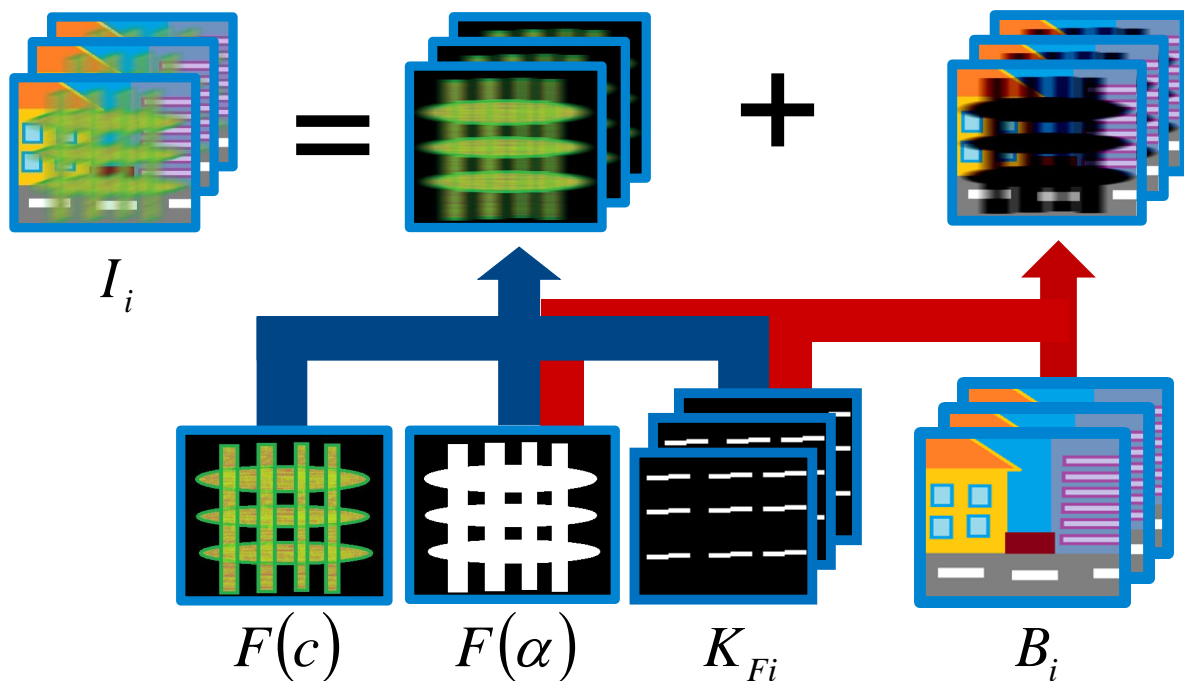


図 3.1 モデルの説明

提案手法では、図 3.2 のようにそれらの未知の要素をそれぞれ導出し、前景のテクスチャ $F(c)$ 以外の要素を求めることで前景のテクスチャ $F(c)$ の導出を可能にする。

I では、特徴点を抽出し、同じ平面上にある特徴点のクラスタの抽出と平面領域の推定を行う。モーションブローのある場合、長いシーンでの特徴点のトラッキングは困難であるため、短いシーケンスに分けて処理を行い、その後、シーケンスごとにそれらの結果を統合する。このとき、統合を繰り返すにより誤差が蓄積されてしまうため、統合した結果を用いて再度平面領域の推定を行うことで、蓄積された誤差を軽減する (図 3.3 K_{F_i})。また、背景のトラッキング結果から各フレームにおける背景画像を導出する (図 3.3 B_i)。その後、観測画像と背景画像から前景画像の抽出と前景のアルファマップを推定する (図 3.3 中段)。II では、推定した平面の動きを用いて、ピクセルごとに、動

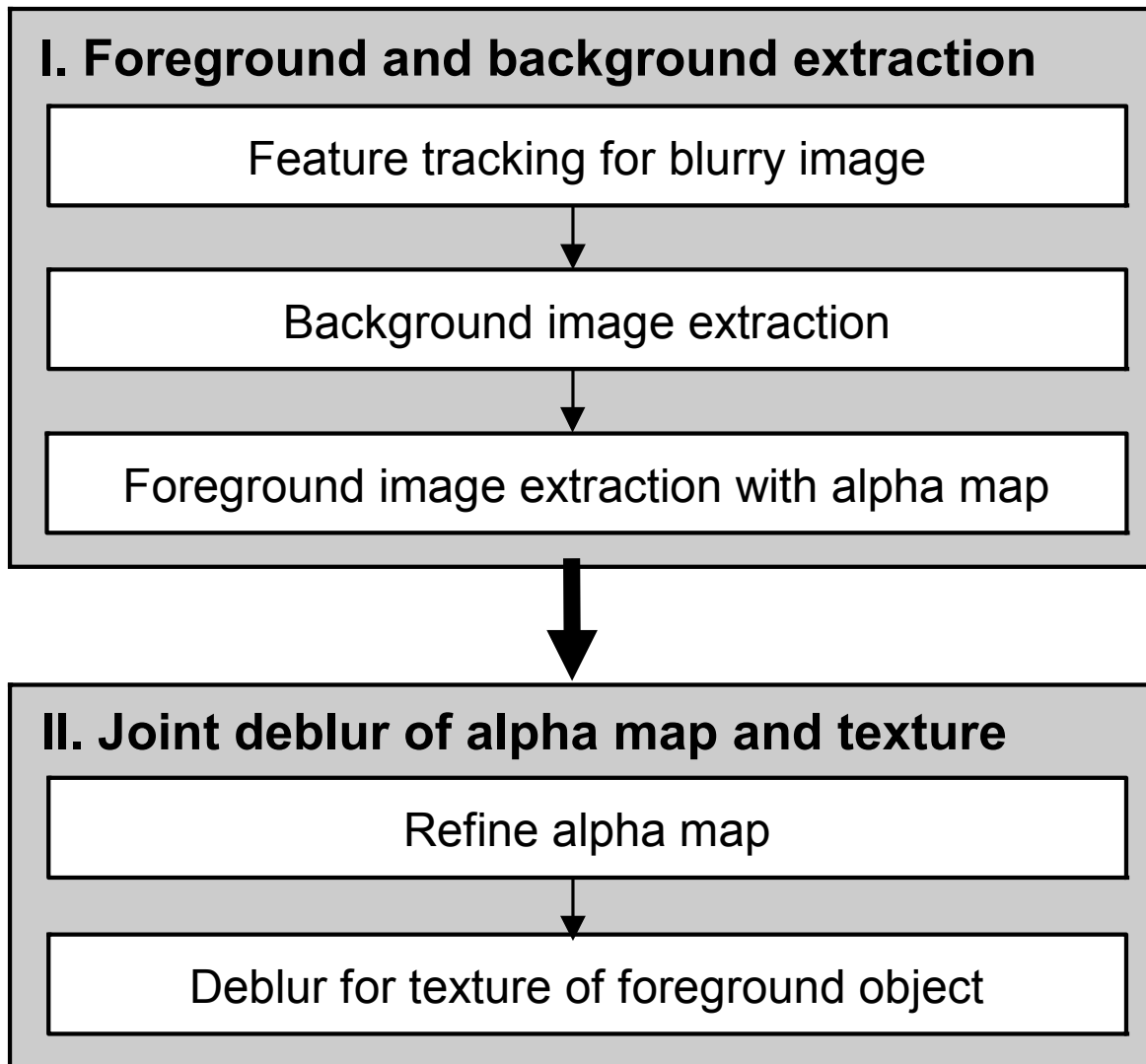


図 3.2 全体の処理のフローチャート

きベクトルを計算し, アルファマップの高精度化を行う.(図 3.3 $F(\alpha)$) その後, 前景部分のブレ除去や超解像を行う.(図 3.3 $F(c)$) 以下のセクションで具体的な手法について述べる

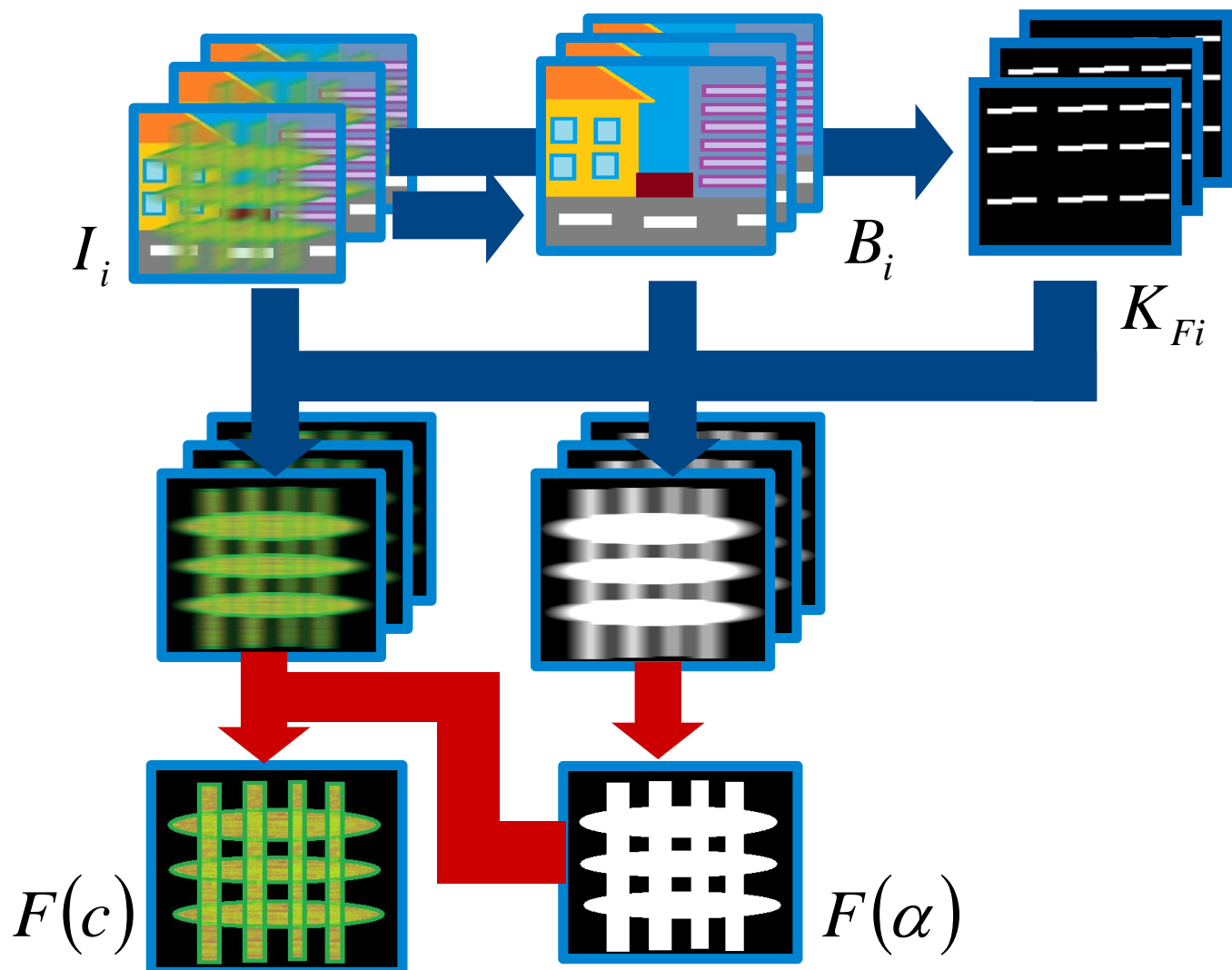


図 3.3 全体の処理の流れ. 青矢印部が図 3.2 における I に, 赤矢印部が図 3.2 における II に該当する.

第 4 章

ブレのある動画像からのブレ除去と超解像の同時処理

今日ではビデオカメラはあらゆるところで一般的に使用されており、ビデオシーケンスから一枚の画像を取り出す需要は増えている。ビデオカメラの解像度は通常デジタルカメラのそれよりも低く、ビデオシーケンスから単純にフレームを切り出すだけでは品質が低い。さらに、ビデオカメラには高い FPS が必要とされているため、ビデオカメラには必然的にモーションブラーが含まれフレーム間の位置合わせに誤差を生じさせてしまう。これは、複数フレームの超解像にとっては重大な問題である。

この章では、モーションブラーを含むビデオシーケンスから高解像度の画像を復元する方法について説明する。ビデオカメラのフレームレートが高いため、連続するフレームにおける物体の動くは小さく、短い区間のシーケンスではブレがあっても特徴点のトラッキングが可能である。そのため、我々は長い区間のシーケンスでのトラッキングを安定化させるために分割-統合アプローチを採用する。また、MAP 推定に基づく複数フレームを用いたブレ除去と超解像を同時処理する方法を提案する。実験結果ではこの手法の優位性を証明する。

4.1 はじめに

監視や個人識別などの目的でビデオカメラの動画からクオリティの高い一枚の静止画を取り出したいことがある。しかし、動画から単純にフレームを切り出しただけの静止画では、目的を満たすには不十分であることが多い。これは、一般に、動画は静止画に比べて低解像度であることや、撮影されている物体の動きや、カメラのブレなどによるモーションブラーの発生などが要因である。これらの問題を解決するためブレ除去や超解像処理が研究されている、一般に、これらの手法では、フレーム全体のカメラの動きを推定し、位置合わせすることでブレ除去や画像の超解像を実現してい

る [87, 88, 89]. しかし, シーン内に異なる奥行き の物体や, 独立した動きがある場合, フレーム内に方向や大きさの異なるモーションブレンダーが発生し, 一つの動きだけを仮定するような過去の一般的な手法を適用することが困難になる. さらに, 超解像処理についてはモーションブレンダーがあることで画像の位置合わせに誤差が生じるだけでなく, モーションブレンダーがノイズとして扱われることで超解像処理にも悪影響を及ぼす [2].

この問題を解決するために, 過去にいくつかの方法が提案されている. 例えば, 画像を小さなパッチに分割して処理をしたり, シーンを構成する形状を平面ごとに分割し, その分割した領域ごとに処理をする手法が提案されている [90]. しかし, これらの手法は複数枚の連続するフレーム間において, モーションブレンダーがある状態で, トラッキングを行うことは依然として困難である. そこで本研究では, 連続する長いシーケンスを, 短いシーケンスごとに分割し, 各短いシーケンス毎で計算し, 最後にそれらの結果を統合することで, 長いシーケンスでの処理を実現する手法を提案する, ここでの提案手法の貢献は以下のとおりである,

1. ピクセルごとのブレンダーカーネルの推定手法の提案
2. ブレのある連続画像での高精度トラッキングの実現
3. 複数フレーム画像を入力とする, ブレ除去と超解像処理の同時実現

これにより, 低解像度でブレのあるビデオを入力として高精度な静止画を作成できる.

4.2 関連研究

画像のモーションブレンダーは, ブレンダーのない画像に, ブレンダーの原因となる動きやブレを表すカーネルを畳み込むことでモデル化される. 従って, ブレ除去 (deblur) は, 一般に, LR 法 [91, 92, 12] などの既知のカーネルを用いた逆畳み込み処理によって実現される. また, 近年ブラインドデコンボリューション技術も盛んに研究されている [88, 93]. どちらの研究でも単一のカーネルを想定しているため, 動きや奥行き の異なる複数の物体がシーンに含まれている場合適用できない. 近年では, シーンをセグメンテーションし, セグメントごとに独立したカーネルを推定し, ブレ除去することで, 領域ごとに被写体とカメラの距離が異なるシーンや複数の物体が独立して動くシーンに対応する方法がいくつか提案されている [94, 89]. しかし, これらの方法は単一の画像に対するものであるため, 画像シーケンスに対する一貫性のあるブレ除去は考慮されない. これは, モーションブレンダーがカメラでは深刻に発生するが, ビデオシーケンスでは発生しないからである. 動画についてはいくつかの研究がおこなわれている [95, 96] が, フレームごとに単一のブレンダーカーネルしか考慮しておらず, 私たちのケースには適用できない.

超解像像に関しては, 単画像のみを使う手法 [5] と複数画像を使う手法 [4] がある. 複数画像の方法

は、画質や安定性の点で優れていると考えられている。また、高解像度のパッチと低解像度のパッチの画像データセットを用いた学習ベースの方法もある [97]。複数画像を用いる超解像で重要となるのは各フレームを基準フレームにサブピクセル精度で位置合わせである [87]。正確な位置合わせを実現するためには特徴点を正確にロバストに追跡する必要がある。一般的な特徴抽出 [98, 99] および特徴点トラッキングの手法 [32] の大部分は画像勾配に基づいているため、モーションブラー等によって画像の高周波成分が失われると失敗する。

ビデオシーケンスの超解像に関する別の問題は、シーン中に複数の動く物体が存在する場合である。これまでその問題に対応した方法が提案されているが [100, 90]、モーションブラーを校正していないため手持ちビデオカメラに適用することは困難である。

4.3 提案手法

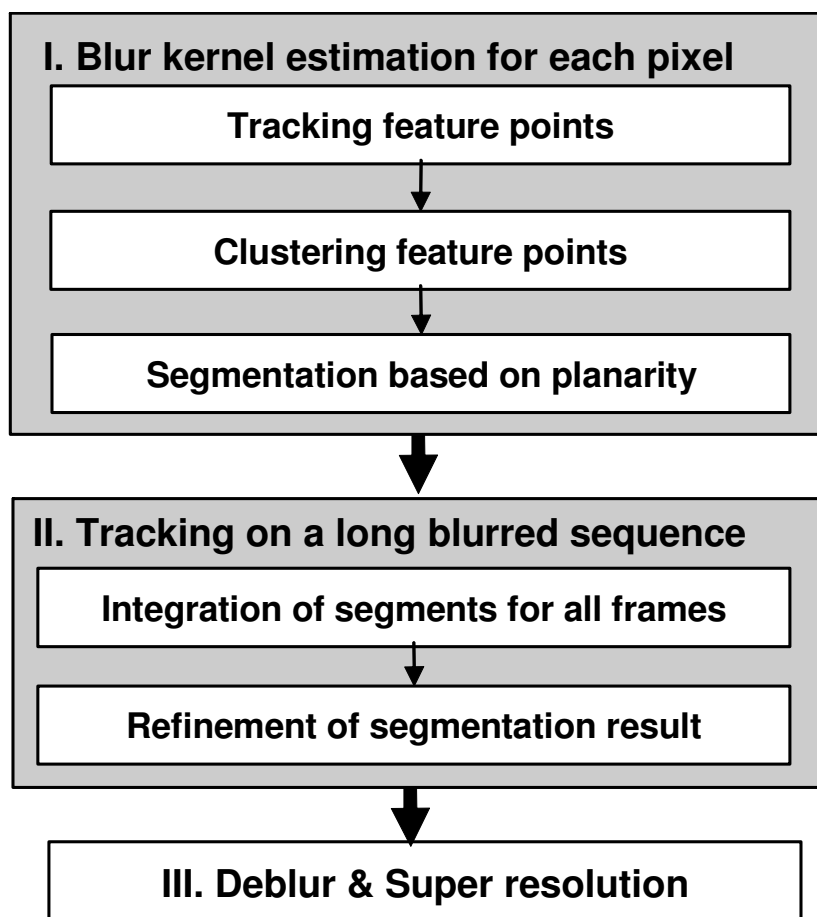


図 4.1 アルゴリズム概要

複数の移動物体が含まれるシーンや、奥行き (depth) の異なる複数の 3 次元物体が含まれるシーン

においては、ブレ除去や超解像処理の前に、シーンを動きごとに分割を行う必要がある。分割した領域ごとにブレ除去と超解像処理を行い、最後に結果を結合することで、画像全体の超解像が実現される。以上を、提案手法では、大きく次の4つの処理で実現する(図4.3)。

I (Sec.4.4.1) では、特徴点を抽出し、同じ平面上にある特徴点のクラスタの抽出、クラスタごとのセグメンテーションを行う。モーションブラーのある場合、長いシーンでの処理は困難であるため、短いシーケンスに分けて処理を行う。II (Sec.4.4.2) では、シーケンスごとに抽出した平面を統合する。統合により誤差が蓄積されてしまうため、統合した結果を初期として、繰り返し処理により蓄積された誤差を軽減する。最後に III (Sec.4.4.3) では、推定した平面の動きを用いて、ピクセルごとに、動きベクトルを計算し、ブレ除去と超解像を同時に行う。以下のセクションで具体的な手法について述べる。

4.4 詳しい方法

4.4.1 ピクセルごとのブレ推定

モーションブラーの除去と超解像を複数のモーションと奥行きがあるビデオデータに適用するためにはシーン中の領域毎にブラーカーネルを推定する必要がある。

ここでは、まずシーン内のオプティカルフローを計算し、それを用いて各ピクセルのブラーカーネルを推定する。オプティカルフローの計算は山口の方法を用いる [90]。この方法は、**1)** 特徴点のトラッキング、**2)** 特徴点を同じホモグラフィ行列で表現できるものでクラスタリング、**3)** 各ピクセルを再投影誤差が最も小さくなるクラスタを割り当てることによるセグメンテーションの3つのステップからなる。このとき、各セグメントに対応するホモグラフィ行列から隣接するフレームの対応ピクセルが分かるため、各ピクセルのオプティカルフローを計算することができる。カメラのシャッタースピードは一般的に短いため、ブラーカーネルの計上は線形であると仮定する。この段階では、ブラーカーネルの方向と長さを求めることができるがカーネルのスケーリングのパラメーターは求めることができないので後で推定する。

4.4.2 ブレのある画像のトラッキング

分割-統合アプローチ

シーン中の動きが激しいとき、モーションブラーや、見え方の変化が大きくなり、トラッキングの誤差が増大するため、フレームを経るごとにトラッキングが困難となる。本手法ではこの問題を解決するため、長いシーケンスを複数の短いシーケンスに分割し、各シーケンスごとに平面領域を抽出し、最後にそれらを統合する方法を提案する。

各シーケンスで得られる平面領域はシーケンスごとに異なる場合が多く、単純に統合することは容易ではない。特に、非剛体の物体の場合や、一つの平面が途中で2つに分かれる場合、元々2つの物体が動き方の関係で一つの物体だと扱われるため、シーケンスの前後で平面領域の数が異なってしまい、各領域を1対1に対応させる方法を用いることはできない。

この問題を解決するために、入力する動画を短いシーケンスに分割するときに、シーケンスの末尾と先端が重複するようにする。次にシーケンスの前後で重複するフレームにおける平面領域の分割結果を比較し、同じピクセルがシーケンスの前後で矛盾しないように、平面領域の再分割を行いシーケンスを統合する(図4.2)。これはシーケンス全体に伝播される。このとき、オクリュージョンやノイズが原因で統合後にどのピクセルからも射影されない領域や、複数のセグメントに属するピクセルから射影される領域が存在することで割り当てることができない領域が発生する。それらのピクセルは不確定とし、周囲の確定したセグメンテーション結果から最も近い位置の値を与える。また幾つものシーケンスの統合を続けると、誤差が蓄積したり、不必要なセグメントが発生したりする問題があるため、次の精度向上手法で解決する。

特徴点トラッキング及びセグメンテーションの高精度化

シーン内の平面構成を表すセグメンテーションの精度は、主にモーションブラーによって引き起こされるトラッキングエラーのために減少する。

この問題を解決するために、特徴点の追跡結果を高精度化する。しかし、ブレの大きさや方向はフレーム毎に異なるため、一般的な方法で精度を向上させることは難しい。我々の方法では、モーションブラーのある画像であっても高精度の位置合わせを行うブレ不変位相限定相関マッチング法(BIPOC)を用いる[101]。しかしながら、BIPOCは原理的に平行運動にしか適用できない。したがって、ホモグラフィ変換によって表現される自由な動きに対応するためには、単にBIPOCを適用するだけで高精度の登録を行うことは困難である。

この問題を克服するために、我々は以下のような3つのステップの反復アプローチを実装する。**1)** 位置合わせの基準となるフレームに位置合わせする様に各フレームをホモグラフィ変換する **2)** BIPOCによるトラッキング精度の向上、**3)** 第2ステップで得られた結果を用いたホモグラフィ変換の更新。精密化ステップを繰り返し行うことにより、高精度の画像位置合わせが達成される。

4.4.3 マップ推定を用いた画像復元

複数枚を用いる超解像手法では、求めたい超解像度画像に対して、サブピクセルの位置ずれや、ピクセル量子化(pixelization)による画質劣化をシミュレートし、これが入力画像と一致するようにMAP推定することが多い。この場合、使用画像にあるモーションブラーはノイズとして扱われる

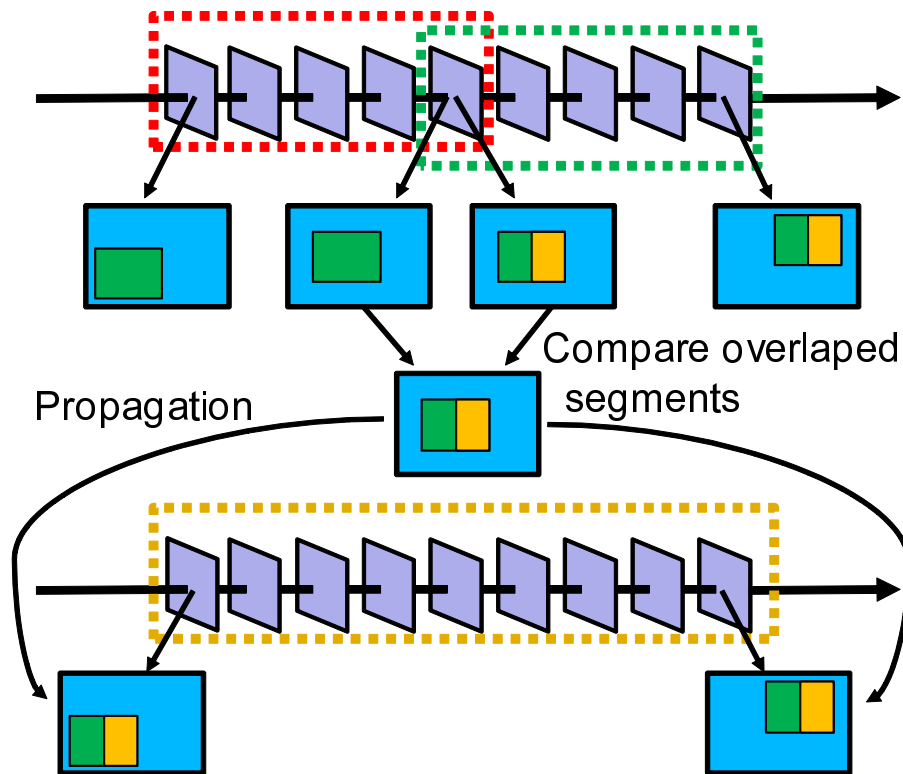


図 4.2 平面候補推定結果の統合.

ため画質に悪影響がある [2].

一方一般的なブレ除去は,1 枚の画像を入力とし, ブラーカーネルを用いたデコンボリューション処理により実現する [92, 91]. このためモーションブラーが大きいとリングングによる画質劣化が大きくなるなどの問題がある.

本論文では, 画質劣化過程に, サブピクセルの位置ズレと, ピクセル量子化に加えて, モーションブラーも考慮することで, MAP 法による, ブレ除去および超解像を同時に実現する手法を提案する. 同時に推定することにより, 超解像におけるモーションブラーによる画質劣化や, ブレ除去におけるリングングが解消されることが期待される. x を高解像度画像ベクトル, y を入力画像ベクトル, N を入力フレーム数, P を高精度画像のピクセル数, s をブラーカーネルのスケールパラメータとすると, 画像復元は以下のような式で表現できる.

$$\arg \min_{\mathbf{x}, s} \left\{ \sum_{k=1}^N \sum_{i=1}^P \|\mathbf{y}_{ki} - \mathbf{D}B_k(s, i)W_k \mathbf{x}_i\|^1 + \lambda \|\mathbf{C}\mathbf{x}\|^2 \right\},$$

ここで, D はサブサンプリングを表す行列, $B_k(s, i)$ はピクセルごとのブラーカーネルを表す行列に値する, W_k はシーンやカメラの動きを表す行列を意味しており, $\|\mathbf{y}_{ki} - \mathbf{D}B_{ki}W_{ki}\mathbf{x}_i\|^1$ は高解像度画像から推測される低解像度画像と実際の観測画像との L1 ノルムである. また, C は高解像度画像の事前情報を表す行列であり, λ は事前情報の寄与度合いを表すパラメータを意味している. 本手法では, 平滑化項としてバイラテラルフィルタを用いる, 式を最急降下法を用いて解く. 実際の処理では, まず x を固定して s を推定し, s を固定して x を推定し, 収束するまでプロセスを繰り返す.

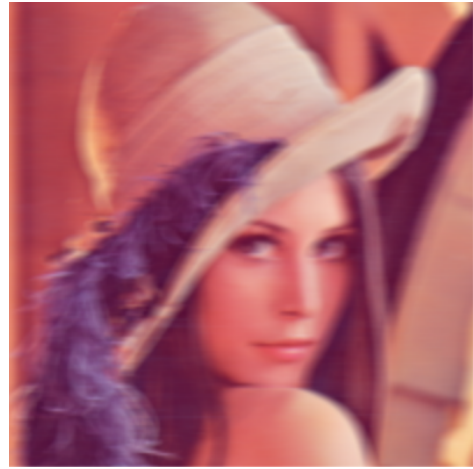
4.5 実験結果

まず, 我々の方法を評価するためにシミュレーション画像を使用する. 我々は, 図 4.3(a) に示すように, 高解像度画像のために 400x400 ピクセル画像を使用する. 次に, ランダムな方向に 17-19 ピクセルの長さの平行移動をシミュレートし, 固定スケール 0.7 の方向に沿ってモーションブラーを適用する. 次に, 図 4.3(b) に示すように, 画像を 200×200 画素にダウンサンプリングした. このようなイメージセットを用いてに 2x2 倍の超解像を行った. 我々の方法では, スケールパラメータは 0.71 と推定され, これは MAP 推定に十分な精度を有する. 図. 4.3(c)-(f) に示すように, 入力枚数が増えるほど, リングングの減少が確認でき, PSNR の向上も確認できる.

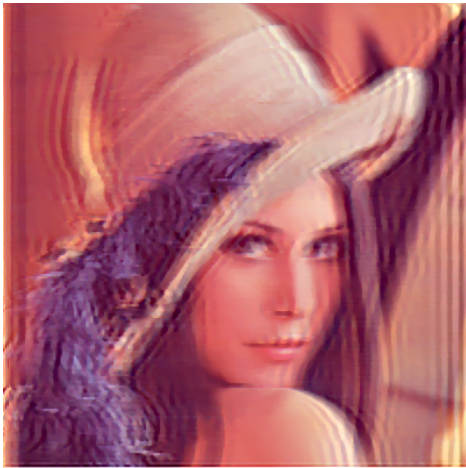
次に, 図 4.4(a) のようなカメラを手を持った状態で縦横方向に自由に動く台に新聞紙を載せて撮影し, 実験を行った. 使用したフレーム数は 30 枚 (350x265 pixel) で 2x2 倍の超解像を行った. 新聞紙はそれぞれ異なる回転と並進運動をしていたため, 各フレームで異なるブラーが生じていた. 図. 4.4(b)-(d) に示すように画像処理の結果を異なるアルゴリズムで比較した. ブレ除去のみを行うもの, ブレ除去したものをを用いて超解像したもの, ブレ除去と超解像を同時に行うものを比較する



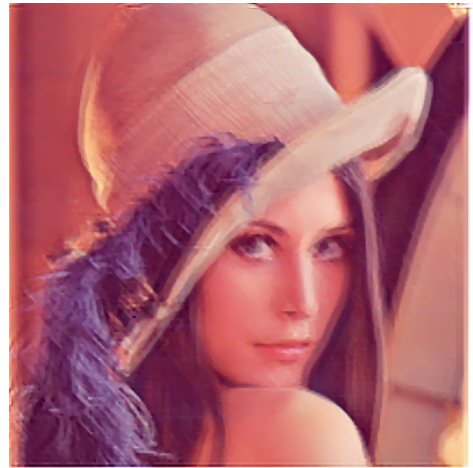
(a) Original (high reso.)



(b) Input (PSNR=23.1dB)



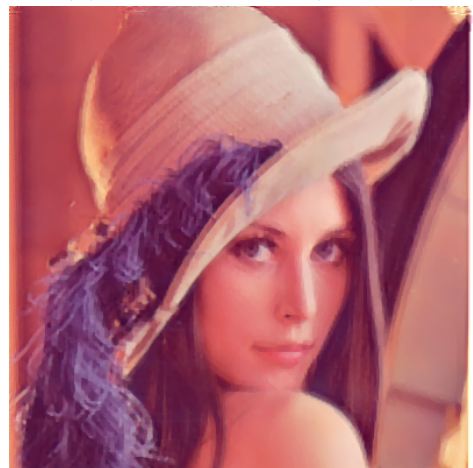
(c) Using 1 frames (22.4dB)



(d) Using 2 frames (23.5dB)



(e) Using 5 frames (24.2dB)



(f) Using 10 frames (24.7dB)

図 4.3 シミュレーション画像によるブレ除去結果.



(a) Input



(b) Deblurring only



(c) SR only



(d) SR + deblur (proposed)

図 4.4 実測画像による結果

と、我々の手法が優れていることが確認できる。最後に、複数の人物が本を持ちながら動くシーンを、固定したビデオで撮影したシーンを入力として、提案手法による超解像を行った。図 4.5(a)-(c)のように使用したフレーム数は 21 枚で、 420×160 ピクセルの画像から 2x2 倍の超解像を行った。シーケンスにはいくつかのモーションが存在するため、短いシーケンスのセグメンテーション結果は異なるが、これらは図 4.6 に示すように一貫して統合されている。

結果からは、超解像とブレ除去を同時に行う手法が、その他の手法の中で一番優れていることが確

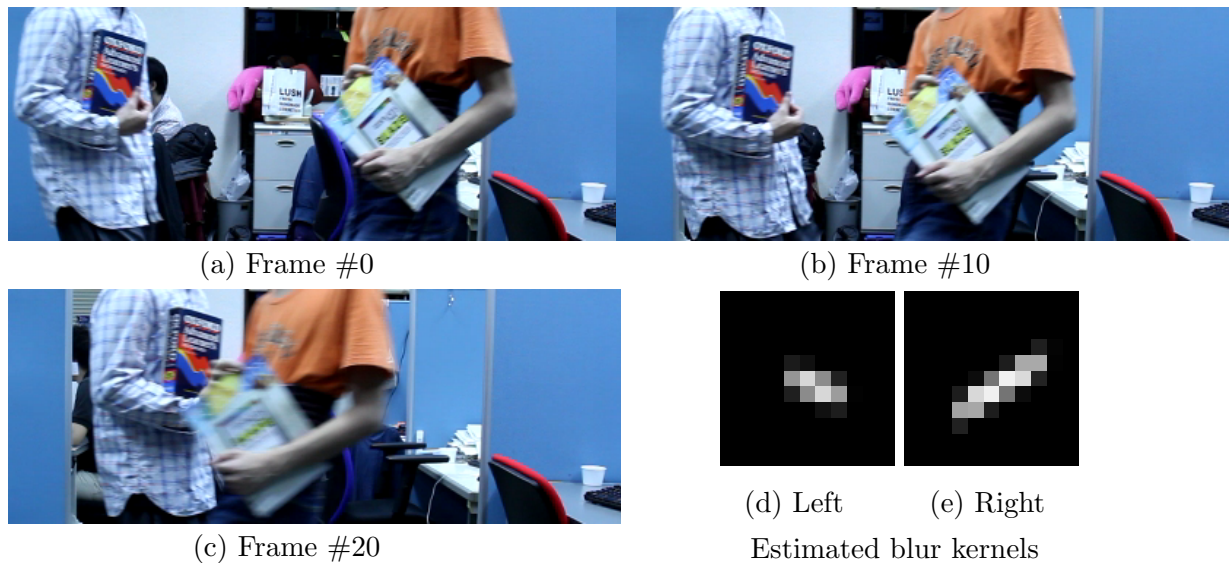


図 4.5 (a)-(c) 入力画像 (d)-(e) フレーム#0 での本の部分で推定されたブローカーネルの一例.

認できる. また, 超解像とブレ除去を同時に行う手法の結果を入力枚数で比較したものを図 4.7, 4.8 に示す.

4.6 まとめ

この章では, 複数の被写体に, 異なるモーションブラーがあるシーンに対応した ブレ除去 および 超解像手法を提案した. 提案手法では, 入力する連続画像を短いシーケンスに分割し, 平面領域の 計算を行い, それらを統合することで長いフレームでの平面トラッキングを行った. また, モーションブラーを考慮した位相相関法 (POC) によるトラッキング結果の高精度化を提案した. さらに, 複数枚を入力として, deblur と超解像を同時に行う手法を提案した. 実験により, 提案手法により高精度な超解像を実現できることを示した.

この処理は, 分割するシーケンスのフレーム数や, ブローカーネルのスケーリング変数など処理結果に大きくかかわるパラメーターがある. これらは, 経験的に設定していたが処理中に望ましい値を自動計算させるようにするのが適切と思われる. 特に, ブローカーネルの推定については高精度化画像の推定と同時に行う手法も多い. また, シーケンスごとの結果の統合については, 重複フレーム以外のセグメンテーション結果や入力画像の輝度情報など, 事前情報を有効活用できていない実装である. 今後は, フレーム間のモーションから, ブローカーネルのスケーリング変数の自動計算や, より高精度の特徴点トラッキング手法, セグメンテーション手法を組み込む予定である.

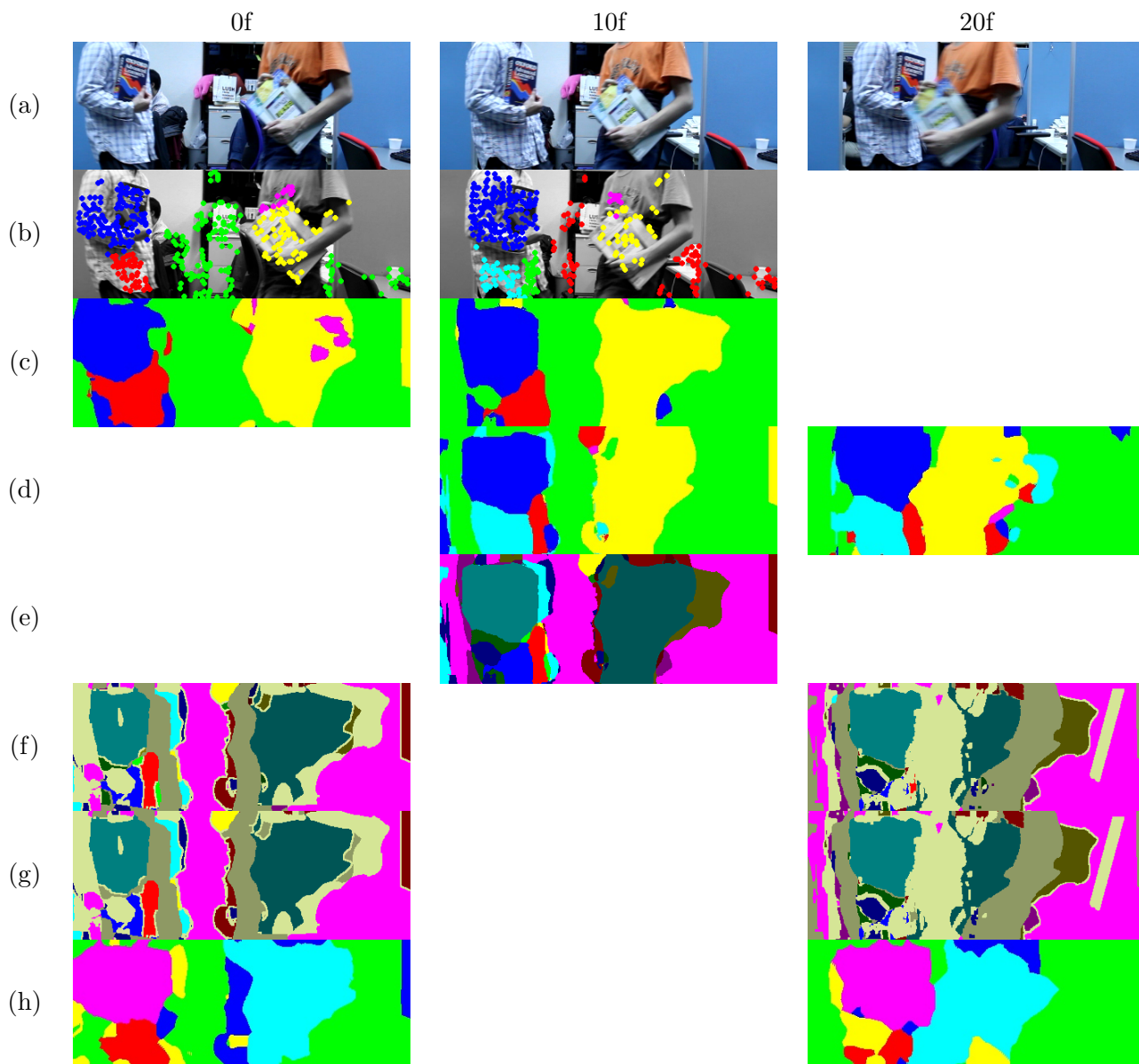
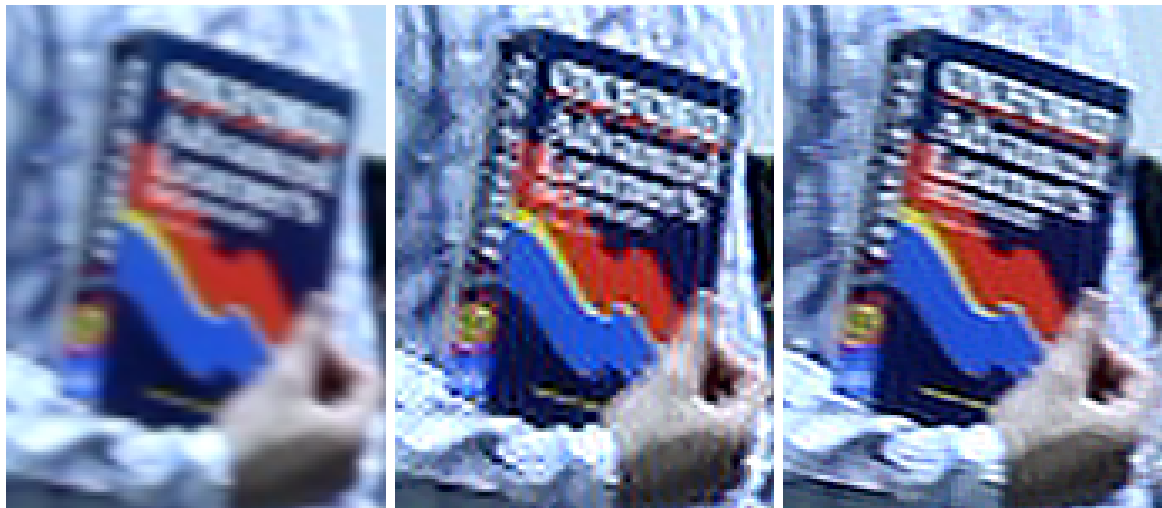


図 4.6 実測画像による平面統合の結果.(a) 入力画像.(b) 特徴点クラスタリング結果.(c)0f~10fのシーケンスでのセグメンテーション結果.(d)0f~10fのシーケンスでのセグメンテーション結果. (e) 重複フレームの重ね合わせ結果.(f)(e)の結果を射影したもの.(g)(f)から不要セグメントをリジェクトしたもの.(h) 統合後のセグメンテーション結果



入力画像

1 枚使用

5 枚使用



10 枚使用

20 枚使用

図 4.7 左側の本のセグメントを高精度化処理したもの (枚数による比較)



入力画像

1 枚使用

5 枚使用



10 枚使用



20 枚使用

図 4.8 右側の本のセグメントを高精度化処理したもの (枚数による比較)

第5章

オクリュージョンを考慮したブレ除去

動画のデータから高品質の画像を復元する技術への要求は高まっている。走行している車や歩行者など、屋外シーンは複数の奥行きや動きを持つ複数のシーンを含むため単にブレ除去の手法を各フレームに用いるだけでは品質の悪い処理結果になる。さらに、前景物体と背景の混色がオクリュージョン部分の境界で発生するため、物体を復元するためには前景物体と背景の色を分割する手法がブレ除去の工程で必要になる。

本手法では、ビデオカメラで撮られたブレのある複数フレームの画像から前景物体のオリジナルの境界を復元する手法を提案する。複数フレームを用いたブレ除去の技術を利用した前景と背景の混色の分解は物体の境界部分を高精度化することで境界線周辺のアーティファクトが低減、効果的に行われる。物体境界の高精度化とブレ除去はそれぞれ MAP 推定を用いた処理になっているため使用するフレームの枚数を増やすと復元画像の精度も向上する。実験結果では提案手法がモーションブラーとオクリュージョンの境界部分の混色があるようなシーンでのブレ除去に成功したことを示している。

5.1 はじめに

都市モデリング、デジタルアーカイブなどのいくつかの重要な目的のため、ビデオカメラを車両に搭載してシーンを撮影することに対する強い需要がある。実際の撮影プロセス中、シーンの一部は、車の動きのために遮蔽されたりブレがあることが一般的に発生する。以前は、これらのオブジェクトは一般的なアプリケーションの障害とみなされ、それらを削除する技術は広く研究、開発された [102, 103]。特に前景オブジェクトは、通常、背景と比較してカメラの近くに表示され、モーションブラーによって頻繁に劣化する。さらに、オクリュージョン境界付近のぼけは、前景と背景の混色を伴い、ひどく劣化する。

高解像度でぼけのない画像をビデオデータから復元するために、多数の解決策が提案されてい

る [89]. 一般に, これらの方法は, フレーム全体に存在する均一なブラーカーネルまたはレジストレーションパラメータを推定することによってブレ除去を行う. しかしながら, この技術は, 異なる深度及び独立した動きを有するいくつかのオブジェクトを含むシーンには適用できない.

この問題を克服するために, 例えば画像をシーンの平面部分を表す小さなセグメントに分割するか [104], または各画素毎に処理する [28] などのいくつかの方法が提案されている. しかし, 前景と背景のテクスチャが混在しているオブジェクト境界付近のオクルージョンは考慮されていないため, 前景のテクスチャを正しく復元することはできない.

本稿では, 動きのぼけと, 前景と背景の混色境界付近の混色を含む低品質ビデオデータから, 前景オブジェクトの境界形状とテクスチャの両方を復元する手法を提案する. この状況では, デプスの影響で背景は動かず前景は平面を仮定できる. さらに, 前景はフェンスのような一つの疎な平面物体と考えることができる. この技法は, アルファマップを使用するオブジェクト境界の精緻化アルゴリズムと, 複数のブラー画像を有する精緻化されたオブジェクト境界を使用する効率的なぼけ除去アルゴリズムの主に 2 つの部分からなる. この論文の貢献は以下の通りである.

1. 時空間解析を用いてビデオデータの各フレームの初期アルファマップ推定を行う.
2. 複数のノイズや誤差のあるアルファマップを用いたアルファマップ, 物体の境界情報の作成.
3. モーションブラーを有する前景オブジェクトと, オクリュージョン境界付近の背景オブジェクトとの混色を除去する技術.

5.2 関連研究

カメラの動きに起因するブレのある画像のブレ除去技術は, 畳み込み処理であるため, 既知のカーネルを用いたデコンボリューション手法として復元技術が提案されている [91, 92, 12]. 近年, ブラインドデコンボリューション技術も盛んに研究されている [93]. どちらのテクニックでも, 単一のカーネルを想定しているため, モーションや深度の異なる複数のオブジェクトがシーンに含まれている場合は適用できない. このような多重オブジェクトのケースに対処するために, シーンセグメンテーションを用いて独立してカーネルを推定する等のいくつかの方法が提案されている [94, 89]. ただし, これらの方法は単一の画像を対象としているため, 画像シーケンス (ビデオ) は考慮されない. これは, モーションブラーがカメラでは頻繁に発生するが, ビデオデータでは発生しないためである.

動画のブレ除去についてはいくつかの研究が行われているが [95, 96], フレームごとに単一のブラーカーネルしか考慮されておらず, 独立したモーションには適用できない. 近年, 映像上の複数の動く物体に対する動きブラー技術が提案されている [104, 28]. この手法は, 各レイヤーのレイヤー

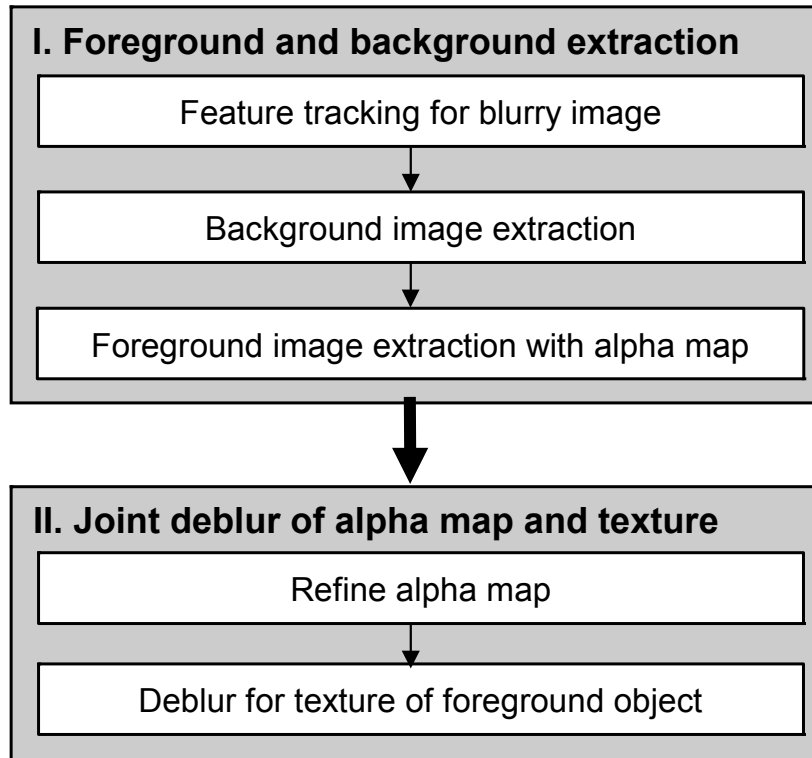


図 5.1 アルゴリズム概要.

分解と適応ブローカーネル推定に基づいている. しかしながら, これらの技術は, ブレのある前景オブジェクトのオクリュージョン境界における混色を考慮しておらず, オブジェクト形状またはテクスチャのいずれも正しく回復することができない. また, ブレのある画像から物体の境界とテクスチャの両方を回復する技術が提案されている [24, 25]. しかし, シャッタースピードを変えて各フレームをキャプチャする特殊なカメラが必要なため, 一般的なビデオデータには適用できない.

5.3 提案手法

5.3.1 アルゴリズム概要

モーションブローでお互いに重なっている 2 つの移動オブジェクトがある場合, 各オブジェクトのテクスチャは, オクリュージョン境界近くで一緒に混色される. この場合, 映像データの i 番目のフレームの撮像画像 I_i は, 以下の式で表すことができる.

$$I_i = \tilde{K}_{Fi} \circ B + K_{Fi} * (\alpha \circ F) \quad (5.1)$$

$$\tilde{K}_{Fi} = (1 - K_{Fi} * \alpha), \quad (5.2)$$

ここで, F および α は前景画像のカラーチャンネルおよびアルファチャンネルであり, B は背景画像であり, K_{Fi} はブラーカーネル, \circ は行列の同じ行と列のピクセル単位の乗算を表す.

このため, ブレの無いシーンを復元するには, 前景と背景の両方を独立して回復する必要がある. したがって, これら 2 つの層の分解が必要である. 分解の重要な手掛かりの 1 つは, 各フレーム上での前景オブジェクトの外観がオブジェクトの動きによるユニークなモーションブラーによって異なって見えていたとしても, 実際のオブジェクトの境界形状は一貫している点である. それを用いて, 本手法は MAP 推定による位置合わせに基づくものであり, 物体の動きに制約を受けることなく良好な結果が期待できる.

実際の方法を以下に示す (図.5.1 参照). ステップ I (第.5.3.2 節), では, 平面のセグメンテーションを画像シーケンスから検出された特徴点をクラスタリングすることで求める. 次に, シーンを背景と前景オブジェクトに分解し, 各フレームの初期アルファマップが作成される. 推定されたアルファマップには誤差があるため, ステップ II (第.5.3.3 節) では統計的な手法 (SSD 最小化) でアルファマップを高精度化する. ステップ II では, 洗練されたアルファマップを使用する前景オブジェクトのブレ除去が適用される.

5.3.2 前景と背景の分離

ブレのあるフレームでの特徴点トラッキング

モーションブラーの除去を複数の動きと奥行のあるビデオデータに適用するためには各領域ごとのブラーカーネルを推定する必要がある. ここでは, まずシーン内のオプティカルフローを計算し, フローを用いて各ピクセルのブラーカーネルを推定する.

ブラーのある入力画像の特徴点トラッキングを効率的かつ長いシーケンスで実現するために, Matsushita らが提案した方法を用いる [104]. この方法は, 4 つのステップからなる. **1)** 全フレームを, 少数のフレーム (たとえば 5 f) からなる複数のブロックに分割し, ブロックごとに特徴追跡を適用する **2)** 同じホモグラフィによって表現される特徴点をクラスタリングする, **3)** 各ピクセルごとに再投影誤差が最小であるクラスタを求めセグメンテーションを行う, **4)** 分割したブロックごとのセグメンテーションとホモグラフィ行列の統合.

対応するホモグラフィを有する複数のフレームを介して統合されたセグメントを使用することによって, ブレ除去に必要な隣接するフレーム間の各ピクセルの位置合わせ行列および動きベクトルを計算することができる. このようにして求められる統合されたシーケンスの各セグメントに対応するホモグラフィ行列を用いて, ブレ除去に必要な隣接するフレーム間の各ピクセルの位置合わせ情報および動きベクトルを計算することができる. ここで, ブラーカーネルの方向と長さは計算できが, ブラーカーネルのスケーリングパラメータは決めることができない事に気を付けたい.

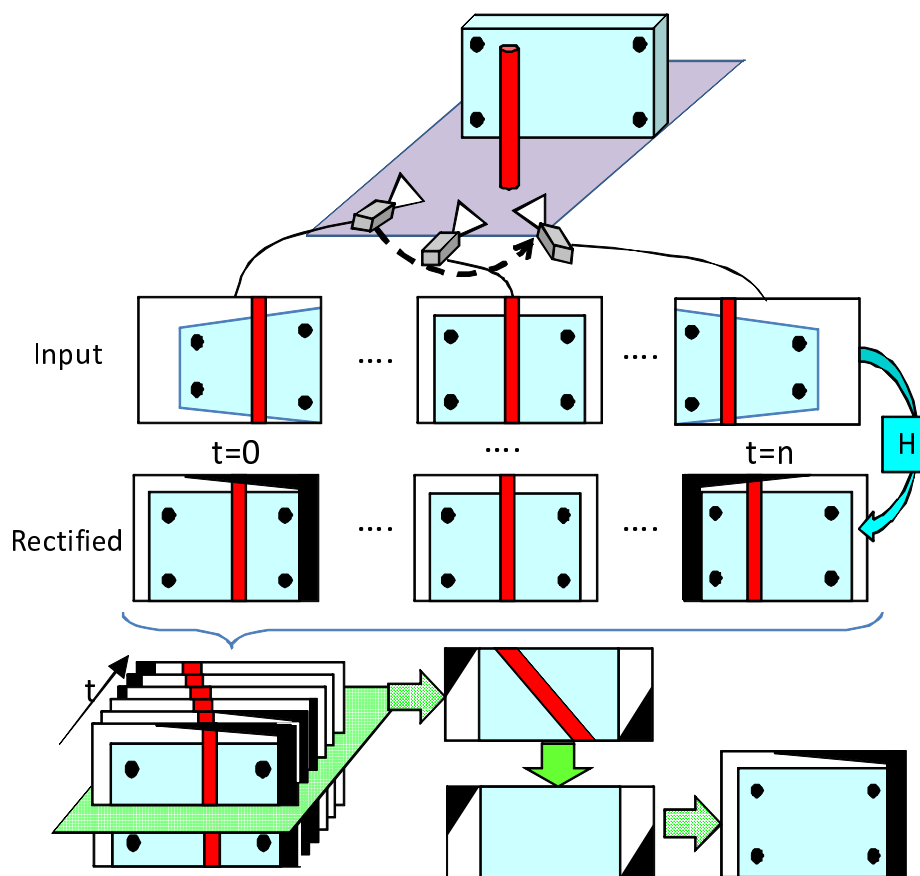


図 5.2 時空間処理を用いた背景検出

背景の抽出

ホモグラフィ変換をビデオシーケンスのすべてのフレームに適用することにより, 図.5.2. に示すように, 各フレームがある基準フレームの背景に位置合わせされた画像を生成することができる. このとき, 2D 座標上の点が前景オブジェクトによって遮られない場合, 時間線に沿ったカラー値は一定になる. 前景物体で頻繁に閉塞されるピクセルの背景色を決定するために, 我々は時間軸に沿ったカラーヒストグラムを作成し, 背景色が統計的に支配的であるという事実を利用するために Vector Median Filter (VMF) [105] を適用する.

アルファマップ付の前景検出

前景またはカメラのいずれかに高速な動きがある場合, 前景には動きブレが含まれている可能性が高くなる. さらに, オクリュージョン境界付近のブレは, 前景と背景の混合色を作り, 混合されたものから前景の色のみを抽出することは困難である. この問題を解決するために, 簡単な画像の差分を使用して, 初期のアルファマップの自動推定を行う. これは, 以前の方法では手動処理であっ

た. アルファマップは Levin の推定手法 [106] を適用する. 今回はホモグラフィ再投影した時のピクセル値の差が小さい部分を初期のアルファマップとした.

5.3.3 アルファマップとテクスチャのブレ除去

前節で推定されたアルファマップの品質は, 前景と背景のオブジェクトの色の類似性, ノイズなどの影響を受けて低下している. ここでは, アルファマップにブレ除去を適用することで問題を解決する.

間違ったアルファマップを用いて前景と背景を分離した場合, 前景と背景の両方がお互いを侵食する. その結果, アルファマップに誤りがある部分では, ブレ除去や位置合わせの工程でノイズが発生する. したがって, 対応するピクセルの分散が減少するようにアルファマップを更新することで正しいアルファマップを求める. 最終的なアルファマップを α , 最終的なテクスチャ $F(c)_{ij}$ とし, 入力フレームの数を N とする. アルファマップの画素数を P , \overline{F}_j は各画素のブラーカーネルとしての F_{ij}, K_{Fi} の輝度の平均とすると, 復元を次のように定式化することができる.

$$\arg \min_a \sum_{i=1}^N \sum_{j=1}^P \|F_{ij} - \overline{F}_j\|^1 + \|I_{ij} - K_{Fi} * (\alpha_j \circ F_j) - B\|^1, \quad (5.3)$$

$$\arg \min_F \sum_{j=1}^P \|I_{ij} - K_{Fi} * (\alpha_j \circ F_j) - B\|^1, \quad (5.4)$$

このプロセスは, 図.5.3 の最下段に示されている.

前景画像と背景画像の画素色が類似している場合, 単純に背景画像との差分を取るアプローチでは前景画像を抽出するときにあいまいさが残る. この場合, 正しいアルファマップがあれば, あいまいさのない前景画像を抽出することが可能である. 我々の方法は, 式 (5.1) により計算された捕捉画像と合成画像との間の差を最小にする MAP 推定に基づいている. 実際の関数は以下に示す.

$$\arg \min_F \sum_{i=1}^N \sum_{j=1}^P \|I_{ij} - (1 - K_{Fi} * \alpha_j) \circ B_j + K_{Fi} * (\alpha_j \circ F_j)\|^1 + \lambda \|C \cdot \alpha_j\|^1, \quad (5.5)$$

ここで \circ は行列の同じ行と列のピクセル単位の乗算を表し, C は正則化のための最終的なアルファマップの事前情報であり, λ はその重みである. この論文では, 我々は, Bilateral Total Variation (BTV) [2]. を適用して評価関数を最適化する. このプロセスは, 図 5.3. の上段に示されている.

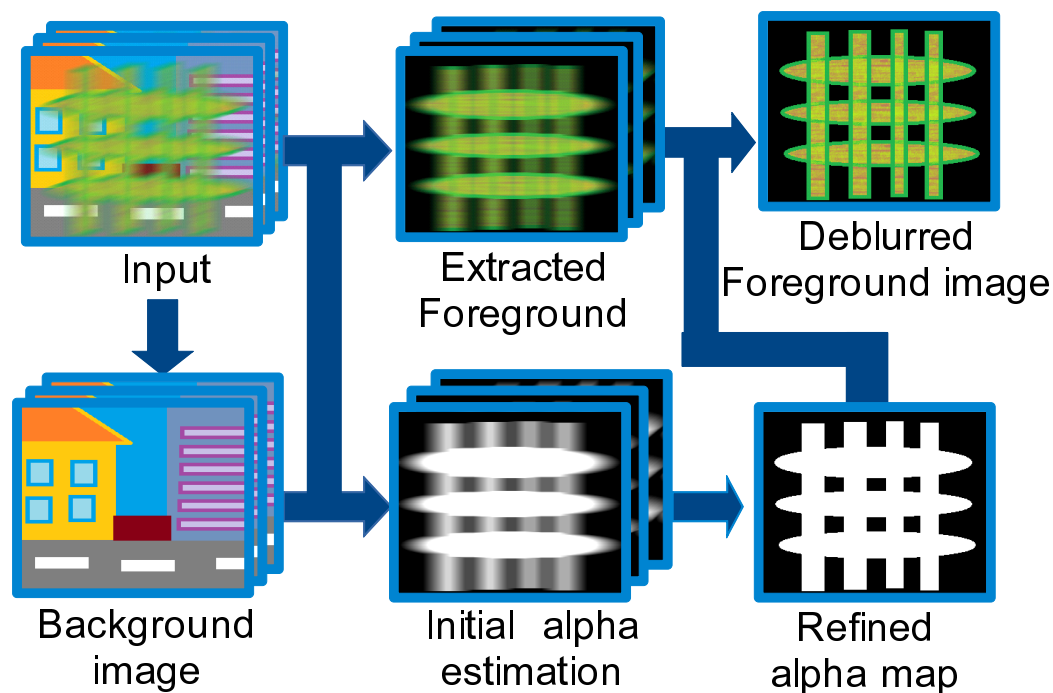


図 5.3 アルファマップとテクスチャのブレ除去の工程.

5.4 実験結果

5.4.1 シミュレーション画像のブレ除去

まず, シミュレーションデータを使用して方法を評価する. 図 5.4(b). に示すように, 前景オブジェクトに高頻度の 2 層オブジェクトを用意する. 次に, 図 5.4 (a) に示すように, 4 方向からランダム方向への直線運動をシミュレートする. このデータセットを使用して, われわれの時空分析を使用してアルファマップ推定を行った. 最初のアルファマップには誤差があるため, 図 5.4 (c) に示すようにアルファマップの細分化を行った. 最後に, 境界付近の混色を考慮したブレ除去アルゴリズムを実施した.

我々はまた, 比較のためにオプティカルフローを用いてピクセル毎にブラーカーネルを近似する Kyoung の最先端のブレ除去技術 [28] を適用する. その結果を図 5.4 (d) に示す. 図 5.4 (c) ~ (d) が示すように, 我々の境界細分化アルゴリズムでは, 特に物体の境界付近にセグメンテーションフリーの手法 [28] よりも優れていると PSNR から確認できる.

また, 同様に図 5.5 のような格子状の物体のデータでも同様の結果が出たが, Kyoung の手法 [28] では前景推定の崩壊が見られた一方, 我々の手法の方が複雑な形状のシーンでも正しく前景物体を求めることが成功していると分かる.

5.4.2 実測画像のブレ除去

次に、前景オブジェクトを速度の異なる任意の方向に移動させた実カメラ画像を用いて実験を行った。実験では、複数フレームのブレ除去に 10 フレームを使用し、輝度値によるセグメンテーションを用いた方法 [104]、オプティカルフローベースの方法と [28] 結果を比較した。撮影シーンについてはテクスチャのある格子状の物体を局所的に複雑なテクスチャを背景に撮影したもの (図 5.6) を基準に、前景領域の中に背景を含まないカードを前景にしたもの (図 5.8)、より実際に存在しうるシーンを想定して屋外でテクスチャの無いシーン (図 5.7)、車載カメラによるシーン (図 5.9) を撮影した。比較する手法 [104, 28] ではオブジェクトの境界正しく推定されていないため、境界付近の結果にアーティファクトがあるが、提案手法では前景と背景の両方が正常にブレ除去処理される。

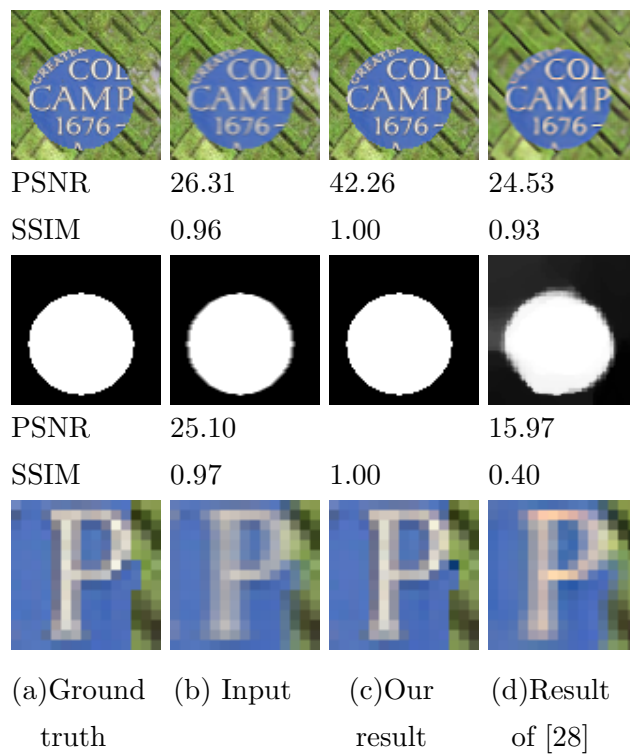


図 5.4 シミュレーション画像 (円) のブレ除去結果。 上段:テクスチャ 中段:前景のアルファマップ 下段:テクスチャの拡大図

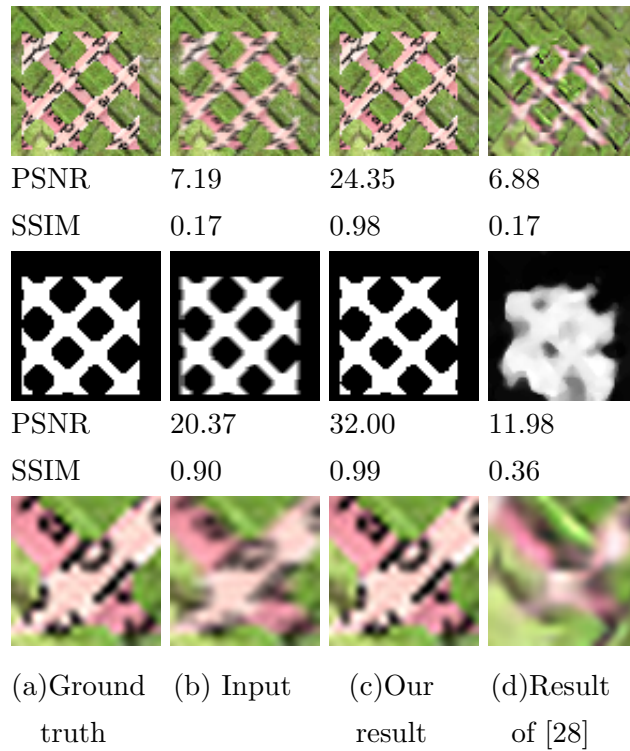


図 5.5 シミュレーション画像 (格子) のブレ除去結果. 上段:テクスチャ 中段:前景のアルファマップ 下段:テクスチャの拡大図

5.5 まとめ

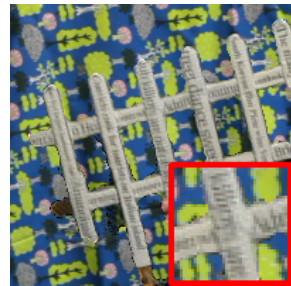
この章では, 移動物体によってオクリュージョンが発生した画像シーケンスに対して, 境界形状推定と画像ぼけの両方を実現する手法を提案した. この方法では, まず, 背景差分法で初期の前景画像を抽出し, 位置合わせ及び前景と背景に分割された画像の画素の分散を最小化することで高精度化される. 最後に, 精緻化されたアルファマップと複数の入力を使用して, 前景色のブレ除去を行った. 実験では, ブレのある画像を明確な境界で正常に復元することができた. 将来の研究として, 互いに重なり合っている複数の動く物体に対応することを計画している.



(a) Input image



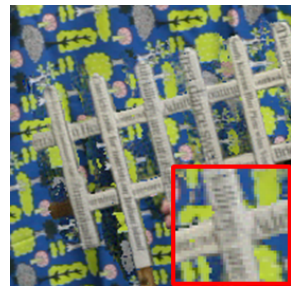
(b) Background image



(c) Our result



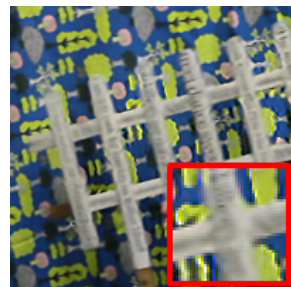
(d) Refined alpha map



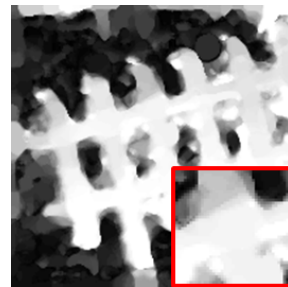
(e) Result of [104]



(f) Alpha map of [104]



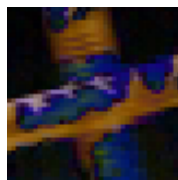
(g) Result of [28]



(h) Alpha map of [28]



(k) Diff
(b)and(e)



(l) Diff
(b)and(g)



(i) Diff
(c)and(g)

図 5.6 屋内シーンでのブレ除去結果.



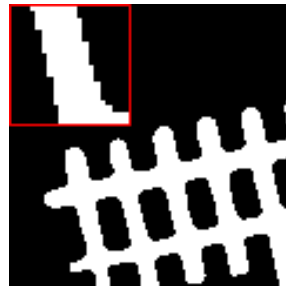
(a) Input image



(b) Background image



(c) Our result



(d) Refined alpha map



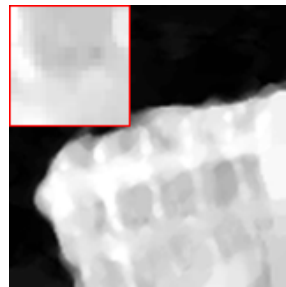
(e) Result of [104]



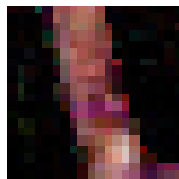
(f) Alpha map of [104]



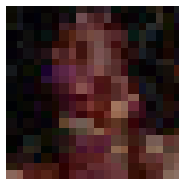
(g) Result of [28]



(h) Alpha map of [28]



(k) Diff
(b)and(e)



(l) Diff
(b)and(g)

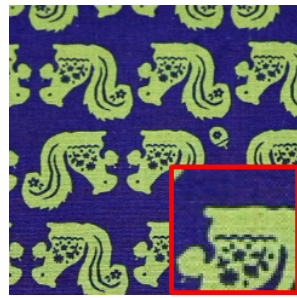


(i) Diff
(c)and(g)

図 5.7 屋外シーンのブレ除去結果



(a) Input image



(b) Background image



(c) Our result



(d) Refined alpha map



(e) Result of [104]



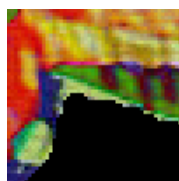
(f) Alpha map of [104]



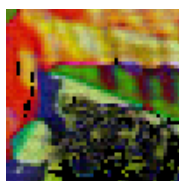
(g) Result of [28]



(h) Alpha map of [28]



(e) Result by [104]



(f) Alpha map by [104]



(i) Diff(c)and(g)

図 5.8 室内シーンでのブレ除去結果 (カード)



(a) Input image



(b) Initial alpha map



(c) Our result



(d) Refined alpha map



(e) Result by [104]



(f) Alpha map by [104]



(g) Result by [28]



(h) Alpha map by [28]



(i) Zoom of (a)



(j) Zoom of (b)



(k) Zoom of (f)



(l) Zoom of (d)



(m) Zoom of (e)



(n) Zoom of (c)



(o) Zoom of (g)



(p) Zoom of (h)

図 5.9 車載カメラシーンでの結果.

第 6 章

考察

本論文では以下の 3 点によって出力画像の精度が向上したと主張している.

1. トラッキング精度・ロバスト性の向上による入力フレーム増加
2. ブレ除去と超解像の同時処理
3. 前景物体部分の推定精度の向上

この章では, 実験結果をもとにそれらの主張の妥当性について考察する.

6.1 入力枚数と結果の相関

図 4.3, 図 4.7, 図 4.8 では画像処理結果を入力枚数で比較した. 入力枚数が増えると画像処理結果の改善が確認できるが, 特に 1 枚使用した場合と 2 枚使用した場合を比較するとリングングの量が大きく減少していることが分かる. その理由を確認するために, 新たに別のシミュレーション画像で画像処理を行う. フレームごとに方向を 30 度ずつ変化させた 13 ピクセルの直線ブレを与えた画像を入力とし, これに 2 x 2 倍の超解像を行った. 初期画像は基準フレームを単純に拡大したものをを用いた. 結果を図 6.1 に示す.

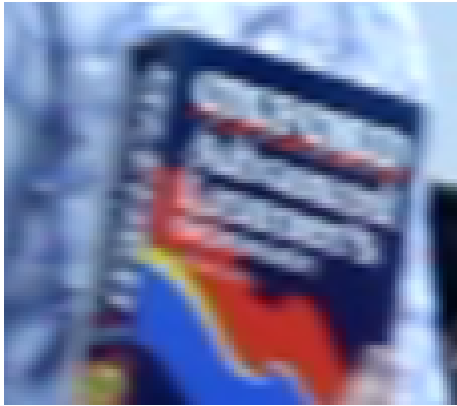
図 4.3 と図 6.1 の PSNR を比較すると, 入力画像の PSNR は図 6.1 の方が低い, 画像処理後の PSNR は高くなっていることが確認できる. 1 枚使用の結果の違いは与えたブラーの角度の差によるものと考えられるが, 5 枚, 10 枚使用の結果の違いはフレーム間のブラーカーネルが図の方が直交状態に近くなったことが原因と考えられる. Feng らの論文 [47] にあるように複数枚入力のブレ除去はフレーム間のブラーカーネルの関係が直交状態に近いほど結果かが向上する.



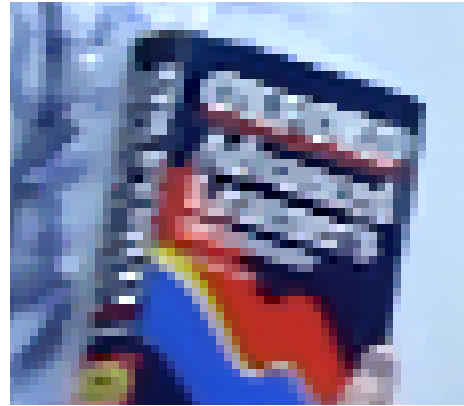
図 6.1 シミュレーション画像による結果 (a) は入力画像 (b) 1 枚使用結果 (c) は 5 枚使用結果 (d) は 10 枚使用結果

6.2 ブレ除去と超解像の同時処理

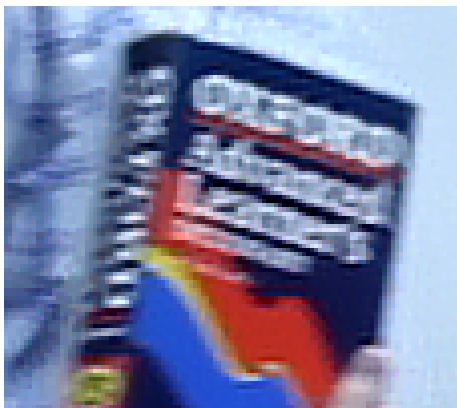
図 4.4, 図 6.2 では画像処理結果を手法別で比較した。ブレ除去のみ、超解像のみ、ブレ除去と超解像を同時に行った場合それぞれが入力画像と比較してブラーが軽減されていることが確認できる。超解像のみでの結果にもブラーが軽減が確認できるのは入力画像中にブラーが比較的小さいフレームがあること、フレーム毎にブラーの方向が違うことによってピクセル毎の輝度値を平均したときにそれぞれのノイズが平滑化されたためだと考えられる。ブレ除去のみと同時処理の違いは、ブレ除去のみの結果にジャギーが確認できる点である。ジャギーは画像を平滑化することで軽減できるが、同時処理画像については、エッジが保持されたままジャギーが無い画像を生成できている点に注目したい。



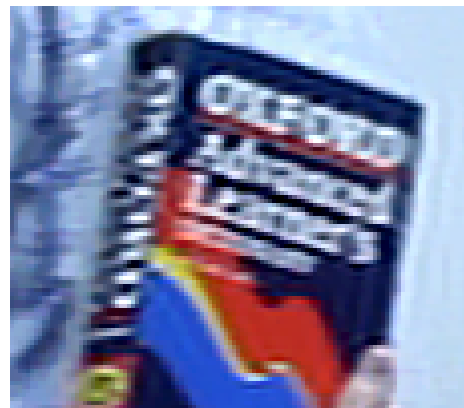
(a) Input



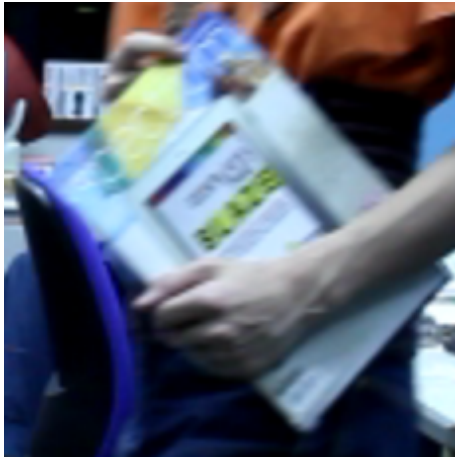
(b) Deblurring only



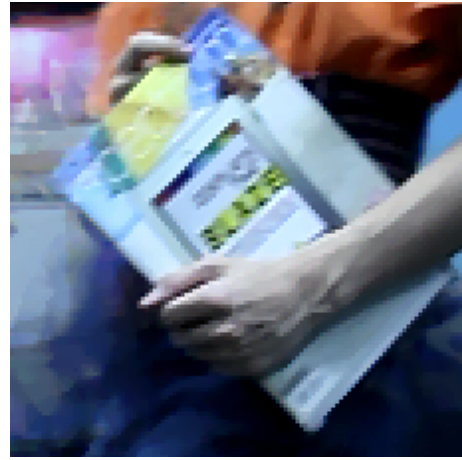
(c) SR only



(d) SR + deblur (proposed)



(e) Input



(f) Deblurring only



(g) SR only



(h) SR + deblur (proposed)

図 6.2 実測画像による結果: (a)-(d) 左側の本の処理前の画像と処理ごとの結果, (e)-(h) 右側の本の処理前の画像と処理ごとの結果.

6.3 前景部分推定と結果の関係

図 5.6-5.8 の結果を比較すると Kyoung らの手法 [28] や松下の先行研究 [104] では前景部分の推定に失敗していることが分かる. 先行研究 [104] はセグメンテーションを位置合わせした各フレームの輝度値差で行っているが, 前景に位置合わせしたときに背景部分の輝度値の差が小さい場合にその部分が前景として扱われている問題が確認できる.

また, Kyoung らの手法 [28] は隣接フレームのオプティカルフローから前景部分を求めている. 先行研究 [104] と比較して背景にテクスチャがあるシーンでは比較的正しく前景部分を推定できる. また, 前景物体の格子が細かいシーン (図 6.3) と太いシーン (図 6.4) で比較を行った. 図 6.3-6.4 のような背景のテクスチャの変化が少ないシーンでは誤推定が多い. 本手法は背景を抽出し, 厳密

なアプローチで前景を推定するため、比較手法で失敗するようなシーンでも正しく前景を求めることに成功した。

図 5.9 のような屋外の車載映像の場合でも同様の傾向がみられるが、図 5.9(c)(d) にあるように提案手法でも他のシーンと比較してアルファマップの推定に失敗している領域が多い。これはカメラの移動が水平方向であるのに対して背景のテクスチャも水平模様で単色であるためアルファマップを変化させてもアルファマップ高精度化の式 (5.3) の値の変化が小さい状況であったためである。また、図 6.5 にあるように、提案手法のアルファマップの推定の推定が正しい部分ではブレ除去時にオクリュージョン部分周辺のリングングが削減されていることが分かる。これによって、背景画像を求め、前景物体と背景物体の混色を考慮した処理を行うことで処理後のノイズが減ることを確認できた。



(a) Input image



(b) Background image



(c) Our result



(d) Refined alpha map



(e) Result of [104]



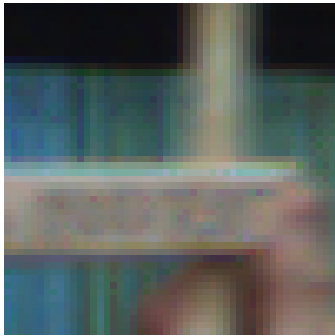
(f) Alpha map of [104]



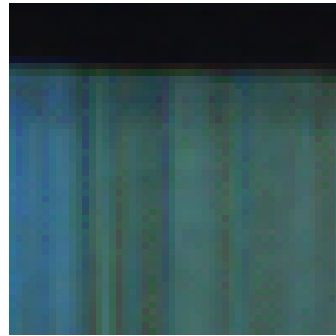
(g) Result of [28]



(h) Alpha map of [28]



(i) Zoom of (b)



(j) Background



(k) Zoom of (c)



(l) Zoom of (d)



(m) Difference
between (b) and (c)



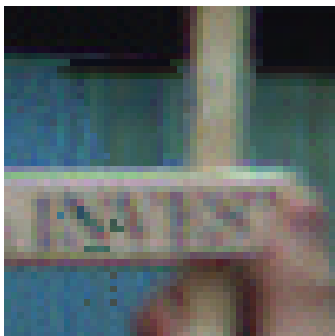
(o) Zoom of (e)



(p) Zoom of (f)



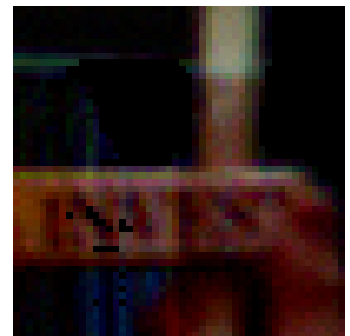
(q) Difference
between (b) and (e)



(r) Zoom of (g)

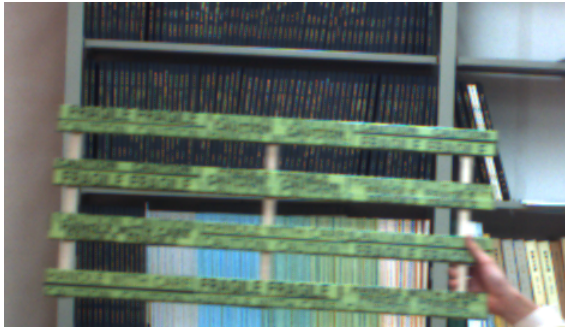


(s) Zoom of (h)



(t) Difference
between (b) and (g)

図 6.3 室内シーンでのブレ除去結果 (格子物体 1)



(a) Input image



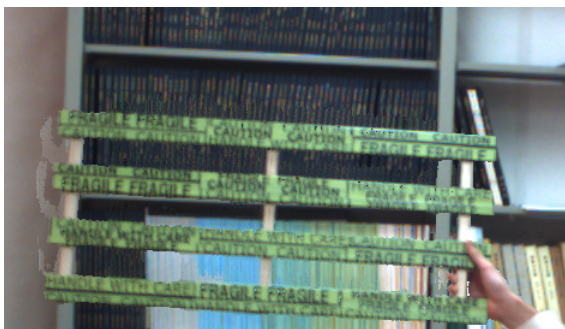
(b) Background image



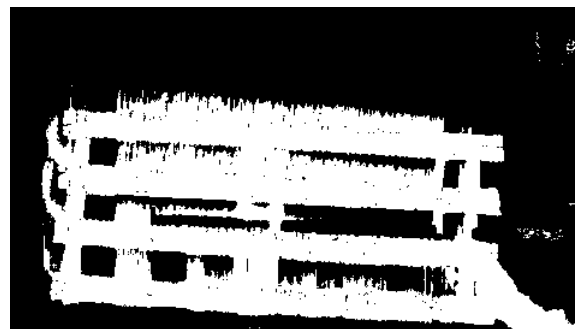
(c) Our result



(d) Refined alpha map



(e) Result of [104]



(f) Alpha map of [104]



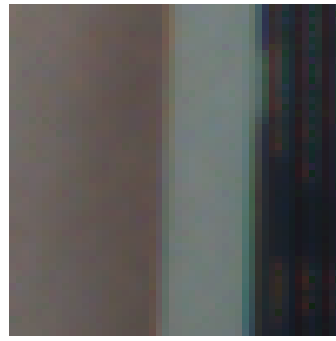
(g) Result of [28]



(h) Alpha map of [28]



(i) Zoom of (b)



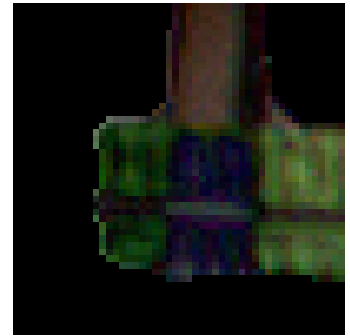
(j) Background



(k) Zoom of (c)



(l) Zoom of (d)



(m) Difference
between (b) and (c)



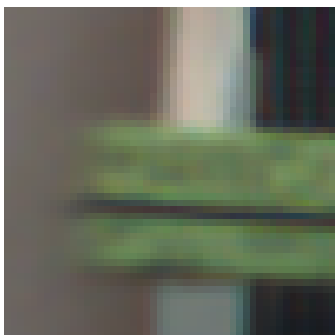
(o) Zoom of (e)



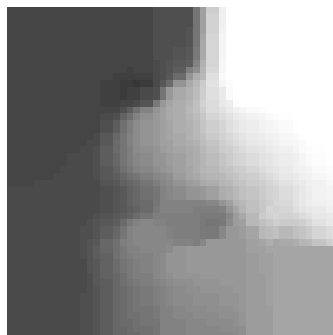
(p) Zoom of (f)



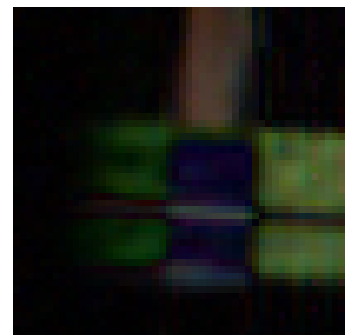
(q) Difference
between (b) and (e)



(r) Zoom of (g)



(s) Zoom of (h)



(t) Difference
between (b) and (g)

図 6.4 室内シーンでのブレ除去結果 (格子物体 2)



図 6.5 車載カメラシーンの境界部分におけるアーティファクトの有無.

第7章

おわりに

本論文では、複数の被写体に、異なるモーションブラーがあるシーンに対応したブレ除去および超解像手法を提案した。提案手法では、入力する連続画像を短いシーケンスに分割し、平面領域の計算を行い、それらを統合することで長いフレームでの平面トラッキングを行う。また、モーションブラーを考慮した位相相関法（POC）によるトラッキング結果の高精度化を提案した。さらに、複数枚を入力として、ブレ除去と超解像を同時に行う手法を提案した。また、前景物体と背景の考慮した画像処理の手法を提案した。実験により、提案手法により高精度な画像処理を実現できることを示した。

本論分の処理は、分割するシーケンスのフレーム数や、ブラーカーネルのスケーリング変数など処理結果に大きくかわるパラメーターがある。これらは、経験的に設定していたが処理中に望ましい値を自動計算させるようにするのが望ましいだろう。特に、ブラーカーネルの推定については高精度化画像の推定と同時に行う手法も多い [107]。また、シーケンスごとの結果の統合については、重複フレーム以外のセグメンテーション結果や入力画像の輝度情報など、事前情報を有効活用できていない実装がある。今後は、フレーム間のモーションから、ブラーカーネルのスケーリング変数の自動計算や、より高精度の特徴点トラッキング手法、セグメンテーション手法を組み込む予定である。また、オクリュージョンを考慮した処理については超解像処理の追加や、前景物体が複数あるシーンを考慮することで、さらに実践的なシーンに対応することができる。そのためにはシーンのモデル化及びアルファマップ及び画像の高精度化処理の方法をより深く検討する必要がある。

謝辞

本論文を執筆するにあたって多くの人からの助力があった。まず研究の根幹部分の指導をしてくださった九州大学の川崎洋先生に心から感謝申し上げます。そして最終的な論文の仕上げについて指導してくださった渡邊睦先生に深謝いたします。また、小野智司先生、佐藤公則先生には副査として多くのアドバイスを頂きました。ありがとうございます。福元伸也助教、事務員の和田さんには非常に手厚いサポートを受けました。本当に助かりました。東京大学の池内克史教授、小野晋太郎先生には隔週のライブミーティングで丁寧なご指導をいただきました。さらに、三嶋道弘助教本学におられた赤木康宏特任准教授には、在学中発表の練習や実験について、さまざまなお助言をいただきました。また、川崎研究室のメンバーだった方々、渡邊研究室のメンバーの方々には多くの勇気をもらいました。最後に、食事、就寝、風呂などあらゆる生活の用意を毎日整えて下さった家族の皆様に、心より感謝いたします。研究生活に限らず、生き方などたくさん心配をかけたかと思いますが、よくここまで耐えてもらえたと、尊敬します。

参考文献

- [1] D. Glasner, S. Bagon, and M. Irani. Super-resolution from a single image. In *ICCV*, 2009.
- [2] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar. Fast and robust multiframe super resolution. *IEEE Trans. on Image Processing*, Vol. 13, No. 10, pp. 1327–1344, 2004.
- [3] M. Turkan, D. Thoreau, and P. Guillotel. Iterated neighbor-embeddings for image super-resolution. In *2014 IEEE International Conference on Image Processing (ICIP)*, pp. 3887–3891, Oct 2014.
- [4] H. S. Lee. A new image super resolution by texture transfer. In *2014 IEEE International Conference on Image Processing (ICIP)*, pp. 3915–3918, Oct 2014.
- [5] A. Singh, F. Porikli, and N. Ahuja. Super-resolving noisy images. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2846–2853, June 2014.
- [6] C. Dong, C. C. Loy, K. He, and X. Tang. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 38, No. 2, pp. 295–307, Feb 2016.
- [7] J. Kim, J. K. Lee, and K. M. Lee. Deeply-recursive convolutional network for image super-resolution. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1637–1645, June 2016.
- [8] Tony Tung, Shohei Nobuhara, and Takashi Matsuyama. Simultaneous super-resolution and 3d video using graph-cuts. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, June 2008.
- [9] Uma Mudenagudi, Ankit Gupta, Lakshya Goel, Avanish Kushal, Prem Kalra, and Subhashis Banerjee. Super resolution of images of 3d scenecs. In *Proceedings of the 8th Asian Conference on Computer Vision - Volume Part II, ACCV'07*, pp. 85–95, Berlin, Heidelberg, 2007. Springer-Verlag.
- [10] L. B. Lucy. An iterative technique for the rectification of observed distributions. *Astron.*

- J.*, Vol. 79, pp. 745+, June 1974.
- [11] William Hadley Richardson. Bayesian-based iterative method of image restoration*. *J. Opt. Soc. Am.*, Vol. 62, No. 1, pp. 55–59, Jan 1972.
 - [12] Michal Irani Irani and Shmuel Peleg. Improving resolution by image registration. *CVGIP*, Vol. 53, pp. 231–239, 1991.
 - [13] Qi Shan, Jiaya Jia, and Aseem Agarwala. High-quality motion deblurring from a single image. *ACM Transactions on Graphics (SIGGRAPH)*, 2008.
 - [14] Michael Hirsch, Christian J. Schuler, Stefan Harmeling, Bernhard Schölkopf. Fast removal of non-uniform camera shake. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 463–470, 11 2011.
 - [15] Z. Hu, L. Xu, and M. H. Yang. Joint depth estimation and camera shake removal from single blurry image. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2893–2900, June 2014.
 - [16] Jin shan Pan, Zhe Hu, Zhixun Su, Hsin-Ying Lee, and Ming-Hsuan Yang. Soft-segmentation guided object motion deblurring. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 459–468, 2016.
 - [17] Jochen Gast, Anita Sellent, and Stefan Roth. Parametric object motion from blur. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1846–1854, 2016.
 - [18] J. Pan, Z. Hu, Z. Su, and M. H. Yang. Deblurring text images via l0-regularized intensity and gradient prior. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2901–2908, June 2014.
 - [19] S. H. Park and M. Levoy. Gyro-based multi-image deconvolution for removing handshake blur. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3366–3373, June 2014.
 - [20] J. Shi, L. Xu, and J. Jia. Just noticeable defocus blur detection and estimation. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 657–665, June 2015.
 - [21] Jian Sun, Wenfei Cao, Zongben Xu, and Jean Ponce. Learning a convolutional neural network for non-uniform motion blur removal. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.
 - [22] H. Zhang and L. Carin. Multi-shot imaging: Joint alignment, deblurring, and resolution-enhancement. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*,

- pp. 2925–2932, June 2014.
- [23] *Image Deblurring with Blurred/Noisy Image Pairs*, Vol. 26. ACM, August 2007.
 - [24] Amit Agrawal, Yi Xu, and Ramesh Raskar. Invertible motion blur in video. In *ACM Transactions on Graphics (TOG)*, Vol. 28, p. 95. ACM, 2009.
 - [25] H. G. Jeon, J. Y. Lee, Y. Han, S. J. Kim, and I. S. Kweon. Complementary sets of shutter sequences for motion deblurring. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 3541–3549, Dec 2015.
 - [26] Sunghyun Cho, Jue Wang, and Seungyong Lee. Video deblurring for hand-held cameras using patch-based synthesis. *ACM Trans. Graph.*, Vol. 31, No. 4, pp. 64:1–64:9, July 2012.
 - [27] Jonas Wulff and Michael J. Black. Modeling blurred video with layers. In *Computer Vision – ECCV 2014*, Vol. 8694 of *Lecture Notes in Computer Science*, pp. 236–252. Springer International Publishing, September 2014.
 - [28] Tae Hyun Kim and Kyoung Mu Lee. Generalized video deblurring for dynamic scenes. In *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*, pp. 5426–5434, June 2015.
 - [29] David G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, Vol. 60, No. 2, pp. 91–110, November 2004.
 - [30] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-up robust features (surf). *Comput. Vis. Image Underst.*, Vol. 110, No. 3, pp. 346–359, June 2008.
 - [31] Kenji TAKITA, Takafumi AOKI, Yoshifumi SASAKI, Tatsuo HIGUCHI, and Koji KOBAYASHI. High-accuracy subpixel image registration based on phase-only correlation (special section, digital signal processing). *IEICE transactions on fundamentals of electronics, communications and computer sciences*, Vol. 86, No. 8, pp. 1925–1934, aug 2003.
 - [32] Carlo Tomasi and Takeo Kanade. Detection and tracking of point features. Technical report, International Journal of Computer Vision, 1991.
 - [33] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce. Non-uniform deblurring for shaken images. *International Journal of Computer Vision*, Vol. 98, No. 2, pp. 168–186, 2012.
 - [34] Norbert Wiener. *Extrapolation, interpolation, and smoothing of stationary time series*, Vol. 7. MIT press Cambridge, MA, 1949.
 - [35] Lu Yuan, Jian Sun, Long Quan, and Heung-Yeung Shum. Image deblurring with blurred/noisy image pairs. In *ACM SIGGRAPH 2007 Papers*, SIGGRAPH '07, New

- York, NY, USA, 2007. ACM.
- [36] Jian-Feng Cai, Hui Ji, Chaoqiang Liu, and Zuowei Shen. Blind motion deblurring using multiple images. *J. Comput. Phys.*, Vol. 228, No. 14, pp. 5057–5071, August 2009.
 - [37] Jia Chen, Lu Yuan, Chi-Keung Tang, and Long Quan. Robust dual motion deblurring. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, June 2008.
 - [38] O. Shahar, A. Faktor, and M. Irani. Space-time super-resolution from a single video. In *CVPR 2011*, pp. 3353–3360, June 2011.
 - [39] E. Shechtman, Y. Caspi, and M. Irani. Space-time super-resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 4, pp. 531–545, April 2005.
 - [40] Patrick Llull, Xuejun Liao, Xin Yuan, Jianbo Yang, David Kittle, Lawrence Carin, Guillermo Sapiro, and David J. Brady. Coded aperture compressive temporal imaging. *Opt. Express*, Vol. 21, No. 9, pp. 10526–10545, May 2013.
 - [41] D. Liu, J. Gu, Y. Hitomi, M. Gupta, T. Mitsunaga, and S. K. Nayar. Efficient space-time sampling with pixel-wise coded exposure for high-speed imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 36, No. 2, pp. 248–260, Feb 2014.
 - [42] Dikpal Reddy, Ashok Veeraraghavan, and Rama Chellappa. P2c2: Programmable pixel compressive camera for high speed imaging. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pp. 329–336. IEEE, 2011.
 - [43] Ramesh Raskar, Amit Agrawal, and Jack Tumblin. Coded exposure photography: Motion deblurring using fluttered shutter. In *ACM SIGGRAPH 2006 Papers*, SIGGRAPH '06, pp. 795–804, New York, NY, USA, 2006. ACM.
 - [44] H. G. Jeon, J. Y. Lee, Y. Han, S. J. Kim, and I. S. Kweon. Fluttering pattern generation using modified legendre sequence for coded exposure imaging. In *2013 IEEE International Conference on Computer Vision*, pp. 1001–1008, Dec 2013.
 - [45] Amit Agrawal, Yi Xu, and Ramesh Raskar. Invertible motion blur in video. In *ACM SIGGRAPH 2009 Papers*, SIGGRAPH '09, pp. 95:1–95:8, New York, NY, USA, 2009. ACM.
 - [46] J. Holloway, A. C. Sankaranarayanan, A. Veeraraghavan, and S. Tambe. Flutter shutter video camera for compressive sensing of videos. In *2012 IEEE International Conference on Computational Photography (ICCP)*, pp. 1–9, April 2012.
 - [47] Feng Li, Z. Li, D. Saunders, and Jingyi Yu. A theory of coprime blurred pairs. In *2011 International Conference on Computer Vision*, pp. 217–224, Nov 2011.

- [48] Chin-Chong Tseng and C. Liu. Complementary sets of sequences. *IEEE Transactions on Information Theory*, Vol. 18, No. 5, pp. 644–652, Sep 1972.
- [49] P. Spasojevic and C. N. Georghiadis. Complementary sequences for isi channel estimation. *IEEE Transactions on Information Theory*, Vol. 47, No. 3, pp. 1145–1152, Mar 2001.
- [50] Jakob Bernasconi. Low autocorrelation binary sequences: statistical mechanics and configuration space analysis. *Journal de Physique*, Vol. 48, No. 4, pp. 559–567, 1987.
- [51] M. M. Naghsh M. Soltanalian and P. Stoica. A fast algorithm for designing complementary sets of sequences. *Signal Processing*, Vol. 93, No. 7, pp. 2096–2102, 2013.
- [52] Yohann Tendero, Jean-Michel Morel, and Bernard Rougé. The flutter shutter paradox. *SIAM Journal on Imaging Sciences*, Vol. 6, No. 2, pp. 813–847, 2013.
- [53] Scott McCloskey. Velocity-dependent shutter sequences for motion deblurring. *Computer Vision–ECCV 2010*, pp. 309–322, 2010.
- [54] S. McCloskey, Y. Ding, and J. Yu. Design and estimation of coded exposure point spread functions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 34, No. 10, pp. 2071–2077, Oct 2012.
- [55] Matthew G Parker, Kenneth G Paterson, and Chintha Tellambura. Golay complementary sequences. *Encyclopedia of Telecommunications*, 2003.
- [56] P. Z. Fan, N. Suehiro, N. Kuroyanagi, and X. M. Deng. Class of binary sequences with zero correlation zone. *Electronics Letters*, Vol. 35, No. 10, pp. 777–779, May 1999.
- [57] A. Agrawal and Yi Xu. Coded exposure deblurring: Optimized codes for psf estimation and invertibility. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2066–2073, June 2009.
- [58] Shengyang Dai and Ying Wu. Motion from blur. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pp. 1–8. IEEE, 2008.
- [59] Y. W. Tai, N. Kong, S. Lin, and S. Y. Shin. Coded exposure imaging for projective motion deblurring. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2408–2415, June 2010.
- [60] A. Levin, D. Lischinski, and Y. Weiss. A closed-form solution to natural image matting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30, No. 2, pp. 228–242, Feb 2008.
- [61] Anat Levin, Rob Fergus, Frédo Durand, and William T. Freeman. Image and depth from a conventional camera with a coded aperture. In *ACM SIGGRAPH 2007 Papers*,

- SIGGRAPH '07, New York, NY, USA, 2007. ACM.
- [62] S. Kim, Y. W. Tai, Y. Bok, H. Kim, and I. Kweon. Two-phase approach for multi-view object extraction. In *2011 18th IEEE International Conference on Image Processing*, pp. 2361–2364, Sept 2011.
- [63] Yoav Y Schechner, Shree K Nayar, and Peter N Belhumeur. Multiplexing for optimal lighting. *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 29, No. 8, pp. 1339–1354, 2007.
- [64] Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, Vol. 13, No. 4, pp. 600–612, April 2004.
- [65] Yuri Boykov, Olga Veksler, and Ramin Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 23, No. 11, pp. 1222–1239, November 2001.
- [66] M. Upmanyu, A. M. Namboodiri, K. Srinathan, and C. V. Jawahar. Efficient privacy preserving video surveillance. In *2009 IEEE 12th International Conference on Computer Vision*, pp. 1639–1646, Sept 2009.
- [67] Ankur Chattopadhyay and Terrance E Boulton. Privacycam: a privacy preserving camera using uclinux on the blackfin dsp. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pp. 1–8. IEEE, 2007.
- [68] Stephen R. Gottesman and E. E. Fenimore. New family of binary arrays for coded aperture imaging. *Appl. Opt.*, Vol. 28, No. 20, pp. 4344–4352, Oct 1989.
- [69] Sunghyun Cho and Seungyong Lee. Fast motion deblurring. In *ACM SIGGRAPH Asia 2009 Papers*, SIGGRAPH Asia '09, pp. 145:1–145:8, New York, NY, USA, 2009. ACM.
- [70] Rob Fergus, Barun Singh, Aaron Hertzmann, Sam T. Roweis, and William T. Freeman. Removing camera shake from a single photograph. In *ACM SIGGRAPH 2006 Papers*, SIGGRAPH '06, pp. 787–794, New York, NY, USA, 2006. ACM.
- [71] Ankit Gupta, Neel Joshi, Lary Zitnick, Michael Cohen, and Brian Curless. Single image deblurring using motion density functions. In *ECCV '10: Proceedings of the 10th European Conference on Computer Vision*, 2010.
- [72] M. Hirsch, C. J. Schuler, S. Harmeling, and B. Scholkopf. Fast removal of non-uniform camera shake. In *2011 International Conference on Computer Vision*, pp. 463–470, Nov 2011.
- [73] Y. Li, S. B. Kang, N. Joshi, S. M. Seitz, and D. P. Huttenlocher. Generating sharp

- panoramas from motion-blurred videos. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2424–2431, June 2010.
- [74] Sunghyun Cho, Hojin Cho, Yu-Wing Tai, and Seungyong Lee. Registration based non-uniform motion deblurring. *Comput. Graph. Forum*, Vol. 31, No. 7pt2, pp. 2183–2192, September 2012.
- [75] H. S. Lee and K. M. Lee. Dense 3d reconstruction from severely blurred images using a single moving camera. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 273–280, June 2013.
- [76] C. Paramanand and A. N. Rajagopalan. Non-uniform motion deblurring for bilayer scenes. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1115–1122, June 2013.
- [77] S. Cho, Y. Matsushita, and S. Lee. Removing non-uniform motion blur from images. In *2007 IEEE 11th International Conference on Computer Vision*, pp. 1–8, Oct 2007.
- [78] Leah Bar, Benjamin Berkels, Martin Rumpf, and Guillermo Sapiro. A variational framework for simultaneous motion estimation and restoration of motion-blurred video. In *11th IEEE International Conference on Computer Vision*. IEEE, 2007.
- [79] Jonas Wulff and Michael Julian Black. *Modeling Blurred Video with Layers*, pp. 236–252. Springer International Publishing, Cham, 2014.
- [80] T. H. Kim and K. M. Lee. Segmentation-free dynamic scene deblurring. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2766–2773, June 2014.
- [81] Y. Matsushita, E. Ofek, Weina Ge, Xiaoou Tang, and Heung-Yeung Shum. Full-frame video stabilization with motion inpainting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 7, pp. 1150–1163, July 2006.
- [82] T. Portz, L. Zhang, and H. Jiang. Optical flow in the presence of spatially-varying motion blur. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1752–1759, June 2012.
- [83] H. Jin, P. Favaro, and R. Cipolla. Visual tracking in the presence of motion blur. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Vol. 2, pp. 18–25 vol. 2, June 2005.
- [84] T. H. Kim, H. S. Lee, and K. M. Lee. Optical flow via locally adaptive fusion of complementary data costs. In *2013 IEEE International Conference on Computer Vision*, pp. 3344–3351, Dec 2013.
- [85] Antonin Chambolle and Thomas Pock. A first-order primal-dual algorithm for convex

- problems with applications to imaging. *J. Math. Imaging Vis.*, Vol. 40, No. 1, pp. 120–145, May 2011.
- [86] Andreas Wedel, Thomas Pock, Christopher Zach, Horst Bischof, and Daniel Cremers. Statistical and geometrical approaches to visual motion analysis. chapter An Improved Algorithm for TV-L1 Optical Flow, pp. 23–45. Springer-Verlag, Berlin, Heidelberg, 2009.
- [87] Sina Farsiu, Michael Elad, and Peyman Milanfar. A practical approach to superresolution. In *Visual Communications and Image Processing*, Vol. 6077, 2006.
- [88] Li Xu, Shicheng Zheng, and Jiaya Jia. Unnatural l0 sparse representation for natural image deblurring. In *Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition, CVPR '13*, pp. 1107–1114, Washington, DC, USA, 2013. IEEE Computer Society.
- [89] Tae Hyun Kim, Byeongjoo Ahn, and Kyoung Mu Lee. Dynamic scene deblurring. December 2013.
- [90] Takuma Yamaguchi, Hisato Fukuda, Ryo Furukawa, Hiroshi Kawasaki, and Peter F. Sturm. Video deblurring and super-resolution technique for multiple moving objects. In Ron Kimmel, Reinhard Klette, and Akihiro Sugimoto, editors, *ACCV (4)*, Vol. 6495 of *Lecture Notes in Computer Science*, pp. 127–140. Springer, 2010.
- [91] L B Lucy. An iterative technique for the rectification of observed distributions. *Astronomical Journal*, Vol. 79, No. 6, pp. 745–754, 1974.
- [92] Hadley William Richardson. Bayesian-based iterative method of image restoration. *Journal of Optical Society of America*, pp. 55–59, 1972.
- [93] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. In *SIGGRAPH*, pp. 787–794, 2006.
- [94] Chandramouli Paramanand and Ambasadram N. Rajagopalan. Non-uniform motion deblurring for bilayer scenes. *2013 IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 0, pp. 1115–1122, 2013.
- [95] Sunghyun Cho, Jue Wang, and Seungyong Lee. Video deblurring for hand-held cameras using patch-based synthesis. *ACM Transactions on Graphics*, Vol. 31, No. 4, pp. 64:1–64:9, 2012.
- [96] Yunpeng Li, Sing Bing Kang, Neel Joshi, Steven M. Seitz, and Daniel P. Huttenlocher. Generating sharp panoramas from motion-blurred videos. In *CVPR*, pp. 2424–2431, 2010.
- [97] Weisheng Dong, D Zhang, Guangming Shi, and Xiaolin Wu. Image deblurring and

- super-resolution by adaptive sparse domain selection and adaptive regularization. *Image Processing, IEEE Transactions on*, Vol. 20, No. 7, pp. 1838–1857, 2011.
- [98] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *In ECCV*, pp. 404–417, 2006.
- [99] Sei Nagashima, Takafumi Aoki, Tatsuo Higuchi, and Koji Kobayashi. A subpixel image matching technique using phase-only correlation. In *Intelligent Signal Processing and Communications, 2006. ISPACS'06. International Symposium on*, pp. 701–704. IEEE, 2006.
- [100] Masayuki Tanaka, Yoichi Yaguchi, and Masatoshi Okutomi. Robust and accurate estimation of multiple motions for whole-image super-resolution. In *Image Processing, 2008. IICIP 2008. 15th IEEE International Conference on*, pp. 649–652. IEEE, 2008.
- [101] Ville Ojansivu and Janne Heikkilä. Image registration using blur-invariant phase correlation. *IEEE Signal Process. Lett.*, Vol. 14, No. 7, pp. 449–452, 2007.
- [102] Takuma Yamaguchi, Hisato Fukuda, Ryo Furukawa, Hiroshi Kawasaki, and Peter Sturm. Video deblurring and super-resolution technique for multiple moving objects. In *The Tenth Asian Conference on Computer Vision (ACCV2010)*, Vol. 4, pp. 2227–2240, 2010.
- [103] Roxas Menandro, Shiratori Takaaki, and Ikeuchi Katsushi. Video completion via spatio-temporally consistent motion inpainting. *IPSP Transactions on Computer Vision and Applications (CVA)*, Vol. 6, pp. 98–102, aug 2014.
- [104] Yuki Matsushita, Hiroshi Kawasaki, Shintaro Ono, and Katsushi Ikeuchi. Simultaneous deblur and super-resolution technique for video sequence captured by hand-held video camera. In *IEEE*, 2014.
- [105] Jaakko Astola, Petri Haavisto, and Yrjo Neuvo. Vector median filters. *Proceedings of the IEEE*, Vol. 78, No. 4, pp. 678–689, 1990.
- [106] Anat Levin, Alex Rav-Acha, and Dani Lischinski. Spectral matting. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 30, No. 10, pp. 1699–1712, 2008.
- [107] Jian-Feng Cai, Hui Ji, Chaoqiang Liu, and Zuowei Shen. Blind motion deblurring using multiple images. *J. Comput. Physics*, Vol. 228, No. 14, pp. 5057–5071, 2009.
- [108] J. Jia. Single image motion deblurring using transparency. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, June 2007.
- [109] Q. Shan, W. Xiong, and J. Jia. Rotational motion deblurring of a rigid object from a single image. In *2007 IEEE 11th International Conference on Computer Vision*, pp. 1–8, Oct 2007.

- [110] L. Bar, B. Berkels, M. Rumpf, and G. Sapiro. A variational framework for simultaneous motion estimation and restoration of motion-blurred video. In *2007 IEEE 11th International Conference on Computer Vision*, pp. 1–8, Oct 2007.
- [111] Lu Yuan, J. Sun, L. Quan, and H. Y. Shum. Blurred/non-blurred image alignment using sparseness prior. In *2007 IEEE 11th International Conference on Computer Vision*, pp. 1–8, Oct 2007.
- [112] Qi Shan, Jiaya Jia, and Aseem Agarwala. High-quality motion deblurring from a single image. *ACM Trans. Graph.*, Vol. 27, No. 3, pp. 73:1–73:10, August 2008.
- [113] Lu Yuan, Jian Sun, Long Quan, and Heung-Yeung Shum. Progressive inter-scale and intra-scale non-blind image deconvolution. *ACM Trans. Graph.*, Vol. 27, No. 3, pp. 74:1–74:10, August 2008.
- [114] Yu-Wing Tai, Hao Du, M. S. Brown, and S. Lin. Image/video deblurring using a hybrid camera. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, June 2008.
- [115] X. Huang, Liu Ren, and R. Yang. Image deblurring for less intrusive iris capture. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1558–1565, June 2009.
- [116] J. F. Caiy, H. Jiz, C. Liuy, and Z. Shenz. High-quality curvelet-based motion deblurring from an image pair. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1566–1573, June 2009.
- [117] Yuanyuan Ding, Scott McCloskey, and Jingyi Yu. Analysis of motion blur with a flutter shutter camera for non-linear motion. In *Proceedings of the 11th European Conference on Computer Vision: Part I, ECCV’10*, pp. 15–30, Berlin, Heidelberg, 2010. Springer-Verlag.
- [118] Neel Joshi, Sing Bing Kang, C. Lawrence Zitnick, and Richard Szeliski. Image deblurring using inertial measurement sensors. *ACM Trans. Graph.*, Vol. 29, No. 4, pp. 30:1–30:9, July 2010.
- [119] Li Xu and Jiaya Jia. Two-phase kernel estimation for robust motion deblurring. In *Proceedings of the 11th European Conference on Computer Vision: Part I, ECCV’10*, pp. 157–170, Berlin, Heidelberg, 2010. Springer-Verlag.
- [120] S. Zhuo, D. Guo, and T. Sim. Robust flash deblurring. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2440–2447, June 2010.
- [121] L. Zhang, A. Deshpande, and X. Chen. Denoising vs. deblurring: Hdr imaging techniques

- using moving cameras. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 522–529, June 2010.
- [122] U. Schmidt, K. Schelten, and S. Roth. Bayesian deblurring with integrated noise estimation. In *CVPR 2011*, pp. 2625–2632, June 2011.
- [123] X. Chen, X. He, J. Yang, and Q. Wu. An effective document image deblurring algorithm. In *CVPR 2011*, pp. 369–376, June 2011.
- [124] S. McCloskey. Temporally coded flash illumination for motion deblurring. In *2011 International Conference on Computer Vision*, pp. 683–690, Nov 2011.
- [125] D. Krishnan, T. Tay, and R. Fergus. Blind deconvolution using a normalized sparsity measure. In *CVPR 2011*, pp. 233–240, June 2011.
- [126] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Efficient marginal likelihood optimization in blind deconvolution. In *CVPR 2011*, pp. 2657–2664, June 2011.
- [127] W. Li, J. Zhang, and Q. Dai. Exploring aligned complementary image pair for blind motion deblurring. In *CVPR 2011*, pp. 273–280, June 2011.
- [128] Ayan Chakrabarti and Todd Zickler. Depth and deblurring from a spectrally-varying depth-of-field. In *Proceedings of the 12th European Conference on Computer Vision - Volume Part V, ECCV’12*, pp. 648–661, Berlin, Heidelberg, 2012. Springer-Verlag.
- [129] H. Ji and K. Wang. A two-stage approach to blind spatially-varying motion deblurring. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 73–80, June 2012.
- [130] Hojin Cho, Jue Wang, and Seungyong Lee. Text image deblurring using text-specific properties. In *Proceedings of the 12th European Conference on Computer Vision - Volume Part V, ECCV’12*, pp. 524–537, Berlin, Heidelberg, 2012. Springer-Verlag.
- [131] Zhe Hu and Ming-Hsuan Yang. Good regions to deblur. 2012.
- [132] H. Zhang and Y. Chen. A sparseland model for deblurring images in the presence of impulse noise. In *2012 19th IEEE International Conference on Image Processing*, pp. 3077–3080, Sept 2012.
- [133] Y. W. Tai, X. Chen, S. Kim, S. J. Kim, F. Li, J. Yang, J. Yu, Y. Matsushita, and M. S. Brown. Nonlinear camera response functions and image deblurring: Theoretical analysis and practice. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 35, No. 10, pp. 2498–2512, Oct 2013.
- [134] Sen Wang, Tingbo Hou, John Border, Hong Qin, and Rodney Miller. High-quality image deblurring with panchromatic pixels. *ACM Trans. Graph.*, Vol. 31, No. 5, pp.

120:1–120:11, September 2012.

- [135] Y. W. Tai and S. Lin. Motion-aware noise filtering for deblurring of noisy and blurry images. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 17–24, June 2012.
- [136] Xiang Zhu, Filip Šroubek, and Peyman Milanfar. Deconvolving psfs for a better motion deblurring using multiple images. In *Proceedings of the 12th European Conference on Computer Vision - Volume Part V, ECCV'12*, pp. 636–647, Berlin, Heidelberg, 2012. Springer-Verlag.
- [137] U. Schmidt, C. Rother, S. Nowozin, J. Jancsary, and S. Roth. Discriminative non-blind deblurring. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 604–611, June 2013.
- [138] T. Yue, J. Suo, X. Ji, and Q. Dai. Optical computing system for fast non-uniform image deblurring. In *2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 937–944, June 2013.
- [139] L. Zhong, S. Cho, D. Metaxas, S. Paris, and J. Wang. Handling noise in single image deblurring using directional filters. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 612–619, June 2013.
- [140] Shicheng Zheng, Li Xu, and Jiaya Jia. Forward motion deblurring. In *Computer Vision (ICCV), 2013 IEEE International Conference on*, pp. 1465–1472. IEEE, 2013.
- [141] Yosuke Bando, Henry Holtzman, and Ramesh Raskar. Near-invariant blur for depth and 2d motion via time-varying light field analysis. *ACM Transactions on Graphics (TOG)*, Vol. 32, No. 2, p. 13, 2013.
- [142] H. Zhang, D. Wipf, and Y. Zhang. Multi-image blind deblurring using a coupled adaptive sparse prior. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1051–1058, June 2013.
- [143] Muhammad Hanif and Abd-Krim Seghouane. Blind image deblurring using non-negative sparse approximation. In *Image Processing (ICIP), 2014 IEEE International Conference on*, pp. 4042–4046. IEEE, 2014.
- [144] Scott McCloskey. Improved motion invariant deblurring through motion estimation. In *European Conference on Computer Vision*, pp. 75–89. Springer, 2014.
- [145] Daniele Perrone and Paolo Favaro. Total variation blind deconvolution: The devil is in the details. 2014.
- [146] Tao Yue, Sunghyun Cho, Jue Wang, and Qionghai Dai. Hybrid image deblurring by

- fusing edge and power spectrum information. In *European Conference on Computer Vision*, pp. 79–93. Springer, 2014.
- [147] Z. Hu, S. Cho, J. Wang, and M. H. Yang. Deblurring low-light images with light streaks. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3382–3389, June 2014.
- [148] Yipin Zhou and Nikos Komodakis. A map-estimation framework for blind deblurring using high-level edge priors. In *European Conference on Computer Vision*, pp. 142–157. Springer, 2014.
- [149] Tomer Michaeli and Michal Irani. Blind deblurring using internal patch recurrence. In *European Conference on Computer Vision*, pp. 783–798. Springer, 2014.
- [150] Ali Mosleh, JM Pierre Langlois, and Paul Green. Image deconvolution ringing artifact detection and removal via psf frequency analysis. In *European Conference on Computer Vision*, pp. 247–262. Springer, 2014.
- [151] Jinshan Pan, Zhe Hu, Zhixun Su, and Ming-Hsuan Yang. Deblurring face images with exemplars. In *European Conference on Computer Vision*, pp. 47–62. Springer, 2014.
- [152] Lu Fang, Haifeng Liu, Feng Wu, Xiaoyan Sun, and Houqiang Li. Separable kernel for image deblurring. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pp. 2885–2892. IEEE, 2014.
- [153] Heesoo Myeong, Stephen Lin, and Kyoung Mu Lee. Alpha matting of motion-blurred objects in bracket sequence images. Vol. 8691, September 2014.
- [154] Sung Hee Park and Marc Levoy. Gyro-based multi-image deconvolution for removing handshake blur. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pp. 3366–3373. IEEE, 2014.
- [155] Xinxin Zhang, Ronggang Wang, Yonghong Tian, Wenmin Wang, and Wen Gao. Image deblurring using robust sparsity priors. In *Image Processing (ICIP), 2015 IEEE International Conference on*, pp. 138–142. IEEE, 2015.
- [156] Ali Mosleh, Paul Green, Emmanuel Onzon, Isabelle Begin, and JM Pierre Langlois. Camera intrinsic blur kernel estimation: A reliable framework. In *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*, pp. 4961–4968. IEEE, 2015.
- [157] Nora Alkhaldi and Joab Winkler. Blind image deconvolution using the sylvester resultant matrix. In *Image Processing (ICIP), 2015 IEEE International Conference on*, pp. 784–788. IEEE, 2015.
- [158] Lei Xiao, Felix Heide, Matthew O’Toole, Andreas Kolb, Matthias B Hullin, Kyros Kutu-

- lakos, and Wolfgang Heidrich. Defocus deblurring and superresolution for time- of-flight depth cameras. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2376–2384, 2015.
- [159] Mingli Zhang, Christian Desrosiers, Caiming Zhang, and Mohamed Cheriet. Effective document image deblurring via gradient histogram preservation. In *Image Processing (ICIP), 2015 IEEE International Conference on*, pp. 779–783. IEEE, 2015.
- [160] Shuochen Su and Wolfgang Heidrich. Rolling shutter motion deblurring. In *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*, pp. 1529–1537. IEEE, 2015.
- [161] Wei-Sheng Lai, Jian-Jiun Ding, Yen-Yu Lin, Yung-Yu Chuang, et al. Blur kernel estimation using normalized color-line priors. In *CVPR*, pp. 64–72, 2015.
- [162] Jianping Shi, Li Xu, and Jiaya Jia. Just noticeable defocus blur detection and estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 657–665, 2015.
- [163] Feng Liu Long Mai. Kernel fusion for better image deblurring. *CVPR2015*, 2015.
- [164] Jian Sun, Wenfei Cao, Zongben Xu, Jean Ponce, et al. Learning a convolutional neural network for non- uniform motion blur removal. In *CVPR*, pp. 769–777, 2015.
- [165] Mauricio Delbracio and Guillermo Sapiro. Burst deblurring: Removing camera shake through fourier burst accumulation. In *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*, pp. 2385–2393. IEEE, 2015.
- [166] Tae Hyun Kim and Kyoung Mu Lee. Generalized video deblurring for dynamic scenes. *arXiv preprint arXiv:1507.02438*, 2015.
- [167] Ziyang Ma, Renjie Liao, Xin Tao, Li Xu, Jiaya Jia, and Enhua Wu. Handling motion blur in multi-frame super- resolution. In *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*, pp. 5224–5232. IEEE, 2015.
- [168] Haichao Zhang and Jianchao Yang. Intra-frame deblurring by leveraging inter-frame camera motion. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4036–4044, 2015.

付録

- ・ 中間審査 (2017 年 11 月 16 日) 発表資料
- ・ 公聴会 (2018 年 2 月 1 日) 発表資料
- ・ ブレ除去のトレンドの推移について

中間審査 (2017 年 11 月 16 日) 発表資料

予備審査
2017/11/16

モーションブラーを含むシーンの
ビデオ映像からの
高精度な画像切り出し手法

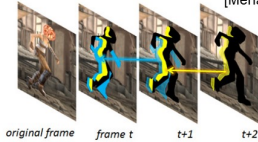
(Technique of high-resolution object
recovery for video with motion blur)

Background

- Demand: capturing out door scene
→occlusion



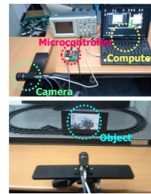
- Common technique: separate foreground and background
[Menandro2014(ECCV)]



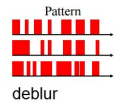
Previous technique

- In So Kweon (ICCV2015)
Use special camera

Hardware Setup



input



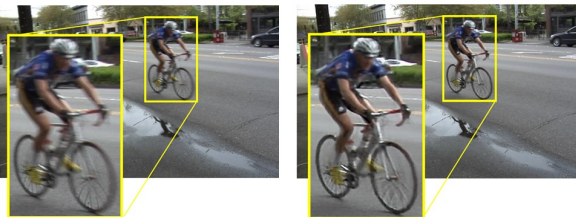
deblur



Previous technique

- Kyoung Mu Lee (CVPR2015)
Use conventional camera

Difficult to restore the complex shape



Purpose

- Deblur for
 - Scene with occlusion
 - Video input
- Recover boundary and texture

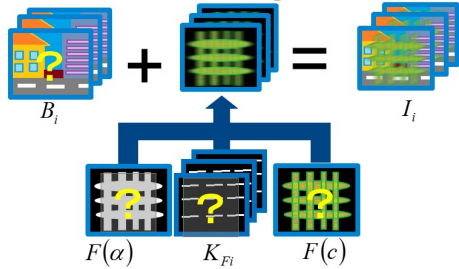
CONTRIBUTIONS

- Estimate correct object motion
- Estimate correct object segmentation
- Increase the experiment scenes and conducted numerical evaluation.

Key idea

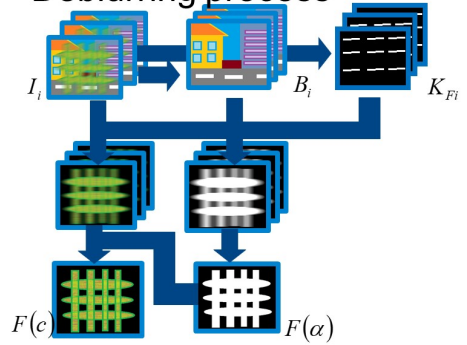
- Improvement of background estimation
- Make consistent alpha map

Motion blurring process

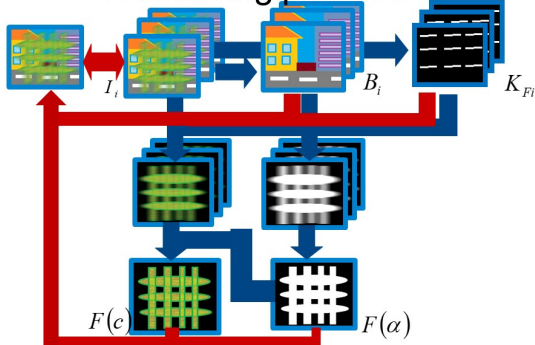


$$I_i = (1 - K_{Fi} * F(\alpha)) * B_i + K_{Fi} * (F(\alpha) \circ F(c)) \quad (r, g, b \in c)$$

Deblurring process

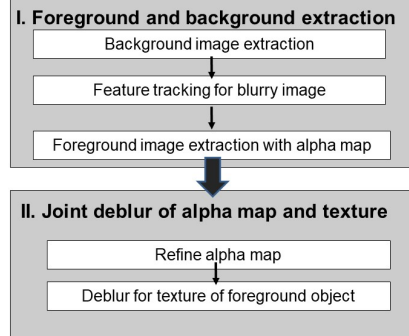


Deblurring process



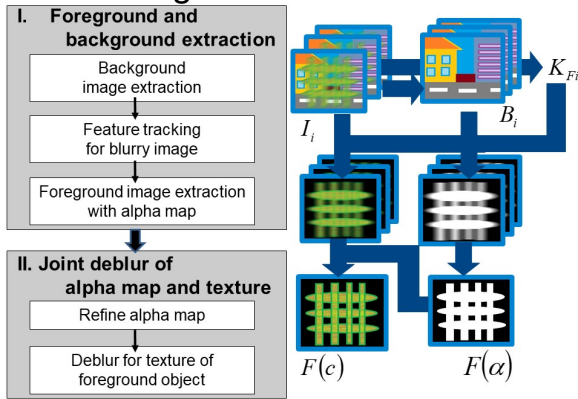
$$\arg \min_{F(c)} \sum_i \sum_j \|I_{ij} - (1 - K_{Fi} * F(\alpha)) \circ B_i + K_{Fi} * (F(\alpha) \circ F(c))\|^2 + \lambda \|CF(\alpha)\|^2$$

Algorithm overview



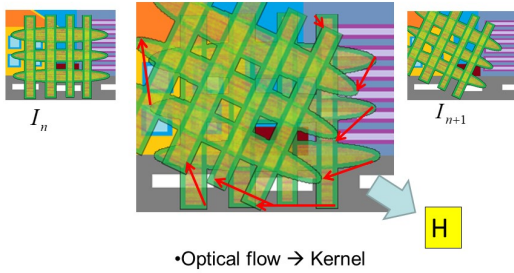
$$\arg \min_{F(c)} \sum_i \sum_j \|I_{ij} - (1 - K_{Fi} * F(\alpha)) \circ B_i + K_{Fi} * (F(\alpha) \circ F(c))\|^2 + \lambda \|CF(\alpha)\|^2$$

Algorithm overview



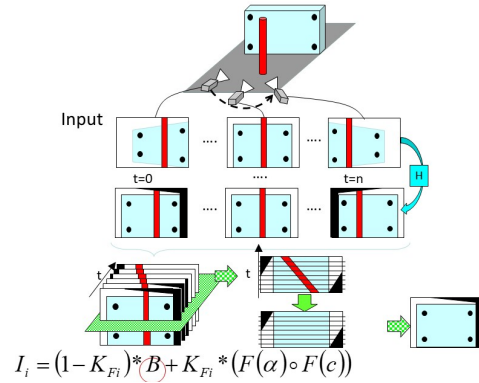
Feature tracking for blurry image

$$I_i = (1 - K_{Fi}) * B + K_{Fi} * (F(\alpha) \circ F(c))$$



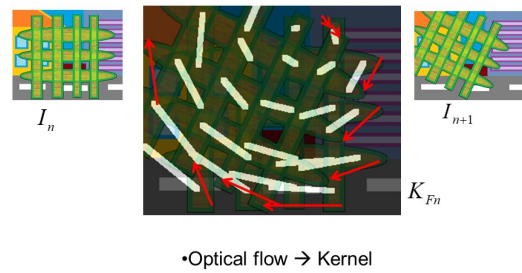
Background estimation

[Kunbayashi, K. Ikeuchi, et al., "Online System of Foreground Obstacle Removal from On-Vehicle Video Using Plane Structure Constraints", ITS symposium2010]



Feature tracking for blurry image

$$I_i = (1 - K_{Fi}) * B + K_{Fi} * (F(\alpha) \circ F(c))$$



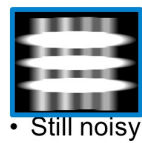
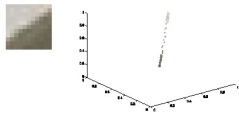
Initial alpha estimation

$$I_i = (1 - K_{Fi}) * B + K_{Fi} * (F(\alpha) \circ F(c))$$



- alpha estimation(Levin 2006)

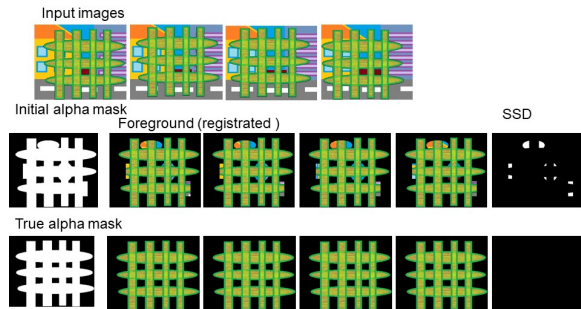
$$\{C_i \in R^3 \mid C_i = \beta_i C_1 + (1 - \beta_i) C_2\}$$



Optimize alpha using multiple images

$$E = \sum_t \sum_{i,j} \{Foreground(t,i,j) + wBackground(t,i,j)\} + \lambda P()$$

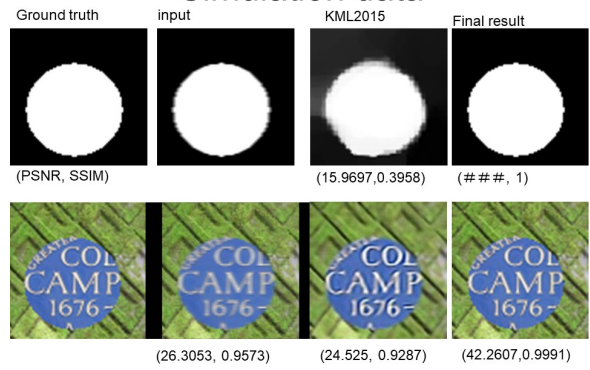
The total SSD of registered pixels The total number of edge pixels in the alpha mask



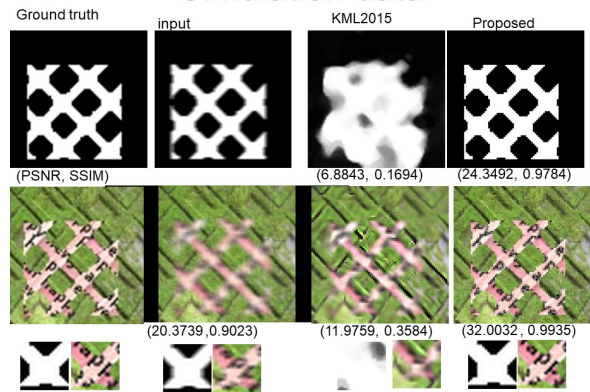
Experimental result



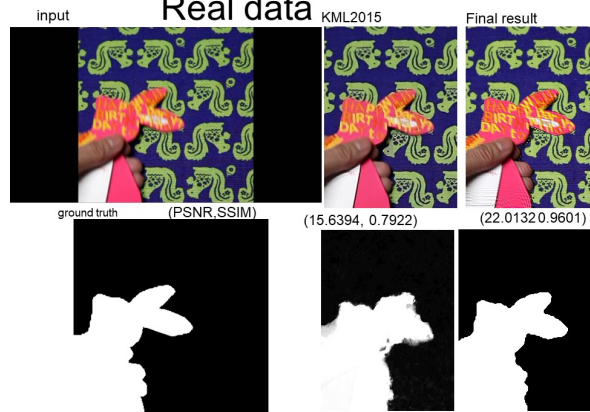
Simulation data

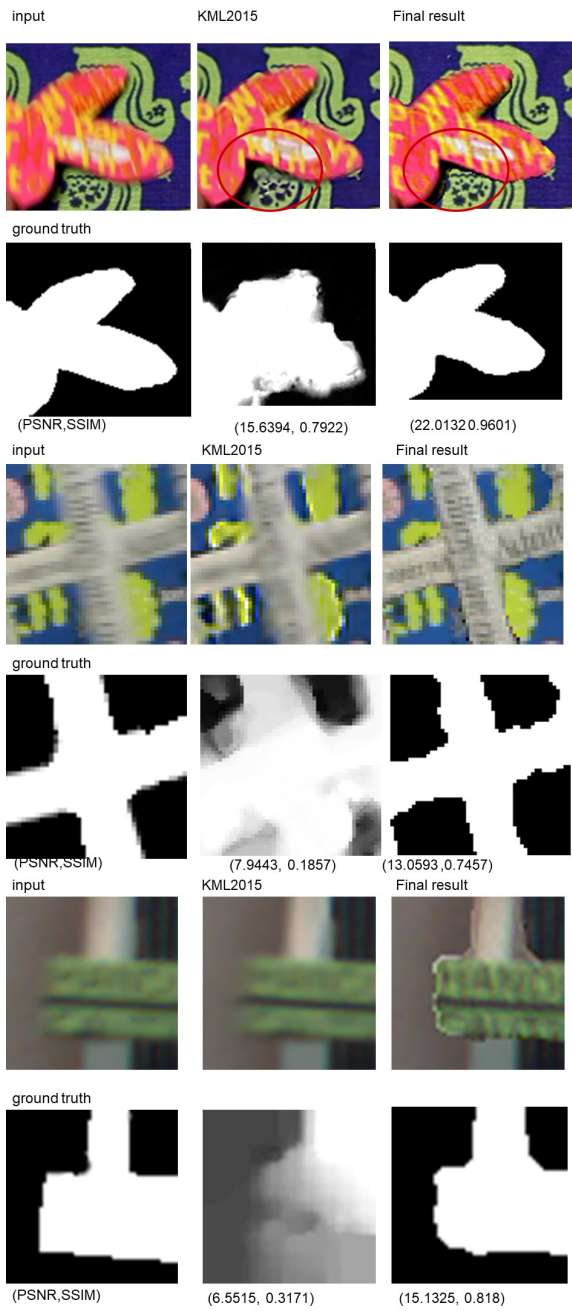


Simulation data

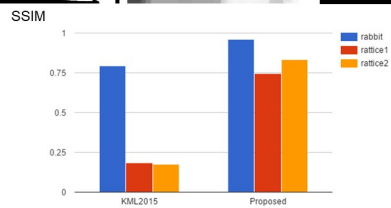
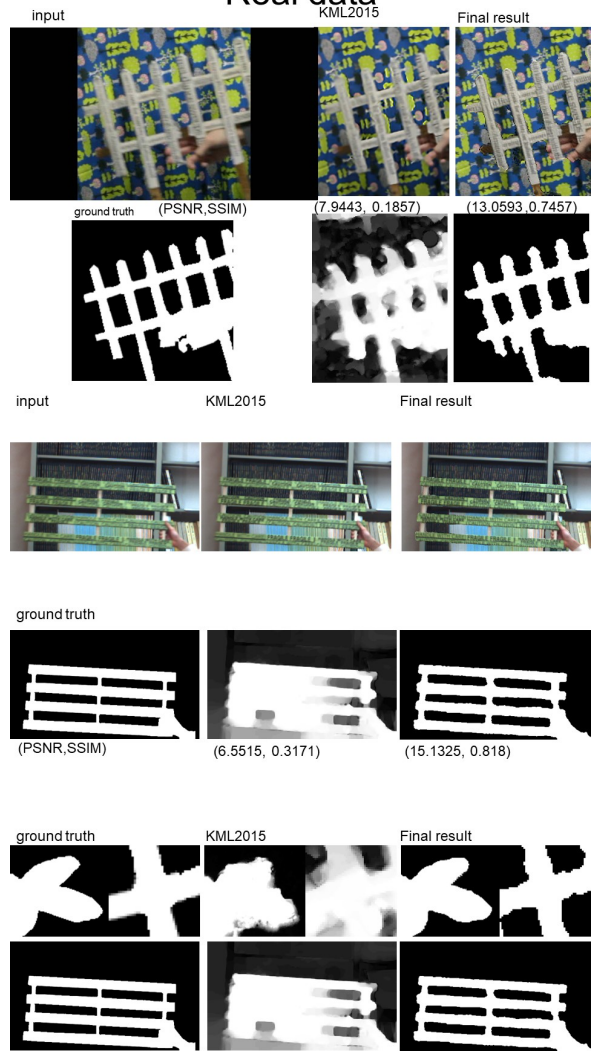


Real data





Real data



Conclusion

- We propose a method which realizes both boundary shape and texture recovery, and successfully restore.

Future Work

- More robust regularization in alpha mask refinement

論文目録

(d)

(1) 著者 松下侑輝, 福元伸也, 川崎洋, 山口拓真, 古川亮, 福田悠人
論文題目「3次元形状を含むシーンにおける手ぶれ映像の超解像処理」
第14回画像の認識・理解シンポジウム(MIRU2011), pp.951-958, 7.2011

(2) 著者 Yuki Matsushita, Hiroshi Kawasaki,
Shintaro Ono, Katsushi Ikeuchi
論文題目「Joint Object Boundary Recovery and Foreground
Image Deblur for Video Including Moving Objects」
第18回画像の認識・理解シンポジウム(MIRU2015), 2, 7.2015

論文目録

主論文

(c)

(1) 著者 Yuki Matsushita, Hiroshi Kawasaki,
Shintaro Ono and Katsushi Ikeuchi
論文題目「Simultaneous Deblur and Super-resolution Technique for
Video Sequence Captured by Hand-held Video Camera」
2014年10月 Proc. ICIP(Paris) pp. 4562-4566 (第4, 5, 6章)

(2) 著者 Yuki Matsushita, Hiroshi Kawasaki, Teruhisa Takano,
Shintaro Ono and Katsushi Ikeuchi
論文題目「Joint technique of fine object boundary recovery and
foreground image deblur for video including moving objects」
2017年5月 Proc.QCAV (Tokyo) Vol.10338 10338 OR
doi: 10.1117/12.2286746 (第7章)

論文と章の対応

1章導入

2章関連研究

3章手法の概要

4章シーケンス中の動きの導出、セグメンテーション MIRU2011

5章動きの高精度化・シーケンスの統合 ICIP2014

6章セグメンテーション高精度化 QCAV2017

7章画像処理 MIRU2015

8章実験結果

9章まとめ

モーションブラーを含むシーンの
ビデオ映像からの
高精度な画像切り出し手法
(Technique of high-resolution object
recovery for video with motion blur)

システム情報科学専攻
松下侑輝

Purpose

- Deblur for
 - Video input
 - Scene with occlusion



- Recover boundary and texture

Related Work

Super-resolution
Reconstruction-base

- Farsiu(ICIP2003)



Learning-base

- Chao(ECCV2014)

3 layer-CNN



Not suppose
blurred scene

Background

Demand:
capturing movie

- blur
- low-resolution
- occlusion
(moving object)

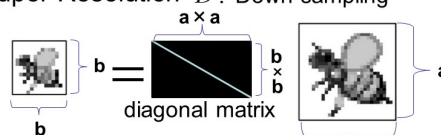


Related Work

I : Input image
 D : Degrade matrix
 L : Latent image

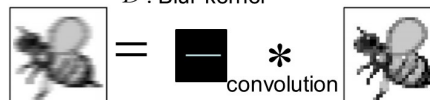
$$I = DL$$

- Super-Resolution D : Down-sampling



- Deblur

D : Blur kernel



Related Work

Blind deconvolution

- Qi Shan(SIGGRAPH2008)

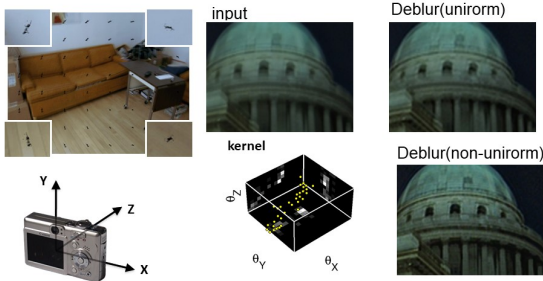


Related Work

Non-uniform blurkernel

Only camera shake

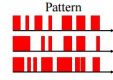
- Oliver Whyte(CVPR2010)



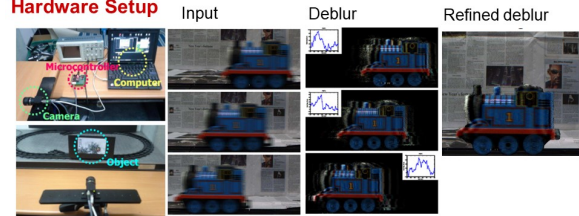
Related Work

Motion deblurring

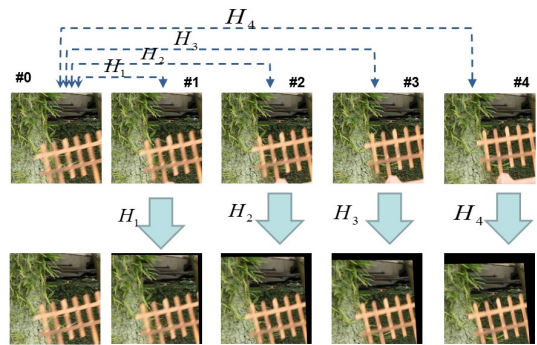
- Agrawal(SIGGRAPH 2009)
 - In So Kweon (ICCV2015)
- Use special camera



Hardware Setup



Homography matrix



Related Work

Motion deblurring (Conventional Camera)

- Kyoung Mu Lee (CVPR2015)

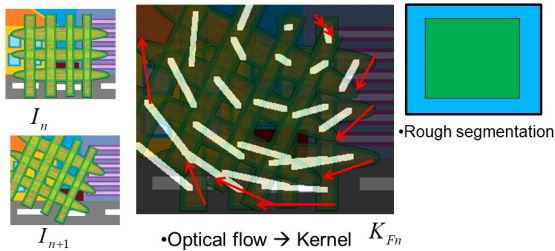
Optical-Flow based segmentation

Difficult to restore the complex shape and occlusion



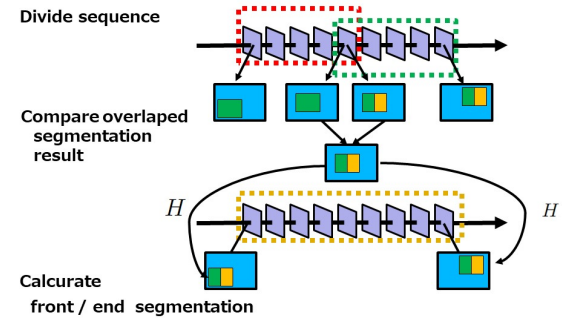
Feature tracking for blurry image

$$I_i = (1 - K_{F_i}) * B + K_{F_i} * (F(\alpha) \circ F(c))$$



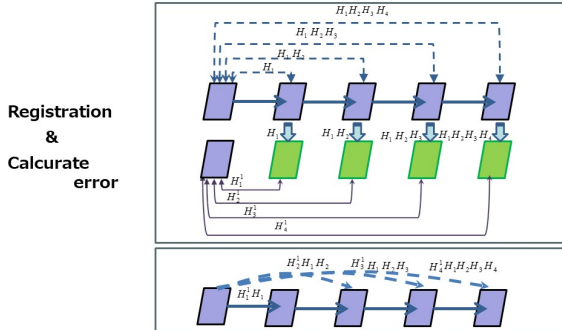
Feature tracking for blurry image

Merge sequence



Feature tracking for blurry image

Refine tracking

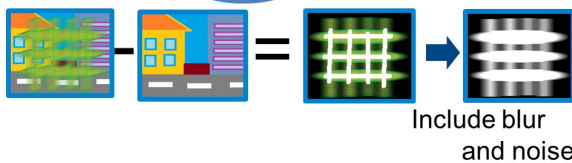


Initial alphamap estimation

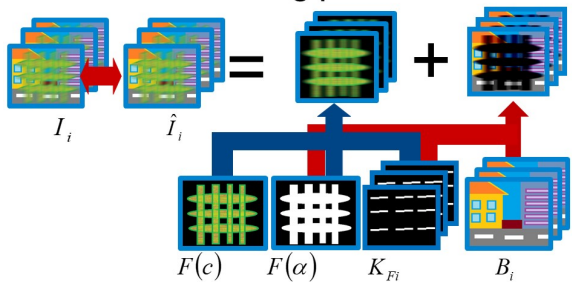
- alpha estimation (Levin 2006)



$$I_i = (1 - K_{Fi}) * B + K_{Fi} * (F(\alpha) \circ F(c))$$



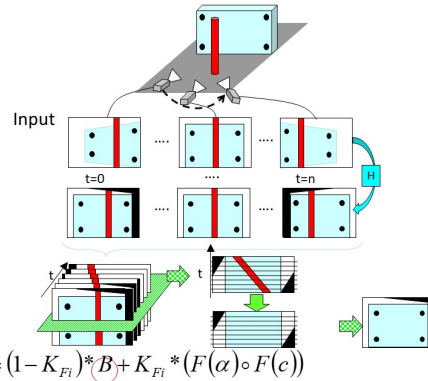
Deblurring process



$$\arg \min_{F(c)} \sum_i \sum_j \|I_{ij} - (1 - K_{Fi}) * B + K_{Fi} * (F(\alpha) \circ F(c))\|^2 + \lambda \|CF(\alpha)\|^2$$

Background estimation

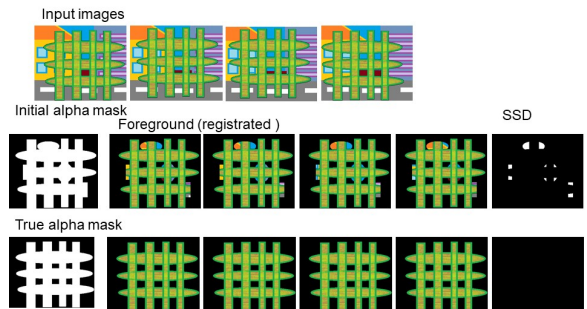
[Kunbayashi, K. Ikeuchi, et al., "Online System of Foreground Obstacle Removal from On-Vehicle Video Using Plane Structure Constraints", ITS symposium 2010]



Optimize alphamap using multiple images

$$E = \sum_t \sum_{i,j} \{Foreground(t, i, j) + wBackground(t, i, j)\} + \lambda P()$$

The total SSD of registered pixels The total number of edge pixels in the alpha mask



Inspection of tracking refinement



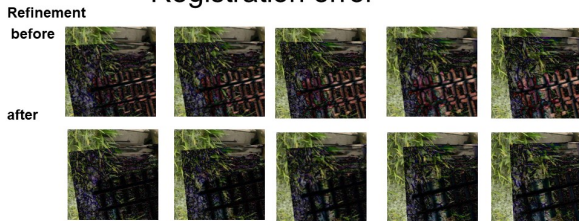
Registration error



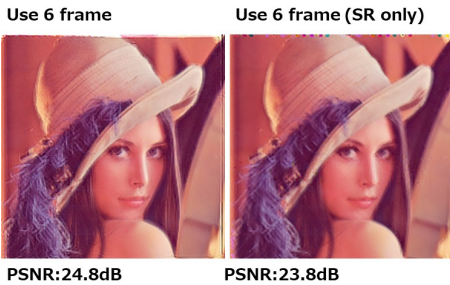
Inspection of tracking refinement



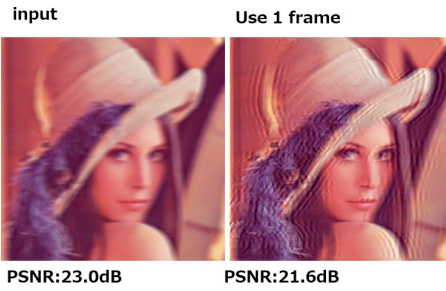
Registration error



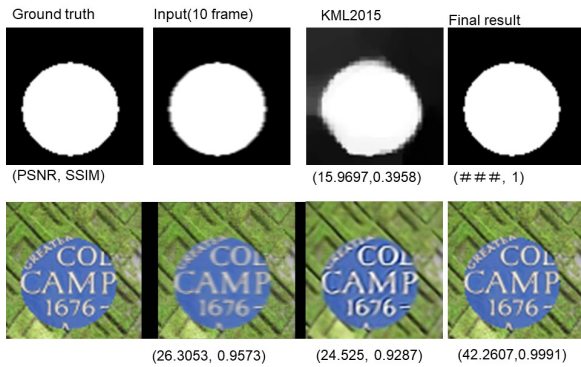
Inspection of number of frame

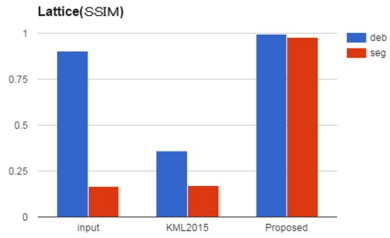


Inspection of number of frame

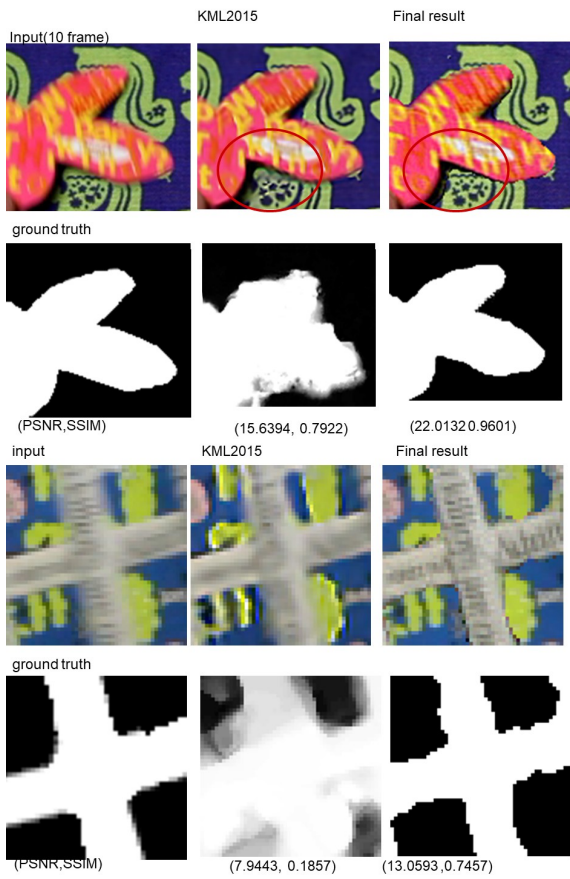
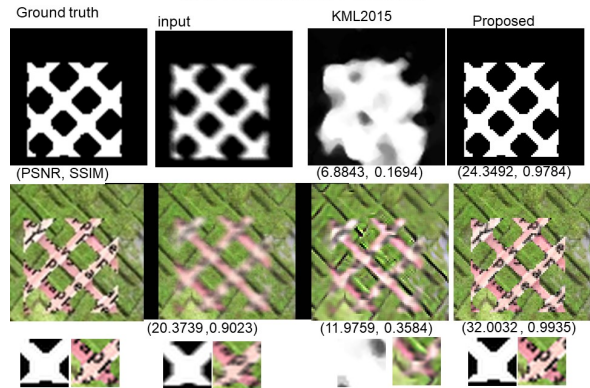


Inspection of alphamap refinement

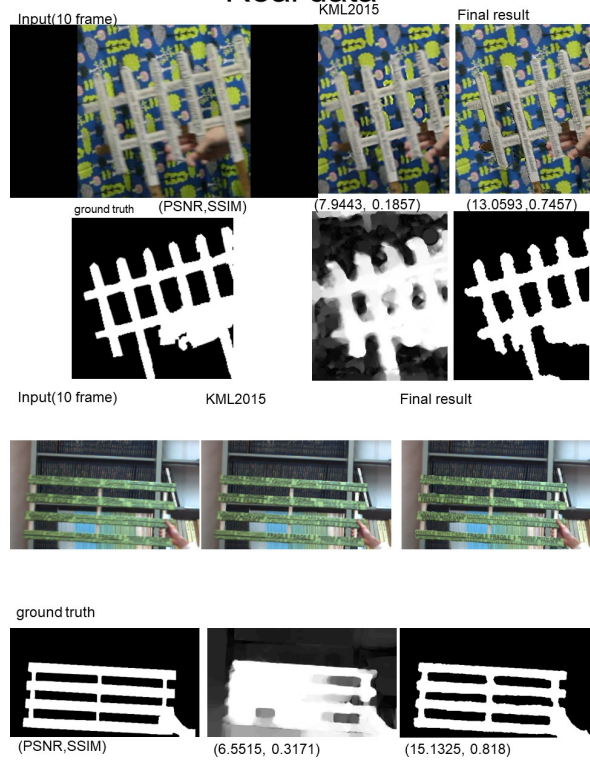


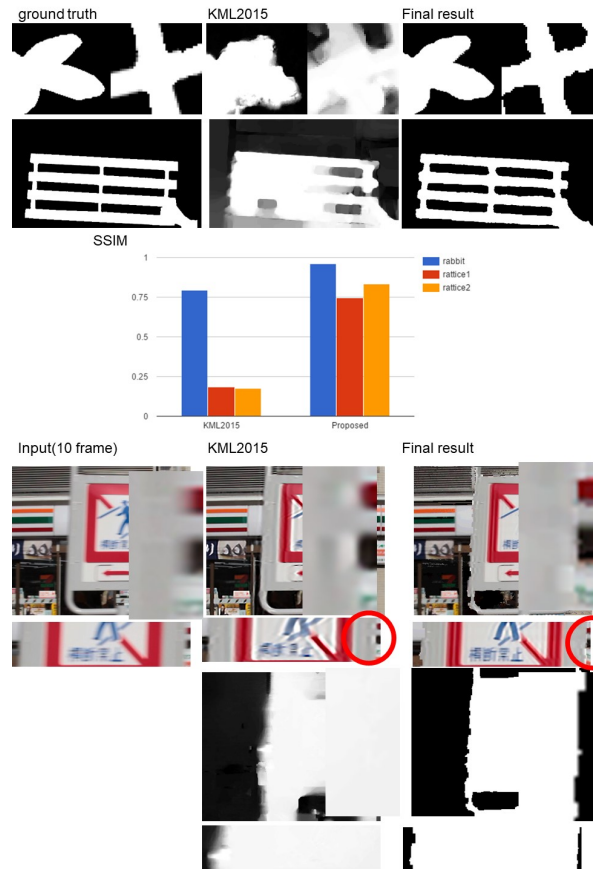
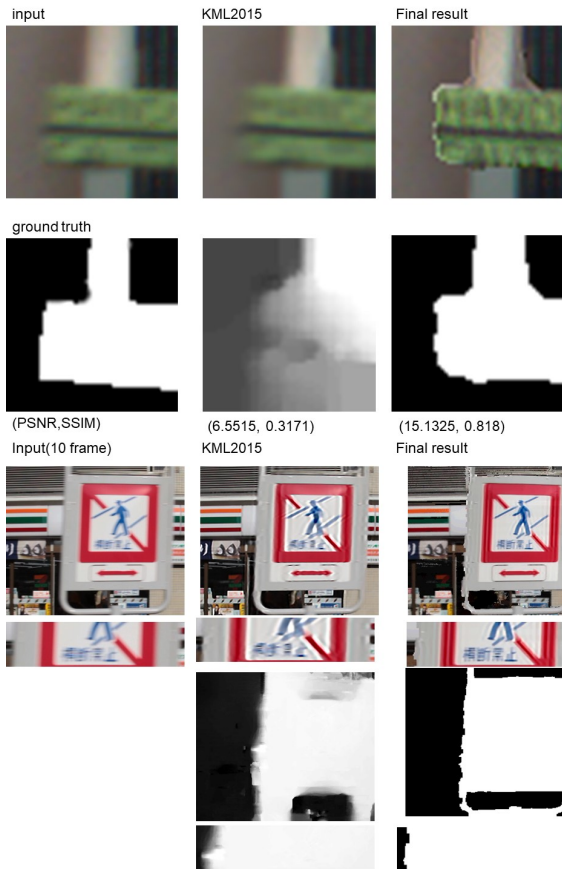


Simulation data



Real data





Discussion

- The more number of input frame, the better processing result.
- Robust alphamap estimation to complex shape object.
- Using background intensity, reduce ringing artifact.

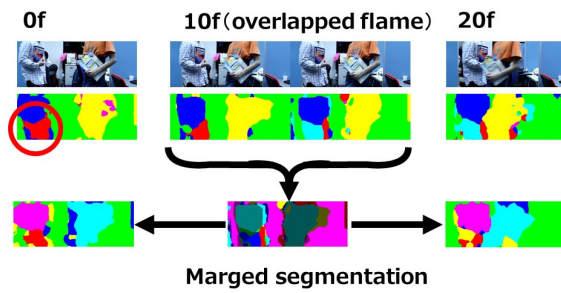
Conclusion

- We propose a method which realizes both boundary shape and texture recovery, and successfully restore.

Future Work

- More robust regularization in alpha mask refinement
- Multi-plane Model

Multi-Plane scene -segmentation



Support slides

Optimization (SR)

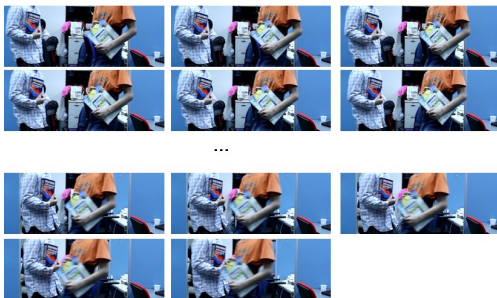
$$\hat{x} = \arg \min_{\mathbf{x}} \left\{ \sum_{k=1}^N \sum_{i=1}^P \|y_i - DB_k i W_k x_i\|^1 + \lambda \|C\mathbf{x}\|^2 \right\}$$

y: latent image
N: image number **x**: input image
P: pixel number **D**: downsampling
B: blur
W: scene move

Blur kernel and registration are Calculated b^H

Multi-Plane scene -input image

Input 20frame



Multi-Plane scene -input image

Input





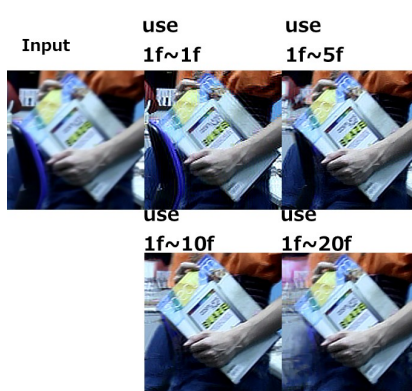
Multi-Plane scene -result Left book



Multi-Plane scene -result Right book



Multi-Plane scene -result Right book



ブレ除去のトレンドの推移について

2008年以前からブレ除去の手法は提案されていたが、2009年を境に発表される論文の数が増加している。第2章で述べたように対象とするブレやシーンの状況も単数フレームを用いて手ブレ等の一様なブレをとるものから、回転や動物体等のようなシーン中に一様でないブレを複数フレームを用いて取るように複雑化する傾向にある。以下の表にブレ除去のトレンドの推移を使用するフレームの数、及び対象とするブレの種類別にまとめる。

2006	単数カーネル	複数カーネル
単数 フレーム	[43] [70]	
複数 フレーム		

2008	単数カーネル	複数カーネル
単数 フレーム	[112] [113]	
複数 フレーム	[114] [37]	

2010	単数カーネル	複数カーネル
単数 フレーム	[53] [117] [119]	[71] [118] [59]
複数 フレーム	[120]	[121]

2007	単数カーネル	複数カーネル
単数 フレーム		[108] [109]
複数 フレーム	[110] [111]	

2009	単数カーネル	複数カーネル
単数 フレーム	[115] [57] [69]	
複数 フレーム	[45] [116]	

2011	単数カーネル	複数カーネル
単数 フレーム	[122] [123] [124] [125] [126]	[72]
複数 フレーム	[127]	

2012	単数カーネル	複数カーネル
単数 フレーム	[128]	[129]
	[130]	[131]
	[132]	
	[133]	
	[134]	
複数 フレーム	[135]	
	[136]	
2014	単数カーネル	複数カーネル
単数 フレーム	[143]	[15]
	[18]	[144]
	[145]	[146]
	[147]	
	[148]	
	[149]	
	[150]	
	[151]	
複数 フレーム	[152]	[79]
		[22]
		[153]
		[154]

2013	単数カーネル	複数カーネル
単数 フレーム	[137]	[138]
	[139]	[83]
		[140]
		[141]
複数 フレーム	[142]	[76]

2015	単数カーネル	複数カーネル
単数 フレーム	[155]	[156]
	[157]	[158]
	[159]	[160]
	[161]	[162]
	[163]	[164]
複数 フレーム	[165]	[166]
	[167]	[168]