

市街地映像データベースを用いた 自車位置推定に関する研究

内 山 寛 之

目次

| | | |
|-------|--------------------------------|----|
| 第 1 章 | 序論 | 1 |
| 1.1 | はじめに | 1 |
| 1.2 | 本研究の目的と意義 | 3 |
| 1.3 | 本論文の構成 | 5 |
| 第 2 章 | 市街地映像データベースと自車位置推定技術 | 7 |
| 2.1 | はじめに | 7 |
| 2.2 | 市街地映像データベースの形態・周辺技術 | 7 |
| 2.3 | 市街地映像データベース中の移動物体の問題 | 11 |
| 2.3.1 | 移動物体による弊害 | 11 |
| 2.3.2 | 映像中の移動物体除去に関する研究 | 12 |
| 2.4 | 自車位置推定技術 | 13 |
| 2.4.1 | GPS 測位 | 13 |
| 2.4.2 | 姿勢センサ・速度センサによる自律航法 | 14 |
| 2.4.3 | レンジファインダによる自車位置推定 | 14 |
| 2.4.4 | カメラによる自車位置推定 | 15 |
| 2.5 | まとめ | 20 |
| 第 3 章 | 市街地映像データベース中の移動物体除去 | 21 |
| 3.1 | はじめに | 21 |
| 3.2 | 映像からの移動物体除去手法 | 21 |
| 3.2.1 | 提案手法の基本戦略 | 21 |
| 3.2.2 | 映像間のレジストレーション | 23 |
| 3.2.3 | 移動物体を含まない部分画像の選択 | 28 |

| | | |
|-------|---------------------------------------|----|
| 3.2.4 | モザイクキング | 31 |
| 3.3 | 実験 | 32 |
| 3.3.1 | 実験条件 | 32 |
| 3.3.2 | 実験結果 | 34 |
| 3.4 | 考察 | 35 |
| 3.4.1 | 実験結果に対する考察 | 38 |
| 3.4.2 | 本手法の対象とする範囲 | 42 |
| 3.4.3 | 最適なサブウィンドウの大きさについての考察 | 45 |
| 3.4.4 | 移動物体除去に用いる映像の違いの影響 | 46 |
| 3.4.5 | 複数映像の収集方法 | 47 |
| 3.4.6 | メディアンの選択手法の考察 | 47 |
| 3.5 | まとめ | 49 |
| 第4章 | 全方位市街地映像データベースと車載通常カメラ映像の照合による自転車位置推定 | 51 |
| 4.1 | はじめに | 51 |
| 4.2 | 提案手法の基本戦略 | 52 |
| 4.2.1 | 従来手法を異種カメラ間対応付けに適用する場合の問題点 | 52 |
| 4.2.2 | 基本戦略 | 54 |
| 4.3 | 異種カメラ間の対応付け手法 | 55 |
| 4.3.1 | 画像間の照合 | 55 |
| 4.3.2 | 映像間の照合コスト | 57 |
| 4.3.3 | 傾斜制限 | 58 |
| 4.3.4 | 拡張 DP マッチングのアルゴリズム | 58 |
| 4.3.5 | データベース映像の正規化 | 60 |
| 4.4 | 実験 | 61 |
| 4.4.1 | 実験条件 | 61 |
| 4.4.2 | 実験結果 | 63 |
| 4.5 | 考察 | 63 |
| 4.6 | カメラ方向が未知の場合の異種カメラ間対応付け | 68 |
| 4.6.1 | 手法概要 | 68 |

| | | |
|-------|-----------------------------|-----|
| 4.6.2 | 実験 | 69 |
| 4.7 | 移動物体除去映像の自車位置推定への適用 | 73 |
| 4.7.1 | 概要 | 73 |
| 4.7.2 | 実験条件 | 73 |
| 4.7.3 | 結果と考察 | 74 |
| 4.8 | まとめ | 76 |
| 第5章 | 2台の車載カメラを用いた車線方向を含む走行位置推定 | 79 |
| 5.1 | はじめに | 79 |
| 5.2 | 2台の車載カメラを用いた位置推定手法 | 80 |
| 5.2.1 | 基本戦略 | 80 |
| 5.2.2 | 画像間の照合 | 81 |
| 5.2.3 | カメラ間の拘束条件を用いた照合結果の統合 | 83 |
| 5.2.4 | 車両の挙動による対応付けの制限 | 84 |
| 5.2.5 | 照合結果の統合のアルゴリズム | 86 |
| 5.2.6 | 三角測量による走行位置の計算 | 86 |
| 5.2.7 | データベース映像の正規化 | 86 |
| 5.3 | 実験 | 87 |
| 5.3.1 | 実験条件 | 87 |
| 5.3.2 | 実験結果 | 89 |
| 5.4 | 考察 | 89 |
| 5.4.1 | 実験結果に対する考察 | 90 |
| 5.4.2 | 車線方向の位置推定誤差に関する考察 | 92 |
| 5.4.3 | 時間帯による影響 | 94 |
| 5.5 | まとめ | 95 |
| 第6章 | むすび | 97 |
| 6.1 | 総括 | 97 |
| 6.2 | 今後の課題と展望 | 98 |
| 謝辞 | | 101 |
| 付録 A | 市街地映像データベースの拡張現実型画像情報提示への応用 | 103 |

| | | |
|-------|-----------------------------|-----|
| A.1 | 拡張現実型画像情報提示手法 | 105 |
| A.1.1 | SfM による 3 次元点群の取得 | 105 |
| A.1.2 | 壁面の検出 | 106 |
| A.1.3 | 提示位置の選択と提示画像の重畳 | 109 |
| A.2 | 画像情報提示実験 | 109 |
| A.3 | まとめ | 113 |
| 付録 B | 動的計画法による最適化 | 115 |
| 付録 C | メディアン の 定義 | 119 |
| C.1 | 1 次元のメディアン | 119 |
| C.2 | 多次元のメディアン | 120 |
| 参考文献 | | 123 |
| 研究業績 | | 133 |

第 1 章

序論

1.1 はじめに

市街地における高精度な自車位置推定は，高度道路交通システム（ITS: Intelligent Transport Systems）を実現するうえで最も重要な要素の 1 つである．例えば，カーナビゲーションシステムにおいて，詳細な自車位置を知ることが，ドライバを正確に誘導するために必須である．さらに，リアルタイムで撮影した車外映像*に文字情報を重畳表示する実写型カーナビゲーションシステム [1] が提案されるなど，より高精度な位置情報が求められつつある．また，高精度な自車位置推定が実現できれば，車線の逸脱防止，一時停止すべき箇所で警告を行う等の運転者支援を行うこともできるようになる．現在のカーナビゲーションシステムは主に GPS によって自車位置を推定しているが，建物等による電波の反射や遮蔽のため，普及型 GPS には 5～30 m 程度の誤差が含まれる．車速やジャイロ스코ープ，道路形状などを利用して，推定位置の修正が行われることがあるが，上記の機能を実現するためには 1 m 程度の位置推定精度が必要であり，精度が不十分である．

古来，正確な自己位置を知ることが，人や物が移動するうえで次に進むべき方向を決定するための最も重要な要素であり，様々な試みがなされてきた．位置を推定する最も単純な方法は，人が目視により周囲の景色を観察することである．例えば，原始的な航海技法では，近くに見える陸地の地形を観察することにより，自船の位置を得ていた．この方法は「山あて」と呼ばれ，航海士は山や高木などの目印の見える方向から自己位置を特定す

* 本論文では，1 枚の画像を「画像」と表記する．また，多数の「画像」が連続し，系列となったものを，「映像」と表記する．「映像」中の「画像」を特に「フレーム」と表記することもある．

る [2] . 時代が進むと , 羅針盤や六分儀といった測量機器が発明され , 天体をランドマークとすることにより , 地球上のどの場所においても絶対位置が推定できるようになった [3] . このように , 周囲の風景と事前に得られた風景についての知識を照合することで , 自己の位置を得ることが可能である . 近年 , このような原理のもと , 車載カメラ映像と事前に蓄積した風景映像を照合することで , 自車位置を推定する手法が注目されている [4] .

現在 , 市街地の風景映像を収集し , 蓄積したデータベースが広く利用されている [5] . 様々な機関によって , 大規模な位置情報付き市街地映像の収集が実験的あるいは商業的になされている . 本論文では , このように市街地映像と地図や位置情報とを対応付け , 蓄積したデータベースを「市街地映像データベース」と呼ぶ . 市街地映像データベースと車載カメラ映像を照合することによって , 自車位置推定が可能である . 市街地映像データベースの例として , Google ストリートビュー [5] や MapJack[6] がある . これらのサービスはインターネット上で地図と対応付けた全方位市街地映像の閲覧環境を提供している . また , 市街地映像データベースを利用した研究も活発に行われており [7][8] , 自車位置推定もその研究の 1 つである .

高精度な自車位置推定を行うためには , それに適したデータベースを構築する必要がある . 自車位置推定に適する市街地映像データベースの要件として様々なものがあるが , 重要なものとして次の 2 つが挙げられる .

- (A) 撮影対象である市街地の街並みが遮蔽されていない
- (B) 利用の障害となる被写体が映像中に含まれていない

(A) の要件として , 車両のようにその場に恒常的に存在しない物体によって , 背景が遮蔽されていない映像である必要がある . 車両等によって背後が遮蔽されている場合 , その部分の風景に関する画像情報を抽出することができない等の支障が生じる . また , 単純にユーザが見たい部分が遮蔽されるという問題もある . このため , 市街地映像データベース中の背景は遮蔽されていない必要がある . (B) の要件として , 人物の顔や車両のナンバープレート等 , 被写体のプライバシーを侵害するものが映像中に存在しない必要がある . 市街地映像データベースを広く一般に公開する場合 , 映像中の被写体のプライバシーに留意する必要がある .

以上の (A) と (B) の要件は , 映像を扱うデータベースの機能として基本的なものであり , 重要である . そのため , 本論文では (A) と (B) の要件を満たす市街地映像データベー

スを構築する．

1.2 本研究の目的と意義

本研究は，遮蔽のない高品質な市街地映像データベースの構築と，それを利用した高精度な自転車位置推定手法の開発を目的とする．

先に述べたように，市街地映像データベースを利用することで高精度な自転車位置推定を実現できる．しかしながら，撮影した映像を単純に格納したデータベースには，車両等による背景の遮蔽や被写体のプライバシーの問題など，数多くの問題が存在する．また，市街地映像データベースに新たな機能を追加することで，利便性をさらに向上させることが可能である．そのため，市街地映像データベースをその利用方法に適するように編集・加工することは，実利用を想定するうえで必要不可欠である．そこで本論文では，遮蔽のない高品質な市街地映像データベースの構築と，高精度な自転車位置推定手法の開発の両面から研究を行う．

遮蔽のない高品質な市街地映像データベースとして，市街地映像中に移動物体の存在しないデータベースを構築する．移動物体を除去することで，先に述べたような背景の遮蔽やプライバシーの問題の解決が可能である．また，移動物体の存在しない市街地映像は，自転車位置推定のみならず，路上障害物検出 [7] 等でも必要とされており，市街地の画像を対象とした研究全般に有用である．

具体的には，本論文では以下の内容について述べる．

- (a) 市街地映像データベース中の移動物体除去（第 3 章）
- (b) 全方位市街地映像データベースと車載通常カメラ映像の照合による自転車位置推定（第 4 章）
- (c) 2 台の車載カメラを用いた車線方向を含めた走行位置推定（第 5 章）

さらに本論文では，市街地映像データベースの一応用として，市街地映像データベース上へ画像情報を提示する手法を示す．ユーザに適切に情報提示を行うことは，ナビゲーションを行う上で必要不可欠である．この画像情報提示手法については，付録 A に収録した．以下，(a)～(c) の研究について，概要を簡単に述べる．

(a) 市街地映像データベース中の移動物体除去

市街地映像データベース中の車両・人物等の移動物体を除去する手法を提案する．市街地映像データベース中の車両や人物には，先に述べたような遮蔽やプライバシーの問題がある．映像中の移動物体を除去することで，これらの問題の解決が可能である．さらに，(b)，(c) で述べる自転車位置推定において，画像照合の障害となる市街地映像データベース中の移動物体を事前に除去しておくことで，位置推定精度の向上が期待できる．

従来，映像中の不要な物体を除去する手法が多く提案されている [9][10][11]．しかしながら，これらの手法は市街地を経路に沿って撮影した映像を適用対象としていないため，市街地映像データベース中の移動物体の除去には適さない．そこで，同一経路を複数回走行して得られた映像を統合することで，市街地映像データベース中の移動物体を効果的に除去する手法を提案する．

(b) 全方位市街地映像データベースと車載通常カメラ映像の照合による自転車位置推定

市街地映像データベースの運転者支援への利用の 1 つに自転車位置推定がある．本研究では，市街地映像データベース中の全方位カメラ映像と車載通常カメラ映像を照合することで，自転車位置を高精度に推定する手法を提案する．市街地映像データベース構築時には広い視野を一度に撮影できる全方位カメラが適する．一方，全方位カメラは上部に突き出した特殊な形状をしており，カメラの構造的な観点から，自転車位置推定を行う一般車には通常カメラが適する．しかしながら従来，データベース構築時と自転車位置推定時で同一種類のカメラを用いたアピアランスベースの自転車位置推定手法 [12] は提案されているが，異種のカメラを用いた自転車位置推定にこれらの手法の枠組みを適用することは，カメラの視野角が異なるため，困難である．そのため，カメラの視野角が異なっても，高精度に映像間の照合が可能な手法を提案する．さらに，(a) の移動物体除去手法を適用した市街地映像データベースを利用することで，位置推定精度が向上することを示す．

(c) 2 台の車載カメラを用いた車線方向を含めた走行位置推定

(b) の自転車位置推定手法は，進行方向の位置情報のみの取得が可能であり，車線方向の位置は取得することができない．そこで，この手法に対し，車線方向の位置も取得可能となるように拡張を行う．具体的には，自転車位置推定時に 1 台の車載通常カメラではなく，2 台の車載通常カメラを利用する手法を提案する．また，2 台のカメラの照合結果を統合

的に利用し，2台のカメラをそれぞれ独立に扱うよりも高精度に自車位置推定を行う手法を提案する．1台のみでは画像照合が不確かな区間でも，2台のカメラの情報を統合的に利用し，補完しあうことで，位置推定精度が向上させることが可能となる．

1.3 本論文の構成

本論文は6章から構成される．第1章は序論であり，本論文の研究背景と目的，意義を述べた．次に，第2章では市街地映像データベースと自車位置推定に関する周辺技術を紹介する．第3章では，市街地映像データベース中の移動物体除去手法について述べる．第4章では，全方位カメラ映像と車載通常カメラ映像の照合による自車位置推定を説明する．また，第3章の移動物体除去手法を適用した市街地映像データベースを用いて自車位置推定を行う実験について述べる．第5章では，この手法を拡張した2台の車載通常カメラを用いて自車位置推定を行う手法を説明する．最後に，第6章で総括と今後の課題，展望を述べる．

第 2 章

市街地映像データベースと自転車位置推定技術

2.1 はじめに

本章では、本論文の研究対象である市街地映像データベースと自転車位置推定に関する周辺技術を紹介する。特に、第 1 章でも触れた市街地映像データベース中の移動物体の問題と、カメラを用いた自転車位置推定については、詳しい説明を加える。

以降、初めに 2.2 で市街地映像データベースの様々な形態や周辺技術を紹介する。次に、2.3 で市街地映像データベース中の移動物体により生じる問題と、映像中の移動物体除去に関する研究について述べる。最後に、2.4 で自転車位置推定技術の概要について述べる。

2.2 市街地映像データベースの形態・周辺技術

市街地映像と地図を対応付けたり、市街地をモデリングしようとする試みがある。本論文では、市街地映像や市街地モデルと地図や位置情報とを対応付けたものを「市街地映像データベース」と呼ぶ(図 2.1)。

本節では、市街地映像データベース中の画像の表現方法、市街地データの収集、データベースの更新について述べる。また、市街地映像データベースの運転者支援や情報提示への応用についても紹介する。

(a) データベース中の画像の表現方法

データベース中の画像の表現方法は、大きく分けて次の 3 つに分類される。

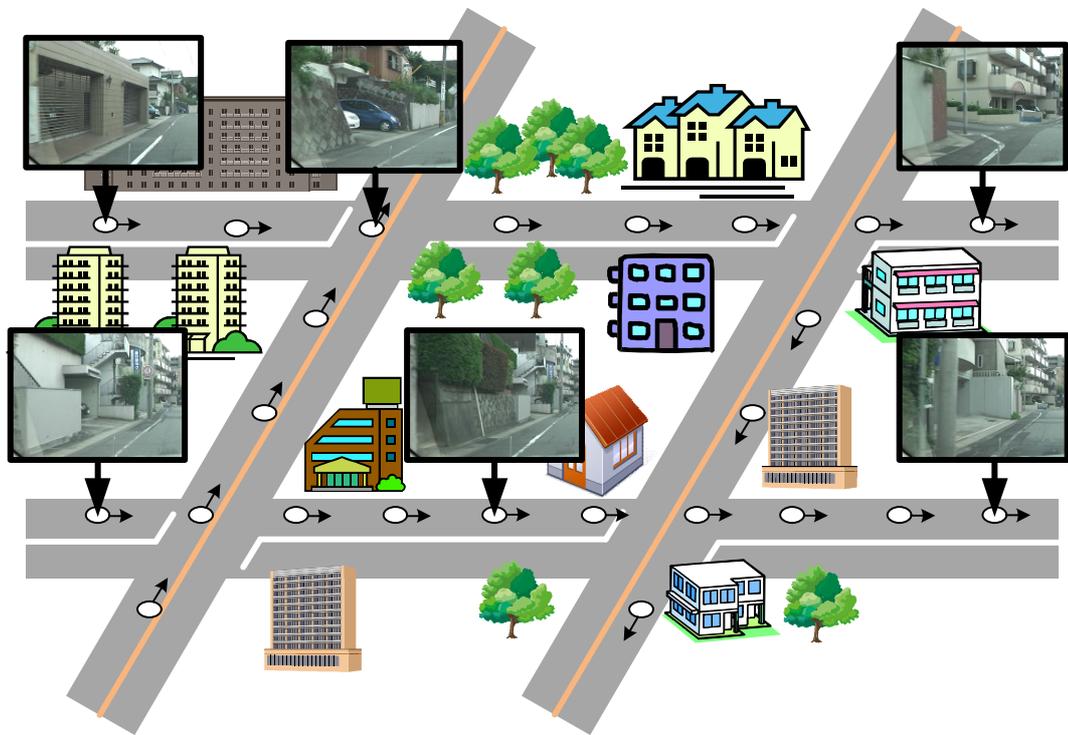


図 2.1 市街地映像データベース

- 未加工の画像 [5][13]
- 加工した画像 [14][15][16]
- 3次元モデル [8][17][18][19]

未加工の画像を利用する試みとして，Google ストリートビュー [5] や MapJack[6] は車載カメラで撮影した全方位カメラ画像をインターネット上で閲覧できるサービスを提供している．また，インターネット上の画像投稿サイトである Flickr[13] や Panoramio[20] には，GeoTag と呼ばれる撮影位置情報が付加された画像が投稿されている．これらの画像投稿サイトから取得した位置情報付き市街地画像を利用した研究もなされており [18]，これらも市街地映像データベースの一種といえる．

加工した画像を利用する試みとして，Zheng ら [14] はラインカメラを用いて車両の側方をスキャンしていくことで，経路に沿ったパノラマ画像を取得し，街並みを再現する手法を提案している．この手法は，側方の 1 ラインのみをスキャンするため，Google ストリートビューのように地点ごとに画像を記憶する場合と比べ大幅にデータ量が削減さ

れる．また，Microsoft[15] は多数の Web 投稿画像を連結し，連結画像として提示する Photosynth というサービスを公開している．その他，実写画像ベースで画像の生成を行う image based rendering を用い，車載全方位カメラ映像から任意視点の映像を生成する研究も行われている [16] ．

市街地映像から 3 次元情報を取り出し，仮想現実に応用する研究が行われている．川崎ら [17] は 2 次元地図上の建物の輪郭と実際の市街地映像を比較することによって，都市の 3 次元形状を復元する手法を提案している．また，石川ら [21] はカメラからの色情報とレーザスキャナからの形状情報を組み合わせ，都市をモデリングする手法を提案している．Agarwal ら [18] は多数の Web 投稿画像から，都市の 3 次元構造を復元する手法を提案している．また，Google ストリートビューのようなパノラマ画像から 3 次元構造を復元する手法も提案されている [8][19] ．

(b) 市街地データの収集

データベース構築のための市街地データ収集の方法として，次の 3 つが利用される．

- 専用車両 [22][23] ．
- 一般車両 [24][25]
- インターネット上の投稿画像 [18][26] ．

市街地の映像を収集することを目的として，高精度に自車位置推定を行う機器と，市街地情報収集のためのセンサを搭載した専用車両が開発されている．データ収集専用の機器を搭載した車両を用いることで，高品質なデータ収集が可能となる．アジア航測(株)[22] は「GeoMaster」というシステムを開発している．このシステムは，慣性航法装置 (INS)，DGPS (Differential GPS)，速度計 (オドメータ) による複合完成測量装置 (HISS) により自車位置を推定し，単眼およびステレオカメラにより道路周辺物体の 3 次元位置情報・画像情報を収集する．また，三菱電機(株) [23] は慣性計測装置 (IMU)，車速パルス，搬送波位相/コード GPS を用いて数 10cm の精度で位置情報を収集しながら，単眼カメラおよびレーザスキャナで市街地情報を収集する「モバイルマッピングシステム」を開発している．

一方，特別な専用車両ではなく，一般車両が自由に走行することで，データベース構築を目指す研究も行われている．一般車両を用いた研究として，佐藤ら [24] や Enami

ら [25] は安価なビデオカメラと普及型 GPS を搭載した車両が市街地を巡回することで、データベース更新のための画像を収集する手法を提案している。佐藤らはさらに、複数回の走行で収集したデータ中の同一地点を画像照合で見つけ出し、その地点における複数回の普及型 GPS の観測を平均化することによって、普及型 GPS の位置推定誤差を減少させる手法を提案している。

その他、インターネット投稿画像を利用した画像収集も行われている [18][26]。近年、GPS 付きのカメラの普及により、位置情報を付加した画像の撮影が容易になっていることが背景にある。インターネット投稿画像を利用した画像収集は、観光地等の写真が大量に撮影される場所では投稿画像の収集が容易であり、収集コストも小さいという利点がある。しかし、それ以外の場所では網羅的に画像を収集することが難しいという問題点がある。

(c) データベースの更新

市街地映像は建物の新築や増改築によって変化していくため、定期的にデータベースの更新を行う必要がある。佐藤ら [24] は、車載の全方位カメラを用い、ある地点における過去と現在の市街地画像の類似度の変化を比較することにより、街並みの変化を検出する手法を提案している。Enami ら [25] や Ishizawa ら [27] は車載の通常カメラを用い、車両の側方を観測することによって街並み変化を検出する手法を提案している。また、文献 [22] では、モバイルマッピングシステムを用いた都市の変化状況の検知について検討している。

一方、車載カメラを用いない市街地映像データベースの更新の研究も行われている。七里ら [28] は、車載全方位カメラで撮影した市街地映像データベースの更新を目的とし、ユーザが市街地で撮影したスナップ写真を自動的に市街地映像データベース中の該当部分に貼り込むシステムを開発している。

(d) 運転者支援への応用

市街地の映像は運転者支援にも用いられている。小島ら [29] は衝突防止システム等の運転者支援を目的とし、車載カメラ画像の道路形状とデジタル道路地図を照合することにより、車両の姿勢および道路の 3 次元形状の推定精度を向上させる手法を提案している。Kyutoku ら [7] は事前に記憶しておいたある地点における過去の映像と現在の映像との差



図 2.2 映像中の移動物体 .

分をとることにより，路上障害物を検出する手法を提案している．位置情報付きの市街地映像データベース中の画像と車載カメラ画像の照合を行うことで，自車位置推定を行う手法も提案されている [24]．その他に，カーナビゲーションシステムにおいて実画像を用いて情報提示を行う実画像カーナビも検討されている [1]．

(e) データベース画像中への情報提示への応用

市街地映像データベース中に画像や文字等の情報を付加することが行われている．三菱電機（株）[30] は道案内のためのアイコンや矢印を重畳した実画像を，携帯電話に配信するシステムを発表している．Google ストリートビュー等の市街地閲覧サービスでは，地図情報と連動して実画像上に建物の名称等が重畳されている．Snavely ら [26] は複数の視点で撮影された投稿写真を対応付け，様々な視点から撮影対象物を閲覧可能な手法を提案しており，ユーザが画像上に文字情報のアノテーションを行い，他のユーザがそれを参照可能である．

2.3 市街地映像データベース中の移動物体の問題

2.3.1 移動物体による弊害

市街地には，自動車（一時的に駐停車しているものを含む），自転車，歩行者等が多数存在する．本論文では，このようなその場に恒常的に存在しないものを移動物体と呼ぶ．

画像中に映りこんだ移動物体が市街地映像データベースを利用する上で大きな問題となる（図 2.2）．例えば，市街地映像データベースを市街地閲覧サービスに利用する場合，見

たい部分を移動物体が遮蔽するという問題がある。自転車位置推定や市街地映像データベースとの差分による路上障害物検出 [7] といった運転者支援への応用でも、移動物体が画像処理対象となる背景を遮蔽するため、移動物体のない市街地映像データベースが必要である。また、都市のモデリングを目的として画像から市街地の 3 次元復元を行う場合にも、移動物体は存在しないことが望ましい。

また、実画像を広く一般に公開するには、画像中のプライバシーに留意することが求められる。例えば、Google ストリートビューでは画像中に人物の顔や車両のナンバープレートが映りこむことが問題となっている [31][32]。そのため、国によっては、プライバシー侵害の懸念のある市街地画像の公開の規制が検討されている他、訴訟に発展した事例もある [33]。プライバシー保護のために、顔やナンバープレートを検出し、ぼかす処理が行われている。しかしながら、検出漏れが多く、現状では十分でない。また、単に顔をぼかしただけでは、身元を隠すのに不十分であるという指摘もある。人や車両そのものを画像中から除去することによって、プライバシー侵害が懸念される物体の存在しない市街地映像データベースが利用可能となる。

2.3.2 映像中の移動物体除去に関する研究

映像中の移動物体除去手法は、以下の 4 つのアプローチがある。

- 同一地点で異なる時刻に撮影した複数の静止画を利用する手法 [9][34][35][36]
- 単一の静止画を利用する手法 [10][37][38]
- 動画像を利用する手法 [11][39][40][41]
- 環境の 3 次元情報を利用する手法 [42]

画像中から不要な物体を除去する手法として、同一地点で異なる時刻に撮影した複数の静止画を利用するものが提案されている。移動物体は特定の地点に長く停留することは少ないため、特定の地点を一定時間連続して観察した場合、移動物体よりも背景の出現頻度が高くなる。そのため、文献 [9][34] の手法ではこの仮定に基づき、投票により背景画像を選択する。また、文献 [36] の手法では、照度の大きい室内では人物の輝度が低いという知見を基に、輝度の高い画素を選択することにより背景を選択している。しかし、これらの手法は移動カメラを対象とした手法ではないため、大規模な市街地映像データベース構築への直接的な適用は困難である。

単一の静止画や動画像（映像）中から不要な物体を除去するインペインティングと呼ばれる手法がある．これらの手法は，あらかじめ人手や自動検出等により除去対象領域を指定する．そして，除去対象領域内の見た目が自然となるように，除去対象領域の周囲の画像を利用して除去対象領域内を補間する．補間の際，単一の静止画を利用する手法では画像を 2 次元画像として扱い，動画像を利用する手法では動画像を時空間画像として扱う．しかしながら，これらの手法では除去対象領域をあらかじめ人手や自動検出等により指定する必要がある他，除去対象領域の検出処理が手法中に含まれている．例えば，Wexler ら [11] の手法は人手で除去対象領域を指定することを前提としており，Flores ら [41] の手法では，人物除去の事前処理として，検出器による人物検出を行っている．膨大な市街地映像中から人手で除去対象領域を指定することは困難であり，1 つの静止画像や動画像から多様に変化する車両や人物の領域を検出すること自体が 1 つの難しい問題である．また，除去対象領域内の画像を周囲から予測するため，現実とは異なる背景が生成されるという問題がある．

その他に，環境の 3 次元情報を利用する手法も提案されている．Früh ら [42] の手法では，市街地の 3 次元モデルの構築を目的としており，まずカメラによる 2 次元画像とレーザスキャナによる 3 次元情報を同時に取得する．そして，3 次元情報を基に前景を検出・除去し，空白部分を周辺の画素で補完する．しかしながら，前景の検出処理が必要となる他，カメラの他にレーザスキャナを必要とする．

2.4 自転車位置推定技術

本節では GPS，姿勢センサ，速度センサ，レンジファインダ，カメラを用いた自転車位置推定技術を紹介する．

2.4.1 GPS 測位

GPS [43][44] は最低 4 機の人工衛星からの信号を受信することにより自転車位置を推定する．地上の基準局を用いない GPS の単独測位では，5～30 m 程度の位置精度が得られる．現在のカーナビゲーションシステムは主にこれにより自転車位置を推定している．位置推定誤差の原因として，電離層と対流圏遅延による誤差，衛星位置および衛星時計誤差，周囲構造物による遮蔽やマルチパス，受信機誤差などがあり，環境に応じて位置推定誤差

は増大する．精度が不十分なため，カーナビゲーションシステムでは自転車位置が道路上に乗るように補正するマップマッチング [45] という手法が併用されている．一時停止すべき箇所で警告を行う等の高度な運転者支援を実現するためには 1 m 程度の位置推定精度が必要であるが，GPS の単独測位では最も条件が良い場合でも数 m 程度の精度であり [44]，精度が不十分である．

GPS の位置推定精度を高める技術として，地上の基準局を併用した DGPS や RTK-GPS (Real Time Kinematic GPS) などがある．DGPS は位置が正確にわかっている定点に基準局を設置し，通信回線を用いて利用者に補正データを知らせることによって位置推定精度を高めている．これによって，誤差は数 m 程度に抑えられる．また，RTK-GPS はコードの伝搬時間だけでなく，搬送波の位相を測定することにより数 cm の精度を実現する．しかし，衛星電波の瞬断やマルチパスに弱いという欠点のほか，現状では非常に高価なため，一般車両が利用するには課題が多い．

2.4.2 姿勢センサ・速度センサによる自律航法

姿勢センサとしてジャイロスコープや加速度センサ，速度センサとして車速パルスなどがある．ジャイロスコープは車両の相対角度や角速度，加速度センサは車両の傾斜を測定する．これらの計測値を積算すれば，車両の移動量を推定することができる．このように，外部の情報を用いずに自転車の動きを推定することを自律航法と呼ぶ．カーナビゲーションシステムですでに実用化されており，GPS による推定位置を補正するために用いられている．これらのセンサを用いた自転車位置推定には，次第に誤差が蓄積されるという問題がある．

2.4.3 レンジファインダによる自転車位置推定

近年，車両にレーザスキャナやミリ波レーダといったレンジファインダを搭載し，車外環境の認識を行う試みがなされている．レンジファインダとは，対象物との距離を測定する装置である．特に，ここでは物体の 3 次元形状が取得できるものについて言及する．以下に，レーザスキャナ，ミリ波レーダ，ソナーの特徴と，それらを利用した自転車位置推定について述べる．

レーザスキャナはレーザパルスを計測対象物に照射し，その往復時間を計測することで

対象物との距離を計測する．そして，同時に照射方向を計測することで計測対象物の 3 次元座標を取得する装置である．レーザは空間分解能が高く，車両，縁石，標識の柱などの形状を精度よく計測することができる．そのため，移動車両検出 [46]，都市空間の 3 次元モデル化 [47] など，車外環境認識に広く用いられている．例えば，渋谷ら [48] は事前に道路形状の時系列データをデータベースに記憶しておき，入力系列との DP マッチングによる照合で自転車位置を推定する手法を提案している．また，Weiss ら [49] は信号機などの柱をランドマークとした手法を提案している．

ミリ波レーダはミリ波を計測対象に照射し，反射してきた電波の伝搬時間を基に，計測対象との距離を測定する．ドップラー効果による周波数変化の観測から，相対速度が測定可能であり，衝突防止を目的とした利用も検討されている．電波を利用するため，霧や降雪時においても精度低下が小さいという利点がある．Dissanayake ら [50] はミリ波レーダを使用して野外環境でデータベース構築と自転車位置推定を同時に行う手法を提案している．しかし，レーザスキャナに比べ，空間解像度が低いという欠点がある．

ソナーは超音波を計測対象物に照射し，その往復時間を計測することで計測対象との距離を測定する．Tards ら [51] は屋内環境において，ソナーを使用してロボットの自己位置を推定する手法を提案している．しかし，レーザスキャナに比べ空間解像度が低いため，屋外における位置推定には大きな課題がある．

2.4.4 カメラによる自転車位置推定

車両の外界を認識するセンサとして広く利用されており，道路標識認識 [52]，障害物検出 [53]，白線検出 [54]，前方車両監視 [55] など，様々な運転者支援への同時利用が可能である．また，比較的安価なため，一般車への普及が期待できる．

カメラを用いた自転車位置推定は，主に次の 2 つの手法に分けられる．

- ランドマークベース [56][57][58][59][60]
- アピアランスベース [12][24][61][63][65]

(a) ランドマークベース

ランドマークベースの手法は，まず環境に存在するランドマークの 2 次元あるいは 3 次元的位置を記憶しておく．そして，入力カメラ画像中で検出された複数のランドマーク

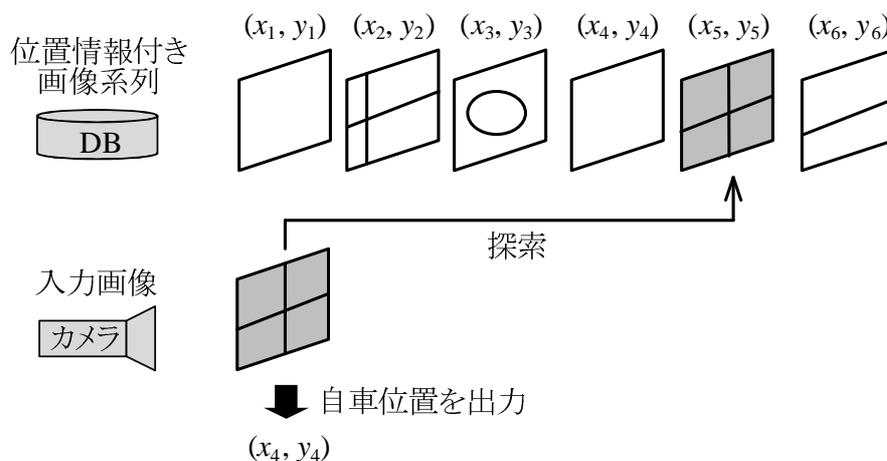


図 2.3 アピアランスベースの自転車位置推定 .

と照合することにより、それらの位置関係から自車両の 2 次元あるいは 3 次元的位置を推定する手法である。ここで、ランドマークとは、環境中のある位置に固定された特徴点である。まず、ランドマークとして人工的に設置した目印 [56] が挙げられる。他に、物体のコーナの位置に関する情報を利用したもの [57] や、SIFT 特徴などの局所記述子を利用したもの [58] など、自然特徴点を利用したものが挙げられる。自転車位置の推定方法として、あるランドマークをステレオで観測する手法 [66] や、複数のランドマークを単眼カメラで観測する手法 [57] がある。人工的なランドマークは設置のために人的コストがかかるという問題があり、自然特徴点を利用したものは、テクスチャの乏しい画像では安定して特徴点を抽出することが困難であるという問題がある。

(b) アピアランスベース

アピアランスベースの手法では、物体の“見え方”を利用する。“見え方”とは、ある地点のある方向から見た画像のことをいう。アピアランスベースの位置推定手法では、図 2.3 のように、あらかじめ画像そのものと撮影地点の位置情報をデータベースに記憶しておく。位置推定時には、入力カメラから得られた入力画像に最も類似する画像をデータベース中から探索することで自転車位置を得る。ランドマークベースの手法ではデータベース中のランドマークに位置情報が付加されている一方、アピアランスベースの手法ではデータベースの画像そのものに位置情報が付加されており、これを利用する。アピアランスベースの手法には被写体の位置を知る必要がないことや、特徴点の抽出・照合が不要で



(a) 全方位カメラと通常カメラ



(b) 全方位カメラ画像



(c) 通常カメラ画像

図 2.4 全方位カメラと通常カメラ。

あるという利点がある．このため，低解像度画像など，画像の品質が低い場合でも比較的頑健に位置推定が可能である．画像照合の特徴量としては，画像そのもの [61]，固有空間法で次元削減をした画像 [67]，エピポーラ平面画像 [63] 等，様々なものが用いられる．

アピランスペースの手法で使用されるカメラとして，全方位カメラと通常カメラがある．全方位カメラとして魚眼レンズで構成されたカメラや，双曲面ミラーで構成されたカメラ [64]，複数の通常カメラのモザイクキングにより全方位映像を構成するものが使用される．双曲面ミラー型の全方位カメラと通常カメラの外観および取得される画像の例を図 2.4 に示す．図 2.4(a) は全方位カメラを車両の屋根に，通常カメラを車内に搭載した例である．(b) の左の画像は全方位カメラで取得した画像であり，右の画像はそれをパ

ノラマ展開した画像である。(c)の画像は通常カメラで取得した画像である。用途に応じて、データベース構築段階、位置推定段階ともに全方位カメラ [61] あるいは通常カメラ [62] を使用する手法が提案されている。データベース構築段階、位置推定段階ともに通常カメラを用いるものは、両段階でカメラ方向を一致させる必要があり、また視野角が狭いため、車両の方向変化に弱いという問題がある。データベース構築段階で全方位カメラを使用する利点として、全方向が一度に撮影できる利点がある。また、死角がないため、車両の方向変化に対応することができる。しかしながら、全方位カメラは上部に突き出した特殊な形状をしており、自転車位置推定時に一般車が搭載するには不向きである。一方、通常カメラの一般車への搭載は容易である。

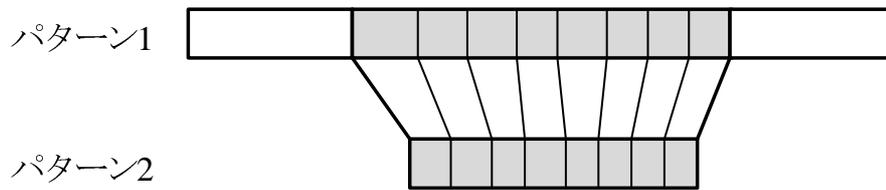
アピアランススペースの自転車位置推定手法では位置推定を画像間の対応付けのみで行う。周囲環境と自転車との幾何的な位置関係は考慮されないため、車両の進行方向の位置の推定は可能だが、車線方向の位置の推定は限定的であるという問題がある。アピアランススペースで 2 次元位置推定が可能な手法の事例として、前田ら [65] は入力、データベースともに全方位カメラを用い、入力画像と同一直線上にない 3 地点上のデータベース画像を照合する手法を提案している。この手法では、画像照合で得られた 3 地点の位置に対し、入力画像とデータベース画像の類似度で重み付け平均することによって、2 次元位置推定を可能としている。しかしながら、データベース中の画像が全て同一直線状に並んでいる画像配置では、この手法の適用は困難である。

画像照合方法としては、DP マッチング [70] を導入することによって時系列情報を利用し、位置決め精度を向上させる手法が数多く提案されている。DP マッチングは非線形な伸縮を許容したパターン間の照合手法である。DP マッチングを基にした映像間照合手法は、自転車位置推定だけでなく、動画検索 [71]、ジェスチャ認識 [72] 等に広く用いられている。単一の画像フレーム同士による照合ではなく、複数のフレーム間で照合を行うことにより、画像間の照合精度が向上する。

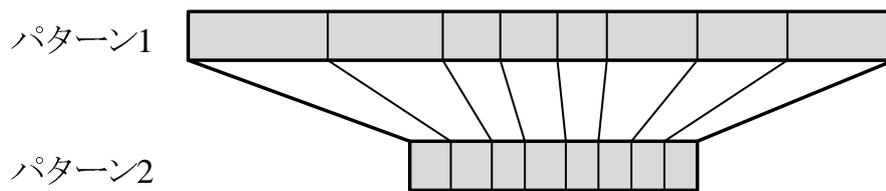
DP マッチングを用いた位置決め手法は、映像の開始点と終了点の扱いによって、大きく分けて次の 2 つの手法に分けられる。

- (a) 映像取得の開始時と終了時の位置は未知である手法 [61][62][63][75]
- (b) 映像取得の開始時と終了時の位置は既知である手法 [24]

手法 (a) は、DP マッチングによるスポッティング認識に基づいた手法によって実現され



(a) スポットティング認識



(b) 開始点・終了点固定の DP マッチング

図 2.5 DP マッチング .

る．スポットティング認識とは，あるパターンに類似した部分を他のパターン中から検出する手法である．例えば図 2.5(a) のように，パターン 1 中からパターン 2 に対応する区間を非線形な伸縮を許容して探索する．自転車位置推定の場合，パターン 1 が市街地映像データベース中の映像，パターン 2 が車載カメラからの入力映像に該当する．パターン 1 の各要素（フレーム）には位置情報が付随しており，パターン 2 の終了点（最新のフレームの取得時）に対応するパターン 1 の要素を見つけることで，最新のフレームの取得時における位置情報を得る．DP マッチングのさらに詳しい原理については，第 4 章の 4.2.1 にて説明する．

スポットティング認識を利用した自転車位置推定手法は，その目的によって様々な種類が提案されている．小野ら [63] や小島ら [61] はスポットティングの一手法である連続 DP [73] を映像間の対応付けに用いている．また，小島ら [61][62] は映像間の複数類似区間の検出が可能な RIFCDP（Reference Interval-free Continuous DP）によるロボットの位置同定システムも提案している．さらに，伊藤ら [74] は RIFCDP の高速化手法である Shift CDP の応用例として，位置推定を挙げている．西村ら [75] は連続 DP の傾斜制限と画像特徴量を改良した Non-monotonic 連続 DP と呼ばれる手法によって，ロボットの向きにかかわらず位置推定を行う手法を提案している．

手法 (b) は，図 2.5(b) のように映像の開始点・終了点を固定して対応付けを行う．事前知識として開始点・終了点の位置を与えるため，開始点・終了点付近の映像間の対応付け精度は手法 (b) より高い．

2.5 まとめ

本章では，市街地映像データベースと自転車位置推定に関する研究と周辺技術を紹介した．特に，本研究の対象とする市街地映像データベース中の移動物体の問題と，カメラによる自転車位置推定については，詳しい説明を加えた．

第 3 章

市街地映像データベース中の移動物体除去

3.1 はじめに

第 2 章で述べたように，市街地映像データベースを利用した研究において，画像中の移動物体による背景の遮蔽が問題となる．また，市街地画像閲覧サービス中の画像には人物の顔や車のナンバープレートも画像に含まれており，これらがプライバシー保護の観点から大きな問題となっている．そのため，移動物体が存在しない画像データベースを用意する必要がある．そこで本章では，自車位置推定のような運転者支援とプライバシー保護のための市街地映像データベースの構築を目的とし，図 3.1 のように走行する車両より得られた全方位カメラ映像から移動物体を除去する手法を提案する．

以降，3.2 で移動物体除去手法について説明し，3.3 で評価実験，3.4 で考察について述べる．

3.2 映像からの移動物体除去手法

3.2.1 提案手法の基本戦略

第 2 章の 2.3.2 で，映像中の移動物体除去の従来研究を，(a) 同一地点で異なる時刻に撮影した複数の静止画を利用する手法，(b) 単一の静止画を利用する手法，(c) 動画像を利用する手法，(d) 環境の 3 次元情報を利用する手法の 4 つに分類して紹介した．第 2 章で述べたように，これらの手法は市街地を経路に沿って撮影した映像を適用対象として

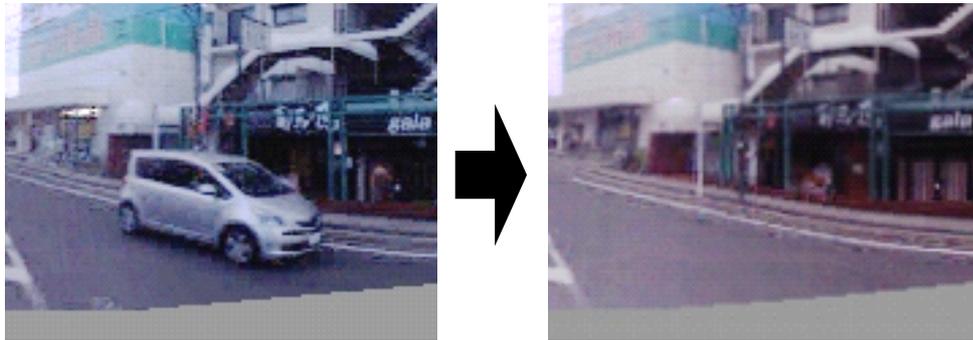


図 3.1 移動物体除去（提案手法の適用例）。

いないため、市街地映像データベース中の移動物体の除去には適さない。そこで本研究では、異なる時刻に撮影した複数の車載カメラ映像を利用することで、移動物体を含まない映像を生成する手法を提案する。複数の映像の収集は、公共の交通機関、運輸業者等の車両にカメラを装着することで実現が可能である。単一の映像ではなく、同一経路の複数の映像を用いることにより、除去対象領域の指定や検出処理が不要な手法を実現する。また、車載カメラで経路に沿って撮影した映像を使用することで、効率的に市街地映像データベースの構築が可能となる。

提案手法の基本戦略を述べる。本論文では、路上に存在する自動車（一時的に駐停車しているものを含む）、自転車、歩行者など、その場に恒常的に存在しないものを移動物体と呼び、それらを除去の対象とする。同一地点で撮影した画像中の同一部分を長時間観察した場合、一般に移動物体が映っている時間はわずかである。そこで、異なる時刻に同一地点で撮影した画像を部分画像に分割し、貼り合わせることで、全体として移動物体を含まない画像を得る（図 3.2）。しかし、実際には、車載カメラにより映像を撮影した場合、走行毎に車両の速度や軌跡が異なるため、映像毎に各フレームの撮影位置は異なる。そのため、映像間のレジストレーション（位置合わせ）処理により、複数の車載カメラ映像から、同一地点で撮影したものと類似する画像群を生成する。

提案手法は合計 N 本の映像を用いる。それぞれの映像は、GPS による位置情報または人手によって、同じ走行区間となるように大まかに切り出されているとする。ここで、移動物体除去の対象となる映像を基準映像、残りの $N - 1$ 本の映像をソース映像と呼ぶことにする。提案手法の流れを図 3.3 に示す。提案手法は、(i) 事前処理、(ii) 映像間のレジストレーション、(iii) 部分画像の選択、(iv) 部分画像を貼り合わせるモザイクキングに

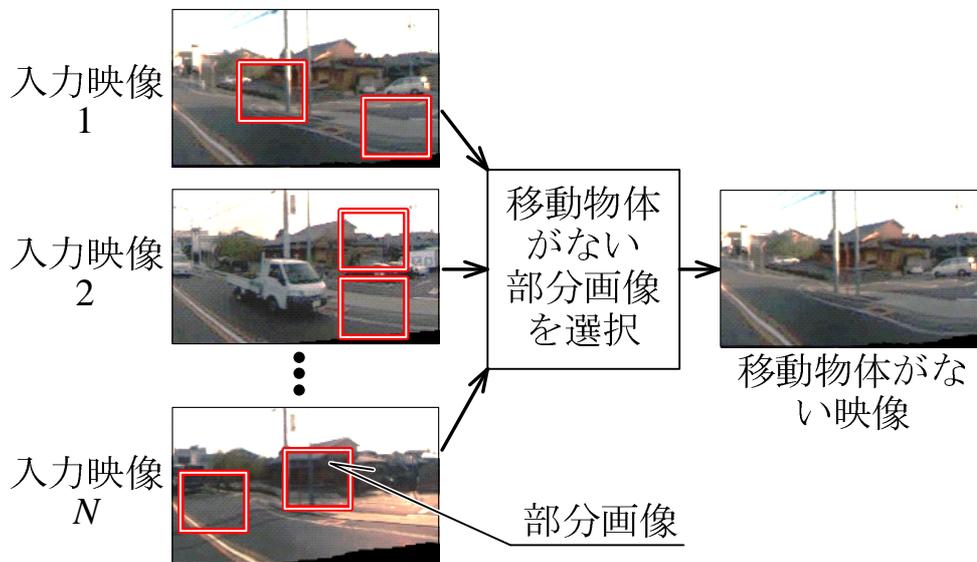


図 3.2 異なる時刻に撮影した同一地点の画像を収集し，移動物体が映っていない部分画像を自動選択して張り合わせるることにより，全体として移動物体が映っていない画像を生成する．

より構成される．カメラの特性に起因し，各映像の明るさや色合いが異なることがある．そこで，まず (i) の事前処理では，映像の各フレームに色ヒストグラム正規化処理を適用し，画像全体の明るさや色合いを正規化する．次に，(ii) の基準映像とソース映像間の時間的・空間的な位置合わせを行うレジストレーション処理により，複数の車載カメラ映像から，同一地点で撮影したものと類似する画像群を生成する．(iii) の移動物体を含まない部分画像の選択処理では，異なる時刻・同一地点の部分画像の中から，最も背景らしい部分画像を選択する．最後に，(iv) のモザイクキング処理により，選択された部分画像の貼り合わせを行う．以下，映像間のレジストレーション，部分画像の選択，部分画像のモザイクキングの詳細について説明する．

3.2.2 映像間のレジストレーション

映像を撮影する車両は移動するため，それぞれの映像フレームの撮影位置は異なる．そのため，異なる映像間では，映像フレームのインデックス番号が同じでも，同じ撮影位置とは限らない．これを解決するために，映像間のレジストレーション処理により，複数の車載カメラ映像から，同一地点で撮影したものと類似する画像群を生成する．

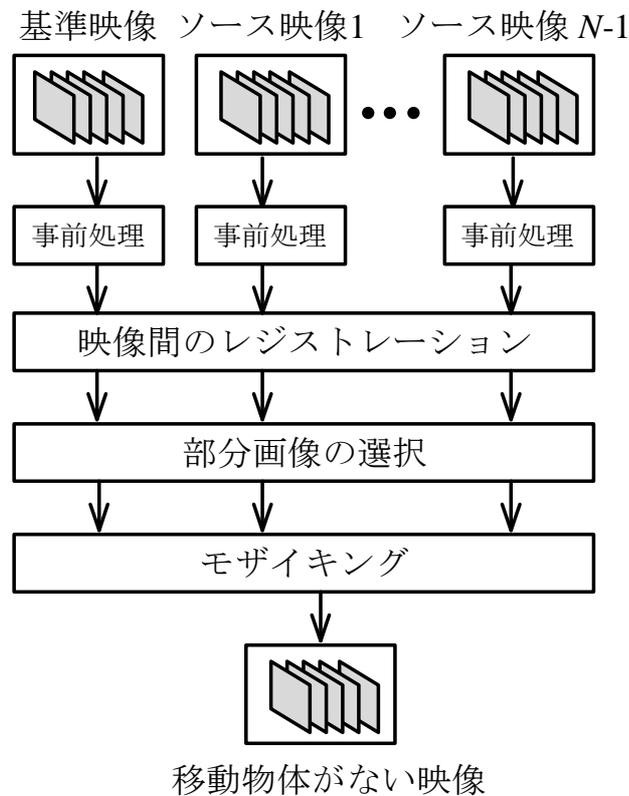


図 3.3 提案手法の処理の流れ .

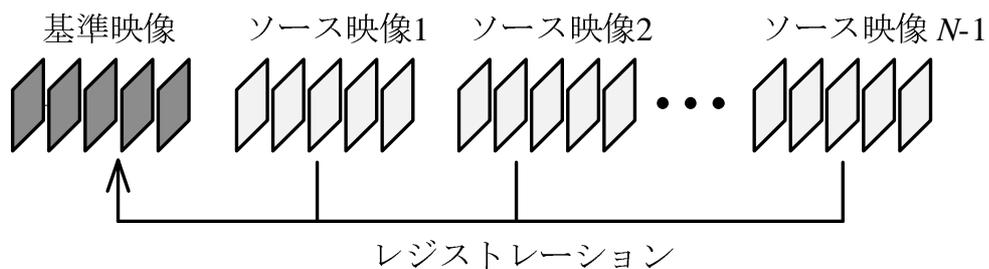


図 3.4 映像間のレジストレーション .

映像間のレジストレーションの手順として、まず始めに、移動物体除去で使用する N 本の映像の中から、任意の 1 本を基準映像として選択する。そして、図 3.4 のように、その他の $N - 1$ 本の映像（ソース映像）を基準映像に一致するようにレジストレーションを行う。このようにすることによって、 $N - 1$ 本のソース映像を基準映像と類似するように変換する。

映像を撮影する車両は道路に沿って走行するが、道路状況は走行毎に異なるため、常に同じ速度で走行することは難しい。このため、映像フレームの撮影間隔に疎密が生じる。これに対処するために、まず始めに、車両の進行方向の位置合わせを行うことを目的とし、映像の時間方向のレジストレーション処理を行う。そしてこの後に、より正確な位置合わせを行うことを目的とし、映像の空間方向のレジストレーション処理を行う。このような2段階処理によって、映像間のレジストレーションを行う。以降、時間方向のレジストレーションと空間方向のレジストレーションについて説明する。

3.2.2.1 時間方向のレジストレーション

まず、全方位カメラ映像間の時間方向のレジストレーションにより、車両の進行方向の位置合わせを行う(図 3.5)。車両の速度は走行毎に異なり、同じ走行中でも一定ではない。そこで、DP マッチングに基づく手法 [24] を適用することでフレーム間の対応付けを行い、このような時間軸の伸縮に対処する。本論文では、映像を撮影する車両は同じ車線内を走行し、車線方向の大きな位置ずれはないと仮定する。そのため、DP マッチングのパターン要素間の距離は、単純に画像の画素値の2乗誤差とする。ここで、パターン要素間の距離を計算する際、建物が映る全方位画像の側面の領域だけを利用し、見えの変化の乏しい車両の正面・背面方向の領域は使用しない。DP マッチングのフレーム間の対応付け結果から、ソース映像が基準映像と同じフレーム数となるように、フレームの挿入・間引きを行う。基準画像の1フレームに対して、ソース映像の複数フレームが対応する場合がある。このような場合には、対応付いた複数フレームの中央のフレームを利用する。また、DP マッチングは、基準映像とソース映像の $N - 1$ 個の対について、それぞれ適用する。

3.2.2.2 空間方向のレジストレーション

映像を収集する車両は、同じ車線を走行したとしても、車線内での走行位置は走行毎に異なる。また、カメラのフレームレートの制限により、DP マッチングにより進行方向の撮影位置を合わせた結果にも誤差が含まれる。そのため、映像の時間方向の位置合わせだけでなく、空間方向の位置合わせも行う必要がある。

市街地環境では建築物が入り組んでいるため、撮影位置が異なると、単純なアフィン変換では画像間の対応付けができない。また、後の部分画像選択処理で、部分画像をきれい

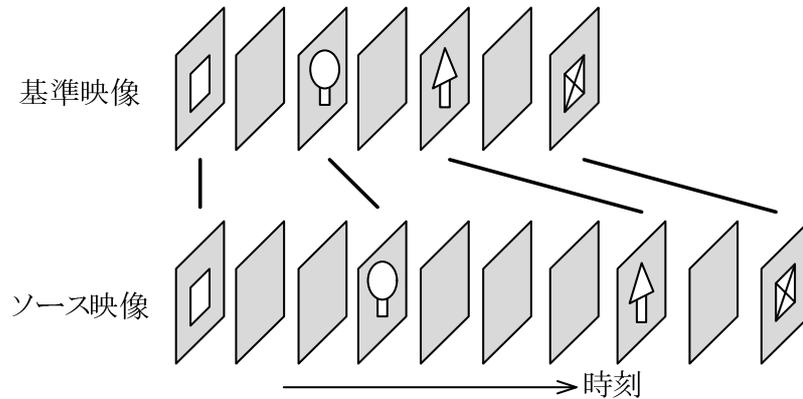


図 3.5 DP マッチングによる映像間の時間方向の位置合わせ .

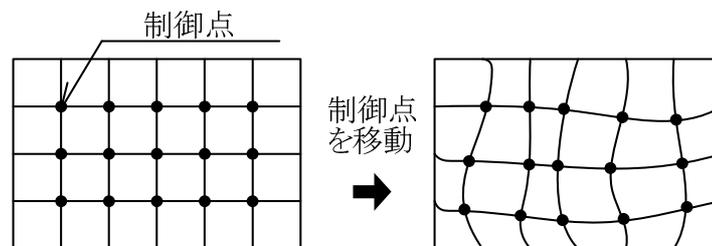


図 3.6 非剛体レジストレーション .

につなぎ合わせるために、ピクセル単位で画像の位置合わせを行う必要がある。そこで、撮影位置の違いによる画像の見えの変化を、画像の非線形な変形で近似することで、映像の空間方向の対応付けを行う。

映像の空間方向の対応付けには、画像の自由な形状変形を行う free-form deformation による非剛体レジストレーション [76] を用いる。画像の見え方の変化は B-spline 曲面で近似する。非剛体レジストレーションでは、まず図 3.6 のように、変形対象となるソース画像上に $n_x \times n_y$ 個の制御点を格子状に設置する。そして、制御点を移動させることでソース画像を変形する。基準画像との照合コストが最小となるように、ソース画像を変形することで、基準画像に一致させる。ここで、ソース画像を A 、基準画像を B とし、ソース画像 A を free-form deformation により変形したものを $\mathcal{F}(A)$ とする。ソース画像 A と基準画像 B 間の照合コスト $C(A, B)$ は、画像間の相違度 $C_d(A, B)$ と B-spline 曲面の

滑らかさに関するペナルティ C_s を用いて次式のように計算する .

$$C(A, B) = -C_d(\mathcal{F}(A), B) + C_s \quad (3.1)$$

ここで , C_s は B-spline 曲面上の各点における 2 次微分の 2 乗和に重みを乗じたものである . C_s の項は , 画像の極端な変形により , 見た目が不自然になることを防ぐ効果がある .

制御点の座標を表すベクトルを $\phi_{i,j}$ ($i = 0, 1, \dots, n_x - 1, j = 0, 1, \dots, n_y - 1$) とする . B-spline 曲面によって , 画素の位置を表すベクトルを $\mathbf{x} = (x, y, z)^T$ から $\mathbf{x}' = (x', y', z')^T$ に変換する関数 $\mathbf{T}(\mathbf{x})$ は次式で表される .

$$\mathbf{T}(\mathbf{x}) = \sum_{l=0}^3 \sum_{m=0}^3 B_l(u) B_m(v) \phi_{i+l, j+m} \quad (3.2)$$

ただし , $i = \lfloor x/n_x \rfloor - 1, j = \lfloor y/n_y \rfloor - 1, u = x/n_x - \lfloor x/n_x \rfloor, v = y/n_y - \lfloor y/n_y \rfloor$ であり , B_l は以下の式のように l 番目の B-spline の基底を表す .

$$B_0(u) = (1 - u)^3/6 \quad (3.3)$$

$$B_1(u) = (3u^3 - 6u^2 + 4)/6 \quad (3.4)$$

$$B_2(u) = (-3u^3 + 3u^2 + 3u + 1)/6 \quad (3.5)$$

$$B_3(u) = (u^3)/6 \quad (3.6)$$

提案手法では , 画像間の相違度 $C_d(A, B)$ として正規化相互情報量を用いる . 市街地のような屋外環境では , 映像を取得する日時や天候により照明条件が異なるため , 照明条件の違いによる明るさの変化が生じる . 具体的には , 画像全体の明るさが異なる他 , 太陽の方向の違いによる不均一な明るさ変化が生じる . そのため , 正確に空間方向の位置合わせを行うためには , 単純な画素間の 2 乗誤差や正規化相互相関のような単純な方法では不十分である . 一方 , 相互情報量を基にした位置合わせ手法は , 画像の明るさが画像全体で不均一に変化する場合でも , より頑健な位置合わせが可能である [77] .

相互情報量は , 2 つの画像間の情報の冗長さを測定するものであり , 値が大きくなるほど , 画像が類似することを示す . 画像間のレジストレーションを目的とし , 相互情報量の正規化形が提案されている . 本研究では , Studholme ら [78] によって提案された画像間の重なり量の変化に頑健な正規化相互情報量を用いる . 画像 A, B 間の正規化相互情報量 $C_{\text{NMI}}(A, B)$ は , 画像 A, B の周辺エントロピー $H(A), H(B)$ と , 画像 A, B 間の結合エントロピー $H(A, B)$ を用いて次式で表される .

$$C_{\text{NMI}}(A, B) = \frac{H(A) + H(B)}{H(A, B)} \quad (3.7)$$

周辺エントロピー $H(A)$ は、画像 A の色ヒストグラムを確率分布 $P(A_i)$ として扱うことによって、次式により計算する。

$$H(A) = - \sum_i P(A_i) \log P(A_i) \quad (3.8)$$

周辺エントロピー $H(B)$ も同様に画像 B の色ヒストグラムから計算する。また、結合エントロピー $H(A, B)$ は画像 A, B の結合ヒストグラムを同時分布 $P(A_i, B_i)$ として扱うことによって、次式により計算する。

$$H(A, B) = - \sum_{i,j} P(A_i, B_i) \log P(A_i, B_i) \quad (3.9)$$

3.2.3 移動物体を含まない部分画像の選択

3.2.3.1 部分画像選択のための目的関数

レジストレーションを行った映像を部分画像単位に分割する。そして、複数の映像の同座標の部分画像群の中から、移動物体が映っていない部分画像を選択する。以降、特定の位置において、部分画像をいずれかの映像から選択する枠のことを、サブウィンドウと呼ぶ。まず、各画像上に $L \times L$ 画素のサブウィンドウを、隣接するもの同士が半分程度重なり合うように設置し、サブウィンドウ内の画素値をベクトルとして表す。このベクトルを以降、部分画像ベクトルと呼ぶ。部分画像ベクトルは、各画素の RGB 値を並べたものであり、 $3L^2$ 次元である。

サブウィンドウは、映像の空間軸および時間軸に沿って並べられており、その位置を (x, y, t) と表記する (図 3.7)。ただし、 $x = 1, 2, \dots, X$, $y = 1, 2, \dots, Y$, $t = 1, 2, \dots, T$ である。また、サブウィンドウの数は $W = XYT$ であり、サブウィンドウのインデックスを w と表記する。そして、以下の処理により、 w 番目のサブウィンドウにおいて、背景らしい部分画像を含む映像のインデックス n_w を決定することで、全体として移動物体の映っていない映像を生成する。

まず、背景らしい部分画像を含む映像を選択する基準として、次の 2 つを用いる。

- (i) 移動物体らしさ
- (ii) サブウィンドウ間の不連続度

具体的には、あるサブウィンドウ w に注目し、インデックス n の映像を選択した場合に

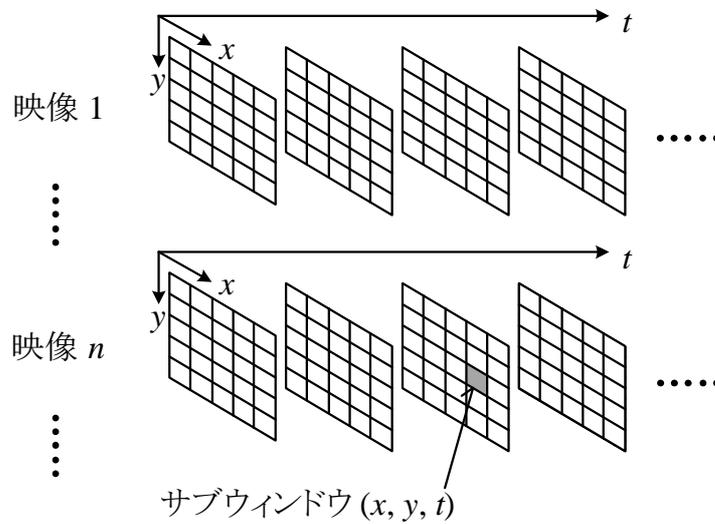


図 3.7 映像とサブウィンドウ .

おける移動物体らしさ $f_w(n)$ と，隣接するサブウィンドウ間の不連続度 $g_w(n)$ を導入する．そして， $f_w(n)$ と $g_w(n)$ の重み付き和を考え，各サブウィンドウにおいて選択される最適な映像のインデックスの集合 $\mathbf{n} = \{n_w\}$ ($w = 1, 2, \dots, W$) を決定するために，次の目的関数を \mathbf{n} に関して最小化する．

$$F(\mathbf{n}) = \sum_{w=1}^W [(1 - \lambda)f_w(n_w) + \lambda g_w(n_w)] \quad (3.10)$$

ここで， λ は重みである．

移動物体らしさ $f_w(n)$ が小さいほど，その部分画像中に移動物体が存在しにくい．また，サブウィンドウ間の不連続度に関する項 $g_w(n)$ を導入することにより，出力映像の隣接するサブウィンドウ間で，入力映像の照明条件の関係を保存する．以降， $f_w(n)$ と $g_w(n)$ について説明する．

3.2.3.2 移動物体らしさ $f_w(n)$

移動物体は同じ地点に留まらないため，異なる時刻かつ同一地点で撮影した画像中の同じ部分を観察した場合，一般的に背景の出現頻度が最も高い．そこで，移動物体らしさ $f_w(n)$ をベクトルメディアンフィルタ [79] を用いて計算する．最頻値ではなくメディアンを用いるのは，サンプル数が少ない場合には最頻値の推定が困難であるためである．ベクトルメディアンフィルタはメディアフィルタの多次元への拡張であり，他のベ

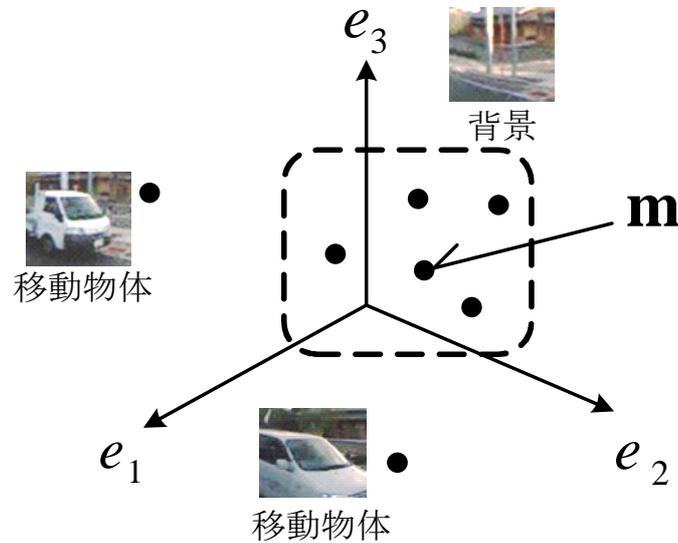


図 3.8 移動物体らしさの計算．メディアンに近い部分画像ベクトルほど，移動物体が存在しにくい．

ベクトルとの距離の和が最小となるベクトルを選択するフィルタである． N 個のベクトル $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N$ を入力とするベクトルメディアンフィルタの出力 \mathbf{m} は，次式で表される．

$$\mathbf{m} = \arg \min_{\mathbf{v} \in \{\mathbf{v}_1, \dots, \mathbf{v}_N\}} \sum_{i=1}^N \|\mathbf{v} - \mathbf{v}_i\|_{L_2} \quad (3.11)$$

ここで， $\|\cdot\|_{L_2}$ はベクトルの L_2 ノルムを表わす．ベクトルメディアンフィルタの詳細な説明を巻末の付録 C に示す．

ベクトルメディアンフィルタは外れ値を除外する性質がある．そのため，図 3.8 のように，中央値に近い部分画像ベクトルほど移動物体が存在しにくいと考えられる．そのため，中央値との距離を，その部分画像ベクトルにおける移動物体らしさ $f_w(n)$ とする． $f_w(n)$ は，サブウィンドウ w でのインデックス n の映像の部分画像ベクトル $\mathbf{v}_w(n)$ と，部分画像ベクトルのメディアン \mathbf{m}_w を用いて，次式で計算する．

$$f_w(n) = \|\mathbf{v}_w(n) - \mathbf{m}_w\|_{L_2} \quad (3.12)$$

3.2.3.3 サブウィンドウ間の不連続度 $g_w(n)$

式 (3.10) の目的関数にサブウィンドウ間の不連続度 $g_w(n)$ を導入することにより，出力映像の隣接するサブウィンドウ間において，入力映像の照明条件の関係を保存する．前述のように，事前処理により画像全体の明るさと色の正規化を行うが，屋外環境では時間帯や天候により，局所的な陰影の有無が異なる．それぞれのサブウィンドウにおいて独立に部分画像選択処理を行った場合，隣接するサブウィンドウ間で異なる照明条件の映像から部分画像が選択されやすくなる．そのため，局所的に周囲と色が異なる部分が発生し，生成される映像の画質が劣化する．そこで，サブウィンドウ間の不連続度に関する項 $g_w(n)$ を加えることにより，隣接するサブウィンドウ間での照明条件の関係を保存しつつ，移動物体を含まない部分画像を選択する．

図 3.9 のように，近傍サブウィンドウでインデックス n_a の映像が選択されている場合，注目サブウィンドウ w に適する映像のインデックス n を決定する場合を考える．具体的には，出力映像の近傍サブウィンドウと注目サブウィンドウの照明条件の関係を，インデックス n_a の映像のものと同じにする．そのためには，注目サブウィンドウ w に適するインデックス n の映像の部分画像として，インデックス n_a の映像中の注目サブウィンドウ w の位置にある部分画像と照明条件が類似する部分画像を選択すればよい．

$A_w = \{n_a\} (a = 1, \dots, M_w)$ をサブウィンドウ w の時間・空間方向を含めた 26 近傍で選択されている映像のインデックスの集合とし，その要素数を M_w とする．また， $\mathbf{v}_w(n)$ を注目サブウィンドウ w で選択されるインデックス n の映像の部分画像とし， $\mathbf{v}_w(n_a)$ をある近傍サブウィンドウで選択されているインデックス n_a の映像の注目サブウィンドウ w と同じ位置の部分画像とする．注目サブウィンドウ w とその近傍のサブウィンドウ間の不連続度 $g_w(n)$ は次式により計算される．

$$g_w(n) = \frac{1}{M_w} \sum_{a=1}^{M_w} \|\mathbf{v}_w(n) - \mathbf{v}_w(n_a)\|_{L_2} \quad (3.13)$$

3.2.4 モザイクキング

部分画像同士が重なり合う領域に，透明度に応じて画像を合成する ブレンディングを適用することで，境界部分が目立ちにくい合成画像を生成する．大量の映像を利用すれば，類似する照明条件の画像が多く存在すると考えられる．そのため，単純な ブレン

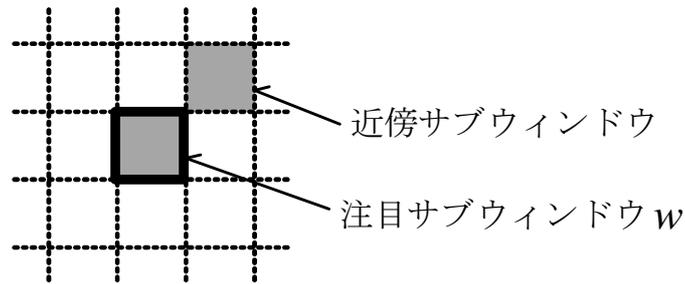


図 3.9 不連続度の計算 .

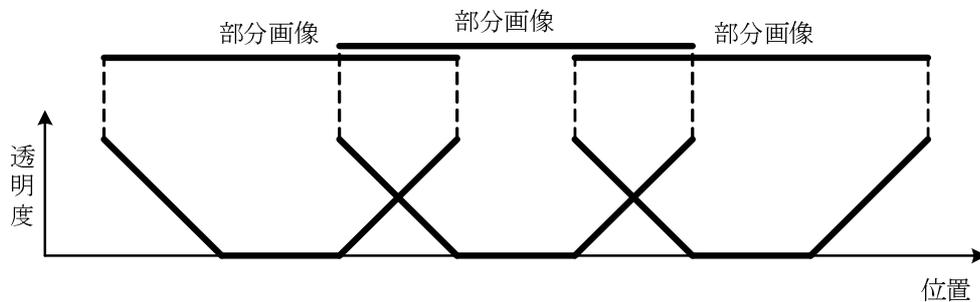


図 3.10 モザイクング . 部分画像の縁から遠いほど透明度が低くなるように , ブレンディングの係数を設定する .

ディングでもモザイクングが十分可能である . ここで , 図 3.10 のように , 部分画像の縁の近くは透明度が高く , 縁から遠いほど透明度が低くなるように , ブレンディングの係数を設定する . ブレンディングの係数に応じた透明度で , 部分画像同士を貼り合わせるにより , 1 つの映像を生成する .

3.3 実験

3.3.1 実験条件

提案手法の有効性を確かめるために , 実際に市街地で撮影した映像を用いて次の 2 つの評価を行った .

- 移動物体除去性能の評価
- レジストレーションの精度評価



図 3.11 実験に用いたデータ．様々な照明条件の映像を含む．

全方位カメラとして Point Grey Research 社製 Ladybug2 を使用し，フレームレート 15 fps で撮影した．実験のために合計 15 本の映像を収集した．その中の無作為に選んだ 1 本を基準映像として用い，残りの 14 本をソース映像として用いた．映像の長さは約 2 分で，走行距離は約 1.0 km である．映像は全て異なる日時に撮影されており，様々な照明条件の映像が含まれる．図 3.11 に実験に用いた画像の例を示す．実験には，球面投影によりパノラマ展開した映像を用いた．解像度は視野角 360 deg あたり 1,024 画素であった．

移動物体除去性能の評価では，従来の画素成分ごとにメディアンを用いた背景選択手法と提案手法中の背景画像選択手法を比較した．比較手法として用いるスカラ値に対するメディアンは，値を昇順にソートした中央値とし，画素のそれぞれの RGB 成分に対して独立に処理を行った．また，ソース映像として使用する映像の数を変化させ，使用する映像数と移動物体の除去性能の関係を調査した．実験手順として，まず N 本の入力映像に対し，提案手法によるレジストレーション処理を行った．そして，提案手法と比較手法により，それぞれ背景画像選択処理を行った．

移動物体除去性能の評価方法として，基準映像中で移動物体が存在する領域の画素のうち，除去映像中に残った画素の数を比較した．具体的には，正解として，まず人手で移動物体の画素をマーキングした基準映像・ソース映像を用意した．そして，それぞれの画素の位置において，提案手法で選択される映像と同じマーキングした映像から画素を取得することにより，除去映像中に残った画素を計数した．本実験では，ランダムに生成した 50 のソース映像の組み合わせについて実験を行い，その平均を評価に用いた．ただし，ソース映像を 14 本用いた場合では 1 つの組み合わせのみを試行した．評価には，映像中から無作為に選んだ車両が映っている 11 地点を用いた．この実験における部分画像選択処理

の際のサブウィンドウのサイズは 30×30 画素とした。また、提案手法における式 (3.10) の重み λ は予備実験により 0.6 とした。

映像間のレジストレーション精度の評価では、DP マッチングのみを用いた場合と、DP マッチングに加え、非剛体レジストレーションも併用した場合について、画像の位置合わせ誤差を比較した。手順として、まず基準画像とソース画像中の建物の角や窓枠等の特徴的な点を人手で対応付けた。そして、基準画像とソース画像の対応する特徴点間の画像上の距離を計算し、それをレジストレーションの誤差とした。1枚の画像につき約 10 箇所の特徴点を選択し、15 地点の画像対について計測を行い、その誤差の平均を評価した。

式 (3.10) の最小化問題について、逐次的なアルゴリズムにより近似解を求める。具体的には、まずランダムに 1 つのサブウィンドウ w を選択し、これに対応する n_w について、目的関数の値が減少するような n_w の値を総当たりで探索する。そして、見つかった n_w を、新たな n_w の値 n_w^{new} とする。この操作は、次式で表わされる。

$$n_w^{\text{new}} \leftarrow \arg \min_{n_w} [(1 - \lambda)f_w(n_w) + \lambda g_w(n_w)] \quad (3.14)$$

この操作を収束するまで繰り返す。いずれのサブウィンドウ w において、 n_w が変化しなくなった時点で、逐次処理を終了する。

3.3.2 実験結果

提案手法と比較手法について、用いた映像数（基準映像数 + ソース映像数）に対する除去成功率を図 3.12 に示す。提案手法では、映像数が 15 のとき 96.5% の移動物体領域の除去に成功した。一方で、比較手法では 78.5% であった。提案手法では、残った移動物体領域の多くが遠くの小さなものや、移動物体の断片であり、得られた移動物体除去映像は目的とするドライバ支援のためのデータベースに十分利用可能であると考えられる。

移動物体除去の例を図 3.13 に、別の地点の移動物体除去の例を 3.14 に示す。(a) が入力となる基準映像で、(b) ~ (d) が提案手法による 15, 9, 7 本の映像の統合結果である。(a) で存在していた路上の移動物体が、(b) ~ (d) で大部分が除去されていることが確認できる。

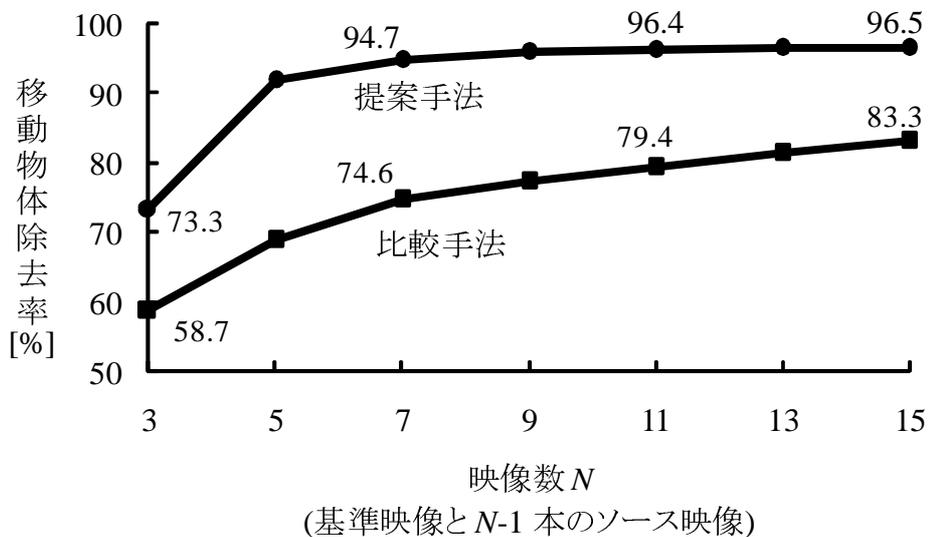


図 3.12 使用した映像数 N と移動物体除去成功率の関係。

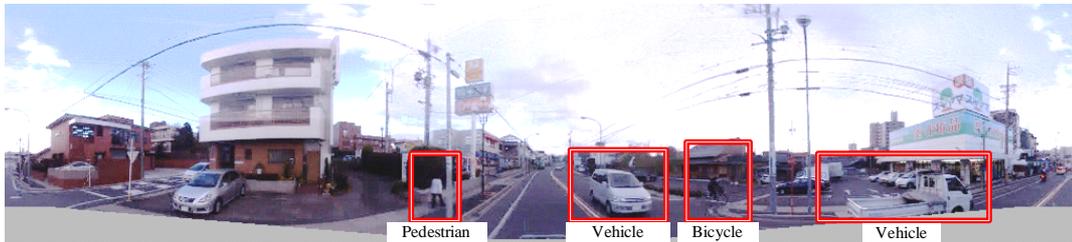
レジストレーションの誤差の累積分布を図 3.15 に示す。誤差が 4 画素以下の割合が DP マッチングのみの場合は 52% であったのに対し、非剛体レジストレーションも適用した場合は 89% であった。DP マッチングに加え、非剛体レジストレーションも用いることで、レジストレーション精度の向上が見られた。

映像間のレジストレーションの適用例を図 3.16 に示す。左列が DP マッチングのみを適用した場合の結果であり、右列が非剛体レジストレーションも用いた場合の結果である。基準画像とソース画像をチェッカボード状に並べて表示している。例えば、2 段目の左列の画像では、建築物の屋根と空の境界がずれているが、非剛体レジストレーションにより、右列の画像のように補正することが可能となる。

3.4 考察

本節では、以下について考察する。

- 3.3 の実験結果に対する考察
- 本手法の対象とする範囲
- 最適なサブウィンドウの大きさについての考察
- 移動物体除去に用いる映像の違いの影響



(a) 基準映像 (入力)



(b) 15本の映像の統合結果 (提案手法)

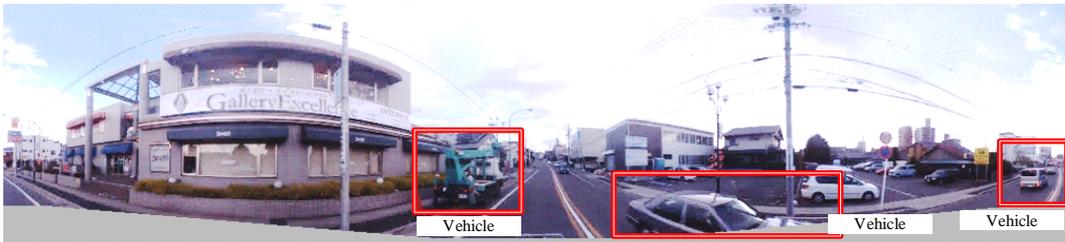


(c) 9本の映像の統合結果 (提案手法)



(d) 7本の映像の統合結果 (提案手法)

図 3.13 移動物体除去の結果の例 1.



(a) 基準映像 (入力)



(b) 15本の映像の統合結果 (提案手法)



(c) 9本の映像の統合結果 (提案手法)



(d) 7本の映像の統合結果 (提案手法)

図 3.14 移動物体除去の結果の例 2 .

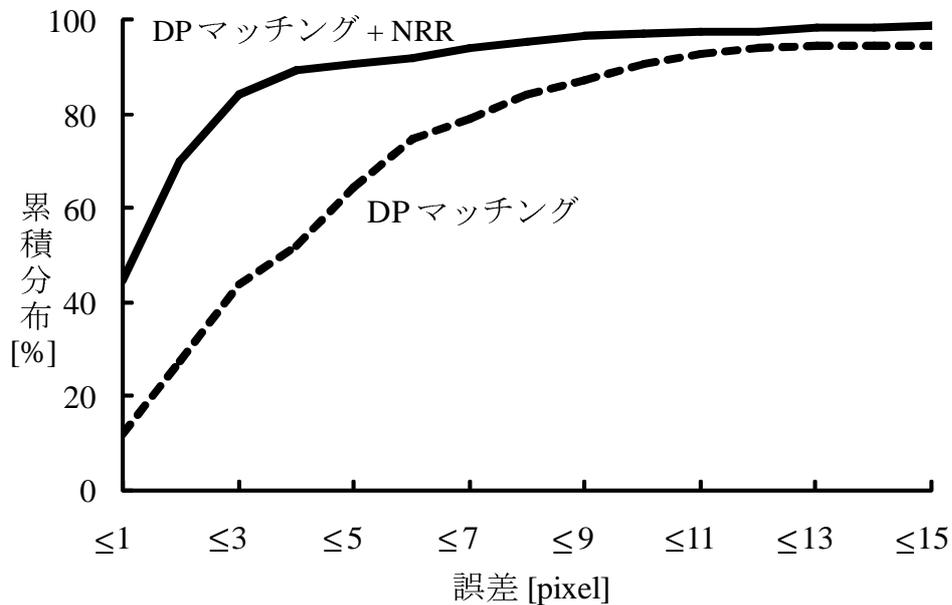


図 3.15 レジストレーション誤差の累積分布．DP マッチングに加え，非剛体レジストレーション（NRR）も用いることで，レジストレーション精度の大きな向上が見られた．

- 複数映像の収集方法
- メディアンの選択手法の考察

3.4.1 実験結果に対する考察

図 3.12 の使用した映像数と移動物体除去率の関係から，映像数が多いほど移動物体除去性能が向上することが確認できる．提案手法では，最低 3 本の映像があればベクトルメディアンフィルタを基にした部分画像選択は可能であるが，9 割以上の移動物体除去の性能を得るためには，5 本以上の映像が必要であった．

提案手法と比較手法による生成画像の一部分を拡大した画像と，除去できなかった移動物体の画素を黒色で表わした画像の例を図 3.17 に示す．提案手法は良好に移動物体が除去できることが分かる．一方，比較手法では，複数の入力映像に存在する移動物体の画素が粒子状に多く残留した．これにより，画素成分単位ではなく領域単位で処理することにより，除去性能が向上することが確認できる．

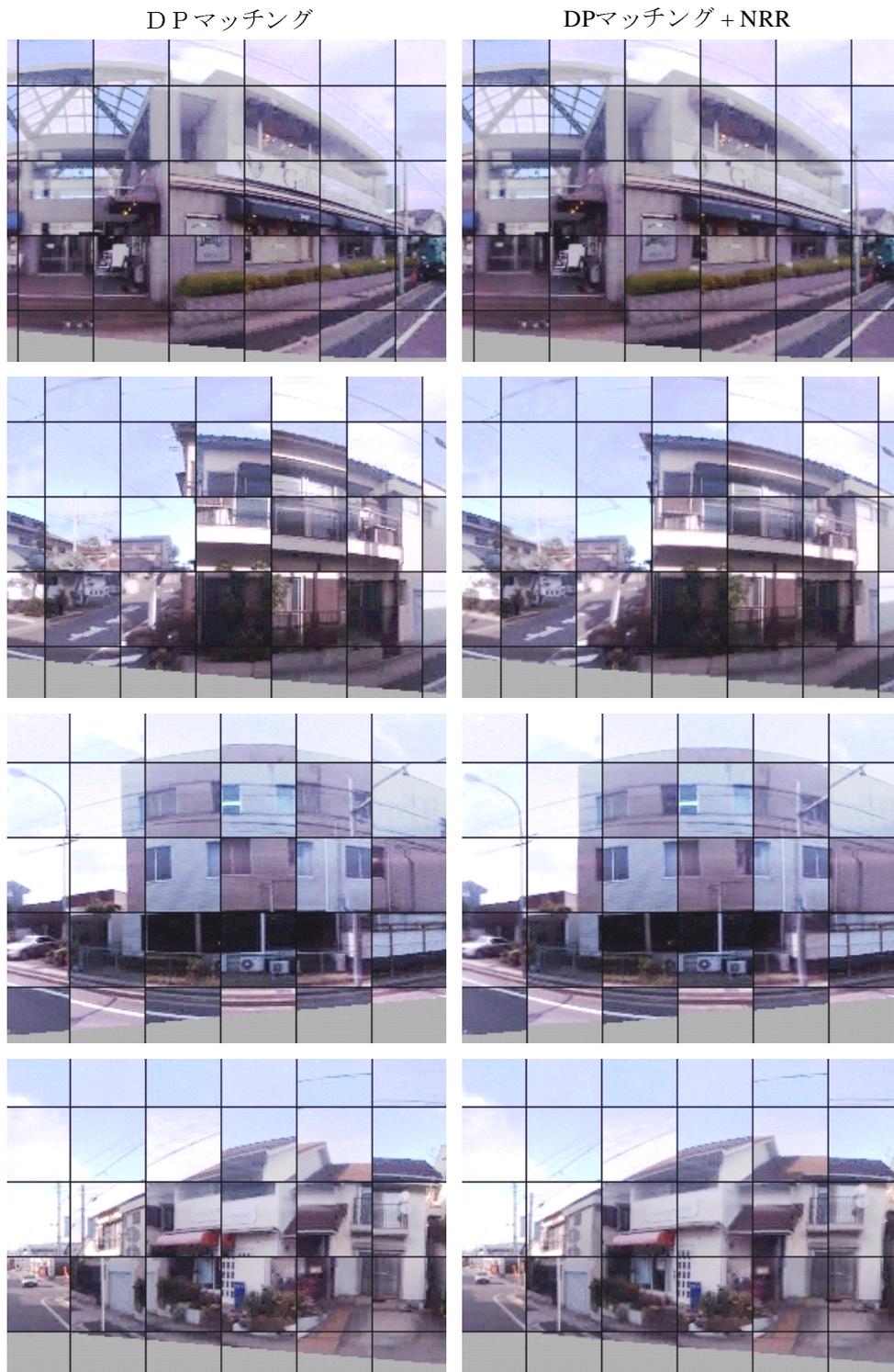
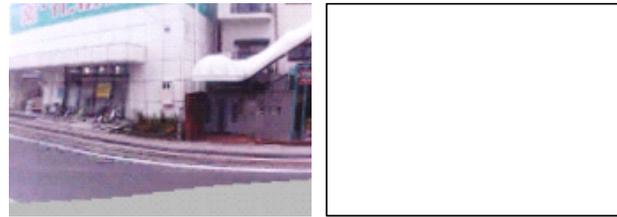
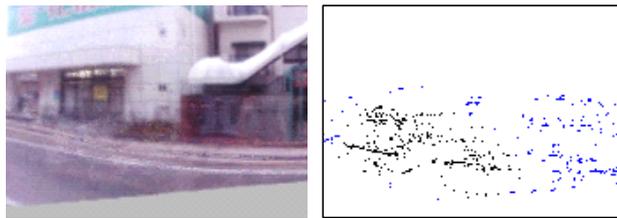


図 3.16 映像間のレジストレーションの適用例．左列に DP マッチングの適用結果を示す．左列の画像に非剛体レジストレーション（NRR）を適用した結果を右列に示す．ソース画像と基準画像をチェッカボード状に並べて表示してある．



(a) 提案手法 (領域単位)



(b) 比較手法 (画素成分単位)

図 3.17 生成された画像と残留した移動物体の画素．画素成分単位ではなく，領域単位で処理することにより，除去性能が向上する．

(a) $\lambda = 0.0$ (b) $\lambda = 0.6$

(c) 元画像

図 3.18 部分画像選択の重み λ による生成画像の違い．

式 (3.10) の部分画像選択の重み λ が 0.0 と 0.6 の場合において生成された画像の一部を拡大したものを，それぞれ図 3.18(a)，(b) に示す．使用した映像数は 9 本である．図 3.18 の建築物の白い壁に注目すると，本来は一樣な色の壁であるにもかかわらず，(a) では局所的に周囲と色が異なる部分が見られる．これは，式 (3.10) のサブウィンドウ間の不連続度 $g_w(n)$ の比重が小さい場合には，照明条件の異なる映像から部分画像が選択されやすくなるためである．一方，適切な重み λ を設定した場合には，(b) のように，改善されている．このように，サブウィンドウ間の連続性を考慮することにより，生成映像の画質が向上することが確認できる．

重み λ と移動物体除去率の関係を図 3.19 に示す．重み λ が大きく，不連続度 $g_w(n)$ の

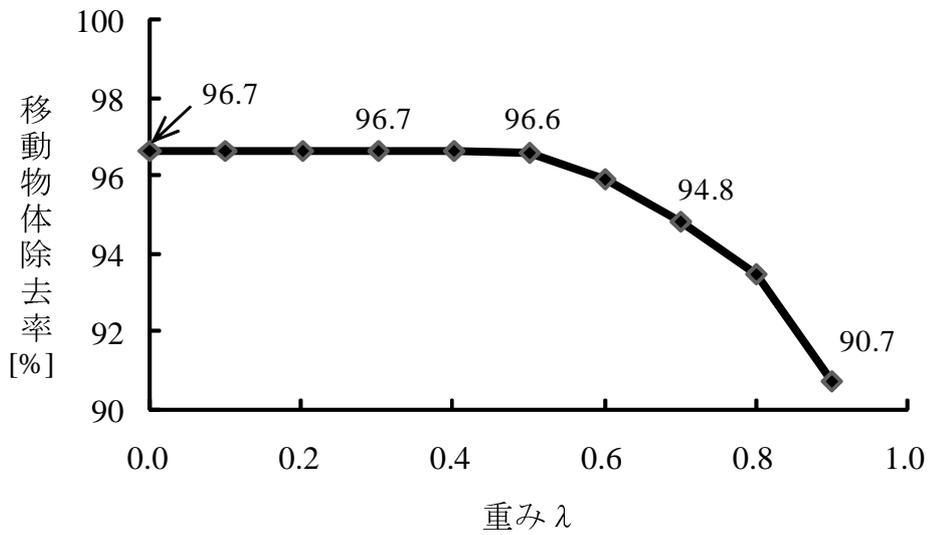


図 3.19 部分画像選択の重み λ と移動物体除去率の関係

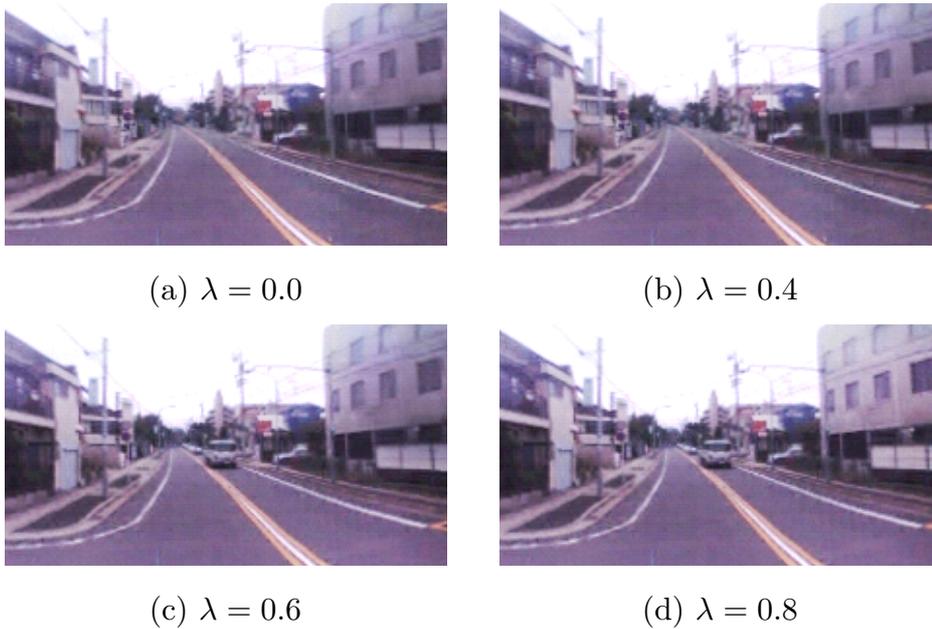


図 3.20 部分画像選択の重み λ の導入による移動物体除去性能の失敗例．(c) , (d) のように不連続度 $g_w(n)$ の重みが大い場合，移動物体除去率の低下が見られた．

比重が大い場合では，移動物体除去率の低下が見られた．図 3.20 に λ を導入したことによる移動物体除去の失敗例を示す．これは，不連続度 $g_w(n)$ を導入した場合には，サブウィンドウ中に移動物体が存在する場合でも，元の入力映像の連続性を保存しようとする

ため、移動物体が存在する部分画像も選択されやすくなるためである。しかしながら、サブウィンドウ間の不連続度 $g_w(n)$ を導入した場合でも、移動物体除去率の低下が十分に小さいことを確認した。

実験に使用した映像の長さはそれぞれ異なるが、マッチングにおいて映像の長さの影響は見られなかった。また、提案手法では、DP マッチングのパターン要素間の距離として、単純に画像の画素値の2乗誤差を用いたため、車線内の走行位置が異なることにより、映像間の対応付け精度が低下することが懸念される。しかし、実験に用いた映像ではそのような問題は見られなかった。異なる車線を走行する状況も想定する場合には、車線方向の位置の違いによる対応付け精度の低下が懸念されることから、DP マッチングのパターン要素間の距離について、さらなる検討が必要である。

CPU Core i7 870 2.93GHz、メモリ 8GB の計算環境で、1対の映像の DP マッチングに要する計算時間は平均 6.2 分、1対の画像の非剛体レジストレーションに要する計算時間は平均 4.7 分、使用する映像が 9 本の場合の部分画像選択処理に要する計算時間は約 6.0 分であった。このように、計算時間の大部分が非剛体レジストレーションであった。しかしながら、照合コストの最小化を改良した手法 [81] や GPU を利用した手法 [80] 等、より高速な非剛体レジストレーション手法が提案されている。そこで、照合コストの最小化に Markov Random Field を利用した非剛体レジストレーション手法 [81] を適用したところ、1対の画像の非剛体レジストレーションに要する計算時間は約 0.5 分であった。そのため今後、非剛体レジストレーション高速化のより詳しい検討が必要である。

3.4.2 本手法の対象とする範囲

3.4.2.1 移動物体除去性能について

ドライバ支援のデータベースといった用途では、時間的に変化しない物体のみから構成された市街地画像が必要となる。また、プライバシー保護の目的では、人物の顔やナンバープレートが映っていない映像が必要である。3.3.2 の実験結果では 96.5% の移動物体を除去できており、目的とする移動物体除去がほぼ達成できた。しかしながら、除去に失敗した例が一部で見られたため、その考察を以下で行う。

提案手法で移動物体除去に失敗した例を図 3.21 に示す。図 3.21(a) は常に駐車場の同じ位置に同じ車両が存在したため、除去に失敗した例である。ドライバ支援のデータベースといった用途では、時間的に変化しない物体のみから構成された市街地画像が必要とな



図 3.21 移動物体除去の失敗例．

るため，このような駐車場の車両は問題とならない．しかしながら，プライバシー保護の観点では問題となるため，その場合には，他の手法を併用する必要がある．

図 3.21(b) は，交通量が多い地点であり，大部分の入力映像中に移動物体が映っていたため，移動物体除去に失敗した例である．本実験の映像には見られなかったが，交通量によっては，すべての入力映像中で前後方向に他車両が映る場合も考えられる．このような問題を防ぐには，交通量が少ない時間帯の映像を利用するなどの対策が必要となる．

また，遠景の移動物体では，サブウィンドウの大きさに比べて移動物体が小さく，除去されない例が見られた．しかしながら，移動物体が小さい場合には，ドライバ支援のデータベースやプライバシー保護といった応用では問題にならないと考える．

3.4.2.2 非剛体レジストレーションについて

提案手法では，非剛体レジストレーションにより，複数の車載カメラ映像から，同一地点で撮影したものと見目が類似する画像群を生成した．従来から，異なる視点で撮影した 3 次元物体の画像間を非剛体変形により対応付けることは広く行われている [82]．以降で，非剛体レジストレーションを市街地画像に適用することの妥当性を考察する．

市街地の建造物の多くは，壁面と道路面で構成されており，特に壁面は道路方向に平行なものと垂直なものが多い．図 3.22(a) のように 2 視点間で観測可能な壁面領域が共通している場合には，非剛体変形で対応可能である．一方，図 3.22(b) のように手前の物体が奥の物体を遮蔽している場合は対応付けが難しい．しかしながら，車線方向への視点位置の変化が十分小さい場合や，車両の速度に対して映像のフレームレートが十分高い場合

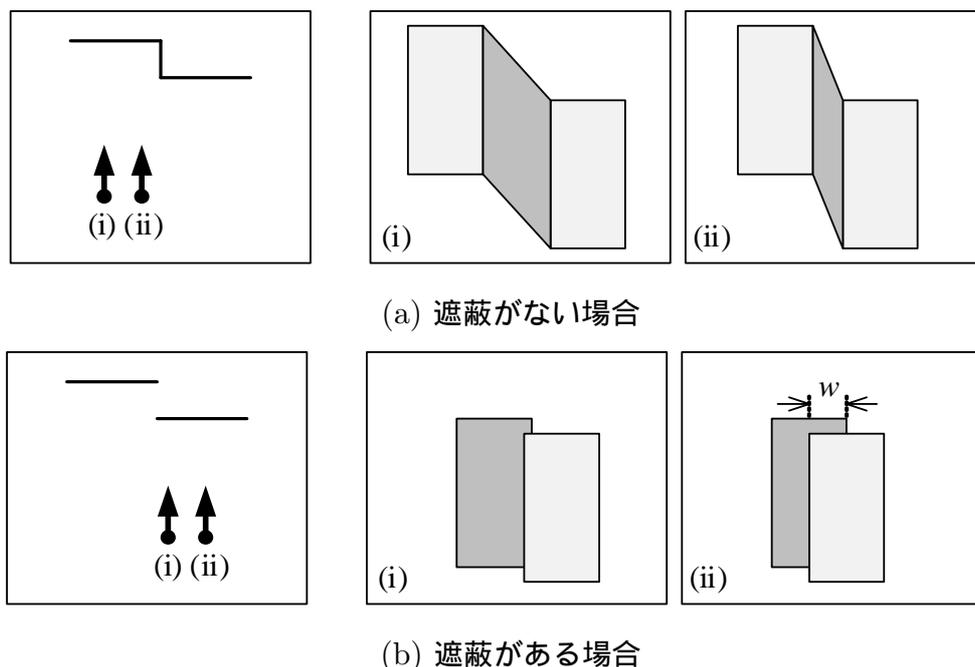


図 3.22 非剛体レジストレーションを市街地映像に適用することの妥当性。(a) の場合には、非剛体変形で対応可能である。(b) の場合には、視点の位置の違いが小さければ、遮蔽は無視できる。

は、遮蔽部分の影響は無視できると考えられる。車線方向の位置の違いがない場合では、車速が 40 km/h，フレームレートが 15 fps，カメラと手前の物体との距離が 3 m，奥の物体との距離が 5 m と仮定すると，進行方向の視点間の距離は最大 0.37 m となり，遮蔽領域の幅 w は最大 2.8 度の視野角に相当すると見積もることができる。この場合，遮蔽領域の幅は十分に小さく，非剛体変形でも十分対応できる。特に，奥行きの違いが小さい場合や，奥行きの違いが大きくても，これらの物体が遠方にある場合には遮蔽領域の幅が小さくなる。一方，奥行きの違いが極端に大きい場合には，遮蔽領域の幅が増加するため，非剛体変形では対応が難しくなる。本実験で，電柱と背後の建物でこのような事例が見られることがあったため，今後それに対応した手法の開発が必要である。

カメラの設置位置の違いや車線方向の位置の違いについても，進行方向の視点位置の違いと同様である。そのため，カメラの設置位置がほぼ同じ，かつ同じ車線内を走行する限り，大部分の典型的な市街地の構造において，非剛体変形で十分対応可能であると考えられる。

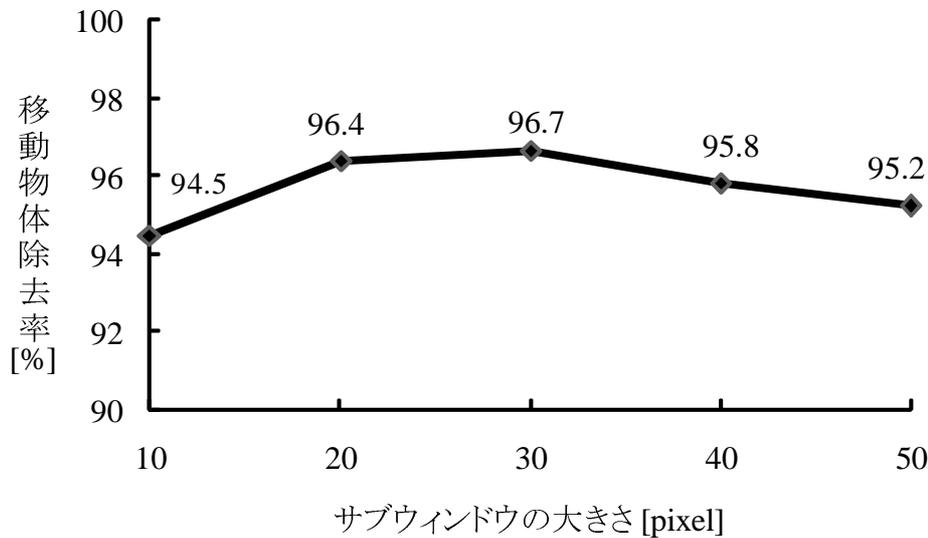


図 3.23 サブウィンドウの大きさと移動物体除去成功率の関係。

3.4.3 最適なサブウィンドウの大きさについての考察

サブウィンドウの大きさによって、移動物体除去の結果が変化するものと考えられる。そこで、サブウィンドウの大きさの違いによる移動物体除去率について調査した。映像は実験 1 と同じものを用い、重みは $\lambda = 0$ とした。サブウィンドウの大きさと移動物体除去率の関係を図 3.23 に示す。図 3.23 のように、サブウィンドウの大きさによって移動物体除去率が異なるという結果が得られた。また、サブウィンドウの大きさより移動物体が小さく映る場合、除去されない例が見られた。このため、最適なサブウィンドウの大きさの決め方に関する指針として、移動物体除去率を考慮する他、除去対象とする移動物体の大きさにあわせる方法が考えられる。例えば、移動物体除去後の映像を自車位置推定のためのデータベースとして用いる場合、小さく映る移動物体は問題とならないため、移動物体除去率が最大となるサブウィンドウの大きさを採用すればよい。一方、遠くの小さい物体も除去したいような映像の利用方法の場合、全体の移動物体除去率を考慮しつつ、移動物体除去率が最大となるサブウィンドウの大きさより、やや小さいものを採用すればよい。

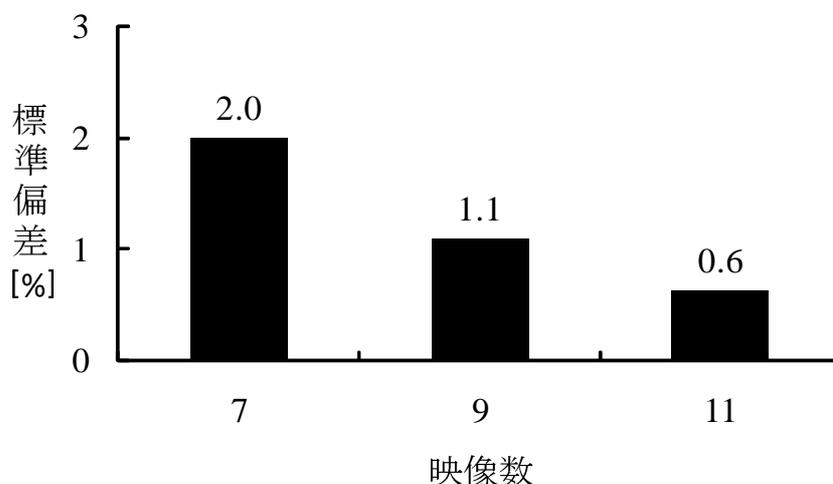


図 3.24 移動物体除去率の標準偏差 .

3.4.4 移動物体除去に用いる映像の違いの影響

実験に使用した映像は、それぞれ撮影日時や撮影環境が異なる。映像の照明条件や撮影位置の違いによる要因が、移動物体除去率に影響を与えるものと考えられる。そこで、利用するソース映像や基準映像が異なることによる移動物体除去性能の影響について調査した。

利用するソース映像の違いの影響を調べるために、ソース映像を様々に変化させ、移動物体除去率の標準偏差を調べた。図 3.24 に映像数が 7, 9, 11 のときの標準偏差を示す。図 3.24 のように、映像数が 7 本で 2.0% (移動物体除去率平均 94.7%)、映像数が 9 本で 1.1% (同 95.9%)、映像数 11 本で 0.6% (同 96.4%) であった。利用する映像数が多いほどソース映像の違いによる移動物体除去率のばらつきが小さくなることが確認できる。この要因として、映像が多い場合、集めた映像の大部分に移動物体が映り込んでしまう状況の発生頻度が少なくなることが考えられる。また、ベクトルメディアフィルタの特性として、映像数が多いほどメディアンを安定に推定できることも要因として考えられる。

また、選択する基準映像の違いによる影響を調べるために、基準映像として昼(午前 11 時)に撮影した映像を選択した場合と、夕方(午後 4 時)に撮影した映像を選択した場合で実験を行った。ただし、それぞれの実験で使用した映像群は同じであり、ソース映像に

は昼と夕方に撮影した映像が含まれている．使用した映像数が9本のとき，基準映像が夕方の場合には移動物体除去率は95.9%であり，昼の場合には93.5%であった．このように，選択する基準映像により移動物体除去率が異なった．そのため，移動物体除去性能の違いは，映像間のレジストレーションの精度によるものと考えられる．レジストレーション精度の違いの要因として，撮影位置の違いの影響の他，映像の色合いや明るさの影響が挙げられる．そのため，基準映像の選び方の指針として，レジストレーションの精度が向上するように選択することが考えられる．例えば，映像群中で色合いや明るさが中間的な映像を選択する方法が挙げられる．

3.4.5 複数映像の収集方法

使用する映像数が多いほど移動物体除去性能が高くなるという結果が得られた．しかしながら，映像数が多いほど，その収集コストは大きくなる．この問題の解決策として，公共の交通機関，運送業者等の車両にカメラを装着することで，低コストに映像を収集する方法が考えられる．収集する映像としては，3.4.2の議論から，交通量が少ない時間帯の映像や，車線等の走行条件が類似する映像が適すると考えられる．

一方，使用する映像数を削減するという方向性も挙げられる．例えば，移動物体らしさの計算の際に，映像の連続性を活かし，注目サブウィンドウに対する時間方向・空間方向の周辺画像領域を使用する方法が考えられる．

3.4.6 メディアンの選択手法の考察

3.4.6.1 概要

メディアンのベクトルを選択する手法として，ベクトルメディアンフィルタを改良した手法が多く提案されている．これらの多くは，マルチチャネルの信号やカラー画像中の雑音除去に用いられており，ベクトルメディアンフィルタより雑音除去性能が良いことが報告されている [83]．そこで，これらのメディアンのベクトルを選択する手法を提案手法に適用し，移動物体除去に利用可能か調査を行った．具体的には，ベクトルメディアンフィルタ (VMF) を用いた手法と，Basic Vector Directional Filter (BVDF) [84]，Directional Distance Filter (DDF) [85]，Directional-Magnitude Vector Filter (DMVF) [86] を用いた手法の比較を行った．これらのメディアンのベクトルを選択手法の特性を以下で説明する．

(a) Basic Vector Directional Filter (BVDF)

Basic Vector Directional Filter は他のベクトルとの角度の和を最小とするベクトルを出力する。角度を用いるため、ベクトルメディアフィルタに比べ、明るさの変化の影響が小さい。

$$\mathbf{m} = \arg \min_{\mathbf{v} \in \{\mathbf{v}_1, \dots, \mathbf{v}_N\}} \sum_{i=1}^N \cos^{-1} \left(\frac{\mathbf{v} \cdot \mathbf{v}_i}{|\mathbf{v}| \cdot |\mathbf{v}_i|} \right) \quad (3.15)$$

(b) Directional Distance Filter (DDF)

他のベクトルとの距離の和と、他のベクトルとの角度の和の重み付き積を最小とするベクトルを出力する。ベクトルメディアフィルタと Basic Vector Directional Filter のそれぞれの項の重み付き積となっている。

$$\mathbf{m} = \arg \min_{\mathbf{v} \in \{\mathbf{v}_1, \dots, \mathbf{v}_N\}} \left[\sum_{i=1}^N |\mathbf{v} - \mathbf{v}_i| \right]^\gamma \left[\sum_{i=1}^N \cos^{-1} \left(\frac{\mathbf{v} \cdot \mathbf{v}_i}{|\mathbf{v}| \cdot |\mathbf{v}_i|} \right) \right]^{(1-\gamma)} \quad (3.16)$$

ここで、 γ は重みである。

(c) Directional-Magnitude Vector Filter (DMVF)

他のベクトルとの距離を、そのベクトルとの和の長さで割ったものの総和を最小とするベクトルを出力する。ベクトルメディアフィルタに正規化項を加えることで、明るさの変化の影響を抑える。

$$\mathbf{m} = \arg \min_{\mathbf{v} \in \{\mathbf{v}_1, \dots, \mathbf{v}_N\}} \sum_{i=1}^N \frac{|\mathbf{v} - \mathbf{v}_i|}{|\mathbf{v} + \mathbf{v}_i|} \quad (3.17)$$

3.4.6.2 実験と考察

15 本から 7 本の映像を選択し、移動物体除去手法を適用した。利用する映像を変えながら複数回実験を行い、それらの除去率の平均を計算した。窓サイズは 30×30 pixel とし、重み λ は 0.0 とした。

移動物体除去率を図 3.25 に示す。図 3.25 から、ベクトルメディアフィルタ (VMF) が最も良い結果が得られた。ベクトルメディアフィルタ以外のフィルタでも移動物体除去は可能であるものの、ベクトルメディアフィルタを超える性能は得られなかった。ベクトルメディアフィルタ以外の手法は彩度 (ベクトルの長さ) を正規化する。このこと

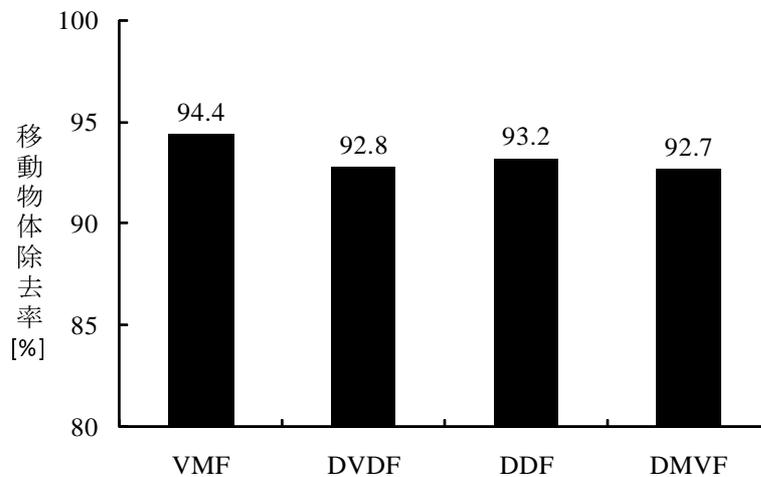


図 3.25 メディアンの選択手法の比較。

から、彩度を正規化するメディアンの選択手法は効果的ではないと考えられる。その原因として、道路面が無彩色であり、車両も無彩色のものが多いため、彩度を正規化することにより両者の区別が難しくなることが挙げられる。現状の照明条件の正規化を行わないベクトルメディアンフィルタでも大きな移動物体除去性能低下は見られないものの、屋外環境では様々な照明条件が想定されるため、性能向上を目指すためには、照明条件に影響されない背景画像の選択手法の開発が必要であると考えられる。

3.5 まとめ

運転者支援等の市街地を対象とした研究や、市街地画像閲覧サービスのための市街地映像データベース構築を目的とし、移動物体を含まない映像の生成手法を提案した。同一経路を複数回走行して得た映像を、時間的・空間的なレジストレーション手法と移動物体らしさに基づく部分画像選択手法により統合した。15本の映像を用いた実験の結果、96.5%の移動物体の除去に成功し、目的とする市街地映像データベース構築に十分な精度を得た。今後、同一経路の大量の映像を収集するために、カメラを搭載した複数の車両が協調するという枠組みが考えられる。その他の課題として、移動物体除去性能と画質の両立が可能な手法の開発や、使用する映像数の削減が挙げられる。

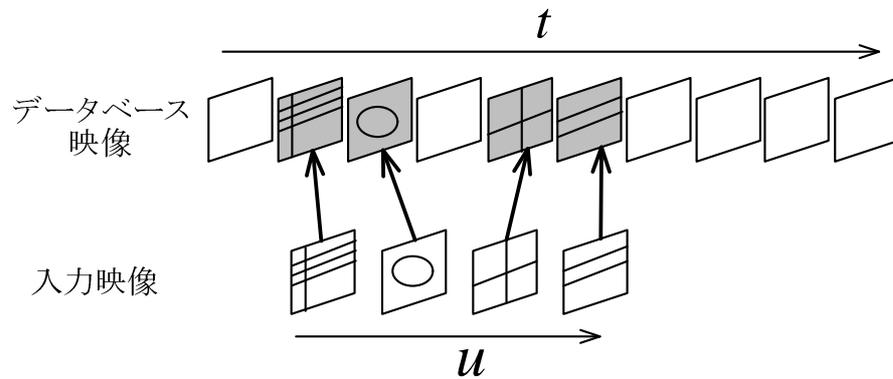
第 4 章

全方位市街地映像データベースと車載通常カメラ映像の照合による自転車位置推定

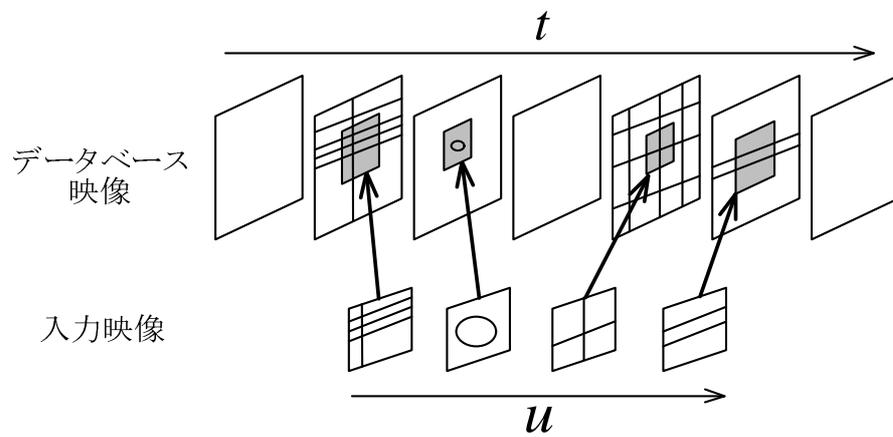
4.1 はじめに

第 2 章で述べたように，従来のアピランススペースの自転車位置推定の手法では，データベース構築時と位置推定時に同一種類のカメラを用いる [4][67][68]．しかしながら，市街地映像データベース構築時は広い範囲を一度に撮影できる全方位カメラが適するが，位置推定時は構造上の問題から通常カメラが適する．そこで本論文では，データベース構築時に全方位カメラ，位置推定時に通常カメラを用いる手法を提案する．ここで提案するカメラ構成では，全方位カメラと通常カメラの視野角の違いにより，画像の見え方が大きく異なる．そのため，画像間の正確な照合が難しい．そこで本研究では，画像の見えの違いに対応すると同時に，映像の時間方向に対応付けることで，全方位カメラ映像と通常カメラ映像のフレーム間に対応付ける．また，市街地には類似する物体が多く存在するため，単一フレームの画像照合では正確な対応付けが難しい．そこで，複数フレームを照合に用い，時系列情報を利用することで，位置推定精度を向上させる．

以降，4.2 で従来手法を異種カメラ間対応付けに適用する場合の問題点と提案手法の基本戦略を述べる．4.3 で提案手法の詳細を説明する．次に，4.4 で評価実験について述べ，4.5 で考察を行う．4.6 でカメラ方向が未知の場合の異種カメラ間対応付け手法とその評価実験について述べる．最後に，4.7 で第 3 章の移動物体除去手法を適用した市街地映像



(a) 従来手法



(b) 提案手法

図 4.1 DP マッチングによる自転車位置推定．従来手法では時間軸方向のみ探索を行う．提案手法では，時間軸方向に加え，空間方向も探索を行う．

データベースを自転車位置推定に利用した実験について述べる．

4.2 提案手法の基本戦略

4.2.1 従来手法を異種カメラ間対応付けに適用する場合の問題点

従来，時系列情報を利用する同一カメラを用いた自転車位置推定手法として，DP マッチングを用いた手法 [24][61][62] が提案されている．DP マッチングは，ある 2 つの時系列パターンに対し，パターンの順序関係を保存しつつ，時間軸の非線形な伸縮を許容する対応付けを求める手法である．これにより，車両の速度の違いに頑健な入力映像とデータベー

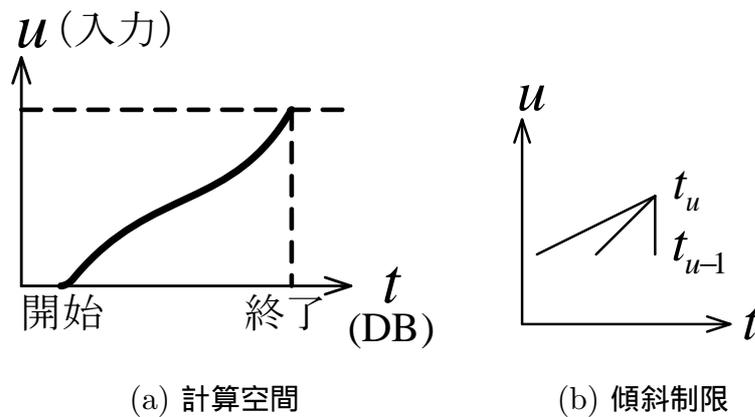


図 4.2 従来の DP マッチングの (a) 映像間の照合コストを最小とする対応付けの軌跡と (b) 傾斜制限の例。

ス映像間の対応付けが実現できる(図 4.1(a)). 佐藤ら [24] は, 全方位カメラ映像同士の対応付けに DP マッチングを用いており, その原理は以下の通りである. まず, データベース画像の系列 $\{A[1], \dots, A[t], \dots, A[T]\}$ と入力画像の系列 $\{B[1], \dots, B[u], \dots, B[U]\}$ が与えられるとする. データベース画像の系列の各要素には位置情報が付加されており, 入力画像の系列の時刻 t とデータベース画像の系列の時刻 u の対応を求めることにより, 入力画像の撮影地点の位置を取得することを目的とする. ここでは, 画像間の非類似度を系列要素間の照合コスト $d(t, u)$ とする. DP マッチングは, $d(t, u)$ の総和を系列全体の照合コストとし, それを最小とする (t, u) の軌跡を動的計画法によって計算する(図 4.2(a)). その (t, u) の軌跡が, 2 つの映像の時刻間の対応となる. DP マッチングによる映像間の対応付けは, 以下のように $d(t, u)$ の累積を評価式とし, それを最小化する t_1, t_2, \dots, t_U を決定する問題として定式化される.

$$F = \min_{t_1, t_2, \dots, t_U} \sum_{i=1}^U d(t_i, i) \quad (4.1)$$

$$d(t, u) = \|A[t] - B[u]\|. \quad (4.2)$$

式 (4.2) の $\|a - b\|$ は a, b 間の距離を表わす. ただし, t_1, t_2, \dots, t_U の選択の方法は, 図 4.2(b) に示すような傾斜制限による制約を受ける.

このように, 同一カメラ映像間の照合を行う場合は, 入力映像とデータベース映像の時間 t, u を対応付ける. しかし, これを全方位カメラ映像と通常カメラ映像のフレーム間の対応付けに適用する場合, カメラ間の視野角の違いが問題となる. また, 車両の位置ず

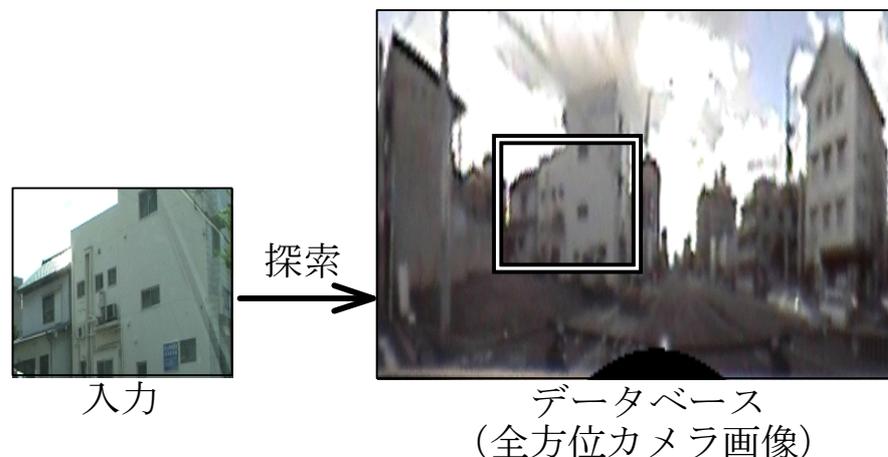


図 4.3 空間方向の探索．パノラマ展開された全方位カメラ画像上にサブウィンドウを設置し，入力画像との照合を行う．

れ，走行中の振動等によって生じる画像の見えの違いも生じる．そのため，データベース映像の時間軸方向だけでなく，空間方向の対応部分も全方位カメラ画像中から探索する必要がある（図 4.3）．また，市街地では建築物の窓枠等，類似する部分が多く存在するため，正確な対応付けを行うためには，空間方向の探索に対しても時系列情報の利用が必要である．

4.2.2 基本戦略

提案手法では，全方位カメラ映像と通常カメラ映像の対応付けのために，次の 2 つのアプローチをとる．

- 全方位カメラ映像と通常カメラ映像の時間軸方向と空間方向の同時対応付け
- 時系列情報を利用した精度の向上

提案手法では，時間軸方向と空間方向の対応付けを定式化し，これを動的計画法で解くことで，図 4.1(b) のように時間方向と空間方向の対応付けを同時に行う．これは，DP マッチングの拡張と考えられるため，ここでは拡張 DP マッチングと呼ぶ．全方位カメラは通常カメラに比べ視野角が大きいいため，全方位カメラ映像にサブウィンドウを設置し，サブウィンドウと入力画像を照合する．そして，サブウィンドウの平行移動と拡大縮小を行う

ことで、見えの変動に対応する。次に、映像間の照合コストの評価式に、全方位カメラ映像上に設置したサブウィンドウの平行移動・拡大縮小を制御するパラメータを導入する。動的計画法を用いてこの評価式を最小化することで、系列全体として最適となる映像間の対応付けを求める。

4.3 異種カメラ間の対応付け手法

提案手法では、全方位カメラ映像と通常カメラ映像の間を拡張 DP マッチングにより対応付ける。そのためには、動的計画法が適用できるように、映像間の照合コストを適切に設定する必要がある。具体的には、最小化する評価式を、動的計画法の漸化式に分解できるような形で表す必要がある。

自車位置推定の全体の流れを図 4.4 に示す。以降、全方位カメラ映像と通常カメラ映像のフレーム間対応付けの詳しい処理の内容を説明する。

4.3.1 画像間の照合

データベース構築時に全方位カメラを使用し、位置推定時に通常カメラを使用する場合、カメラの違いや撮影条件の違いにより、それぞれで得られる映像は見えが異なる。2つの映像間の差異は、次の2つに分けられる。

- 時間的な差異
- 空間的な差異

時間的な差異は、車両の走行速度の違いより発生する映像の時間方向の伸縮である。空間的な差異は、視野角の違いや視点の違いである。

まず、視野角の違いに対応するために、全方位カメラ上にサブウィンドウを設置し、サブウィンドウと入力画像を照合する。本論文では、全方位カメラと通常カメラを設置する位置は異なり、各カメラの水平方向は既知であるという状況を想定する。同一車両の同一位置にカメラを設置し、かつ車両の走行軌跡が同じ場合は、視点の違いを考慮する必要はなく、サブウィンドウの位置や大きさを固定すればよい。しかし実際には、カメラを設置する位置の違いや車両の走行軌跡の違いの他、走行中の振動で生じるカメラの角度変化により視点の違いが発生する。そこで、サブウィンドウの平行移動と拡大縮小を行うこと

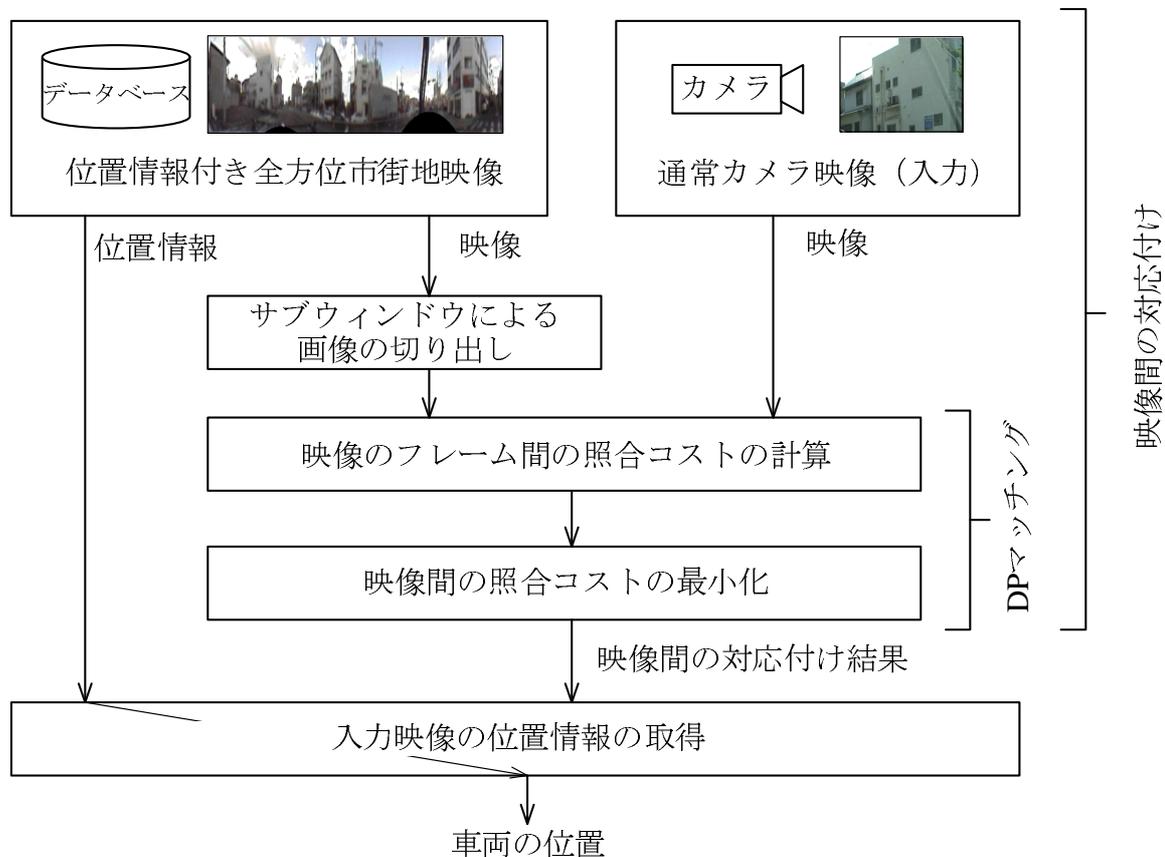


図 4.4 自車位置推定の流れ.

で、見えの違いに対処する。車両が直進する場合には、サブウィンドウの水平方向の平行移動は変化が小さい。しかしながら、鉛直方向の平行移動は車両の走行軌跡の違いによって大きく変化する。そのため、平行移動は鉛直方向のみ行う。

まず、データベース画像の系列 $\{A[1], \dots, A[t], \dots, A[T]\}$ と入力画像の系列 $\{B[1], \dots, B[u], \dots, B[U]\}$ が与えられる。データベース画像系列は、普及型 GPS または人手による位置情報によって、入力画像系列の走行経路を含む区間に切り出されているとする。次に、時刻 t のデータベース画像から、平行移動・拡大縮小を制御するパラメータ v, s に対応するサブウィンドウを切り出す。 v, s は離散値 ($v = 1, 2, \dots, V$, $s = 1, 2, \dots, S$) であり、サブウィンドウの平行移動の量と拡大縮小の倍率は v, s の値に応じ、一定値刻みで設定する。以降、この切り出された画像を $\{A[v, s, t]\}$ と表記する。

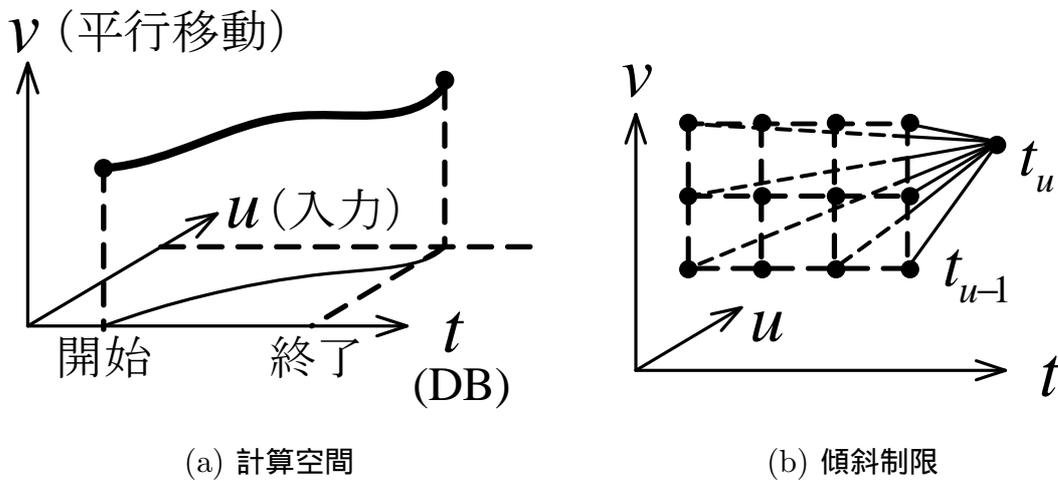


図 4.5 (a) 映像間対応付けコスト最小化処理のための計算空間．時刻 t, u に加え，サブウィンドウの平行移動を制御するパラメータ v を映像間の対応付けに導入する．(b) 極端な変化を抑制する制約を与える．

サブウィンドウ $A[v, s, t]$ と入力画像 $B[u]$ の照合コストは次式のように表される．

$$d(v, s, t, u) = \|A[v, s, t] - B[u]\|. \quad (4.3)$$

カメラ間の色合いの違いを軽減するために，入力画像とデータベース画像の各サブウィンドウに色ヒストグラムの平坦化処理を行う．また，画像間の距離には，サブウィンドウの面積により正規化した画像の画素値からなるベクトル間の L1 距離を用いる．

4.3.2 映像間の照合コスト

データベース映像の時間軸方向のパラメータ t に加え，サブウィンドウの平行移動・拡大縮小を制御するパラメータ $v = 1, 2, \dots, V$ および $s = 1, 2, \dots, S$ を映像間の照合コストの評価式に導入する．自転車位置推定を行うためには，ある入力映像の時刻に対応するデータベース映像の時刻を一意に決定する必要がある．すなわち，ある u に対応する v, s, t の組が 1 つだけ存在する必要がある．そのため，提案する拡張 DP マッチングは，図 4.5(a) のように，照合コストが系列全体で最小となるように， u に対する v, s, t の軌跡を $\{(v_1, s_1, t_1, 1), \dots, (v_u, s_u, t_u, u), \dots, (v_U, s_U, t_U, U)\}$ として求める．図 4.5 では，拡大縮小を制御するパラメータ s を省略し， $v-t-u$ の 3 次元の空間で表示しているが，実際は $v-s-t-u$ の 4 次元の空間である．

ある u に 1 つの v, s, t の組が対応するため，系列全体での照合コストを次式で表わす．

$$F = \min_{\substack{(v_1, s_1, t_1) \\ \vdots \\ (v_U, s_U, t_U)}} \sum_{i=1}^U d(v_i, s_i, t_i, i) \quad (4.4)$$

この式は，後の 4.3.4 節で述べるように動的計画法で解くことができる．動的計画法では，系列全体での照合コストが最小となるように $(v_1, s_1, t_1), \dots, (v_U, s_U, t_U)$ を決定する．また，式 (4.1) と同様に， $(v_1, s_1, t_1), \dots, (v_U, s_U, t_U)$ の選択には，図 4.5(b) のような傾斜制限により制約を与える．

4.3.3 傾斜制限

映像間の誤対応を抑制するため 映像の撮影条件に応じ 式 (4.4) の $(v_1, s_1, t_1), \dots, (v_U, s_U, t_U)$ の選択方法に制約を加える．提案手法では，次のような仮定を設ける．

- (a) 車両の最大速度は限られており，自車位置が短時間に大きく変化することはない
- (b) 車両は前進するのみで後退しない

(a) と (b) の仮定により，入力映像の時刻 $u - 1$ に対応するデータベース映像の時刻 t_{u-1} と，時刻 u に対応する時刻 t_u の差はしきい値 θ_t 以内とする．

$$0 \leq t_u - t_{u-1} \leq \theta_t \quad (4.5)$$

次に，(a) の仮定により，平行移動・拡大縮小の変化量を制限できる．例えば，時刻 u に対応する平行移動のパラメータ v_u は，時刻 $u - 1$ に対応する v_{u-1} の近傍になる．そのため，平行移動の条件はしきい値 θ_v を用いて次式のように表される．

$$-\theta_v \leq v_u - v_{u-1} \leq \theta_v \quad (4.6)$$

同様に，拡大縮小の条件もしきい値 θ_s を用いて次式のように表される．

$$-\theta_s \leq s_u - s_{u-1} \leq \theta_s \quad (4.7)$$

4.3.4 拡張 DP マッチングのアルゴリズム

系列間の照合コストが最小となるように，式 (4.4) を $(v_1, s_1, t_1), \dots, (v_u, s_u, t_u), \dots, (v_U, s_U, t_U)$ について最小化する．その際，照合が系列全体で最適となるように，動的

[Step1] Initialization

計算空間の範囲外に ∞ を代入する．すなわち，範囲が $1 \leq v \leq V$ かつ $1 \leq s \leq S$ かつ $1 \leq t \leq T$ かつ $1 \leq u \leq U$ 外するとき，

$$g(v, s, t, u) = \infty$$

[Step2] Iteration

$u = 1, 2, \dots, U$ について，次の漸化式を計算する．

for $u = 1$

for $v = 1, 2, \dots, V, s = 1, 2, \dots, S, t = 1, 2, \dots, T$

$$g(v, s, t, u) = d(v, s, t, u)$$

for $u = 2, 3, \dots, U$

for $v = 1, 2, \dots, V, s = 1, 2, \dots, S, t = 1, 2, \dots, T$

式 (4.8) を計算する．

[Step3] Output

入力画像 U が入力されたときに，映像間の照合コストが最小となるパラメータ v_U, s_U, t_U を出力する．また， v_u, s_u, t_u ($u = 1, 2, \dots, U - 1$) が必要ならば，バックトラックを行う．

$$(v_U, s_U, t_U) = \arg \min_{(v, s, t)} g(v, s, t, U)$$

図 4.6 拡張 DP マッチングのアルゴリズム．

計画法 [88] を用いて解く．式 (4.4) は付録 B の式 (B.10) と同形である．したがって，式 (4.4) は次の漸化式を $u = 1, 2, \dots, U$ について計算することで，最小化を行うことができる．

$$g(v, s, t, u) = d(v, s, t, u) + \min_{\substack{-\theta_v \leq p_v \leq \theta_v \\ -\theta_s \leq p_s \leq \theta_s \\ 0 \leq p_t \leq \theta_t}} \{g(v - p_v, s - p_s, t - p_t, u - 1)\} \quad (4.8)$$

図 4.5 にこの漸化式の計算空間の概略を示す．このとき，式 (4.4) は次式として求まる．

$$F = \min_{v, s, t} g(v, s, t, U) \quad (4.9)$$

式 (4.9) で最小値を与える v, s, t が，入力画像 $u = U$ に対応するサブウィンドウの制御パラメータおよびデータベース画像となる．このときのデータベース映像の時刻 t に対応する位置が現在の推定位置となる．具体的な拡張 DP マッチングのアルゴリズムを図 4.6 に示す．

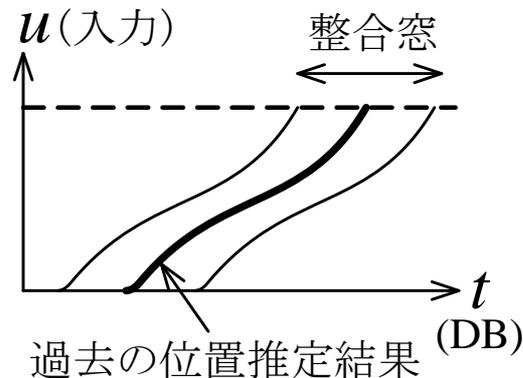


図 4.7 計算範囲の制限.

大規模なデータベースを用いる場合、 t の取り得るすべての範囲に対して式(4.8)を計算することは、記憶容量・計算量の観点から現実的ではない。そこで、DP マッチングの計算空間中で計算を行う範囲を制限する手法である整合窓 [70] を導入することで、計算範囲を絞り込む。具体的には、過去の自車位置結果から一定の範囲内を計算を行う範囲として指定する(図 4.7)。なお、自車位置推定の開始時の初期範囲は、人手または普及型 GPS 等により取得する。

4.3.5 データベース映像の正規化

データベース車両と位置推定車両の速度の差が極端に大きい場合、傾斜制限により映像の時間軸の伸縮の範囲が制限されるため、データベース映像と入力映像の対応付け処理の際に問題となる。具体的には、提案手法で用いる傾斜制限では、入力映像撮影時の速度はデータベース映像撮影時の速度の θ_t 倍以下である必要がある。また、データベース映像撮影時に完全に車両が停止していたフレームが存在する場合、DP マッチングで用いる傾斜制限の種類によっては、対応付けが困難になる場合がある。例えば、図 4.5 に示す入力映像の時間軸方向とデータベース映像の時間軸方向で非対称な傾斜制限では、入力映像撮影時に停車していたフレームが存在しても対応付けが可能であるものの、データベース系列撮影時に停車していたフレームが存在する場合には対応付けができない。そこで、データベース映像と入力映像の対応付け処理を簡便にするために、データベース映像のフレーム間の長さが一定になるように、フレームの間引きと挿入によるリサンプリングを行う(図 4.8)。リサンプリングの間隔は、入力映像撮影時(自車位置推定時)に想定する最大

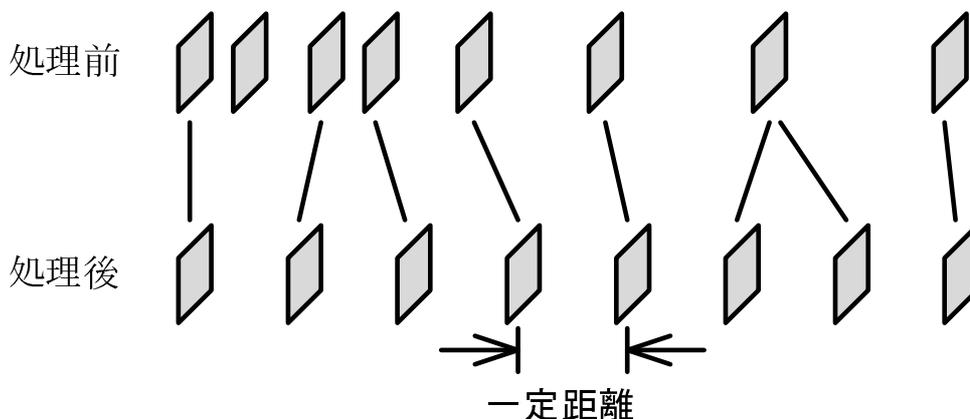


図 4.8 速度正規化のためのデータベース映像の正規化。

速度や映像のフレームレート，傾斜制限のパラメータ θ_t によって決定する．具体的には，入力映像撮影時の車両の最大速度が a ，フレームレートが f_s の場合，データベース映像のフレーム間の長さは $a/(\theta_t f_s)$ とする．このようなりサンプリングによる正規化を行うことで，車両の速度と停車の問題を解決する．

4.4 実験

4.4.1 実験条件

提案手法の有効性を確認するために，位置推定誤差を以下の実験により評価した．実験に用いるカメラの車両への装着位置を図 4.9 に示す．データベース構築には円錐ミラー型の全方位カメラを使用し，路面から 1.5 m の位置に設置した．自車位置推定用の通常カメラは民生用の DV カメラを用い，左前方へ向け，車内の助手席のサイドガラスに固定した．また，通常カメラは路面から高さ 1.0 m の位置に設置し，正確な設置方向は人手で測定した．撮影した映像のフレームレートは全方位カメラ，通常カメラともに 30 fps であった．

経路長が 300 m と 400 m の 2 つの区間を用い，合計で 7 回の実験を行った．映像中には車両の加減速，停止が含まれている．なお，データベース構築用と自車位置推定用の映像は異なる日付に撮影したものを使用した．映像収集期間は約 6 ヶ月である．データベース内の映像の 1 フレームあたりの走行距離が，0.4 m となるように正規化したデータベー

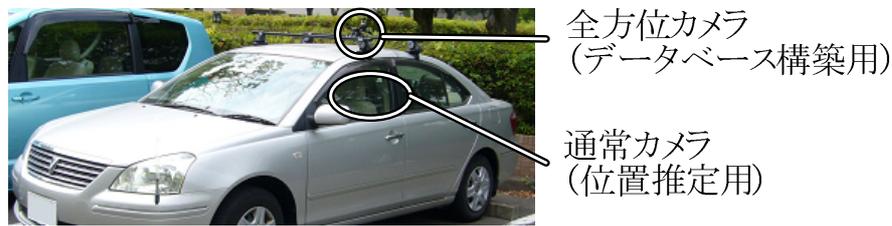


図 4.9 カメラを設置した位置 .

表 4.1 実験条件

| | 平行移動・拡大縮小 | DP マッチング |
|--------|-----------|----------|
| 比較手法 1 | なし | 従来 |
| 比較手法 2 | あり | 従来 |
| 提案手法 | あり | 提案 |

ス映像を用意した .

提案手法におけるパラメータは予備実験により得られた以下の値を使用した . まず , 式 (4.5) の θ_t は 3 とした . これにより , 位置推定車両の速度は 0 ~ 130 km/h に対応可能である . 画像間の照合における画像の平行移動は 1.44 度刻み , 拡大縮小は 0.03 倍刻みとした . また , θ_v と θ_s は 1 とした . カメラの内部パラメータはカタログ値を使用した . 入力画像の下半分は道路領域のため , 上半分のみを画像照合に使用した . 自車位置推定開始時の初期探索範囲として , 真の自車位置を中心とした約 120 m の区間を人手で与えた .

全方位カメラ映像の元の解像度は 640×480 pixel であった . パノラマ展開は球面投影によって行い , 水平方向の解像度は 180 度あたり 250 pixel とした . また , 入力画像の元の解像度は 640×480 pixel であり , パノラマ展開後の全方位カメラの解像度と通常カメラの画角を考慮し , 拡大縮小の倍率が 1 のときの解像度を 61×56 pixel として位置推定を行った .

本実験では , 表 4.1 に示すように , 提案手法を 2 つの比較手法と比較した . 比較手法 1 は画像の平行移動・拡大縮小を行わず , 式 (4.1) の従来の DP マッチングを用いた . すなわち , 画像間の距離を算出する際に , サブウィンドウを一定位置かつ一定サイズに固定し , 系列要素間の照合コスト $d(t, u)$ を算出した . 比較手法 2 は画像の平行移動・拡大縮小を行ない , 式 (4.1) の従来の DP マッチングを用いた . 具体的には , フレーム単位でサ

ブウィンドウの平行移動と拡大縮小を総当たりで行い、最小となる照合コストを系列要素間の照合コスト $d(t, u)$ とした。そのため、映像の空間方向の時系列情報は利用されない。提案手法はサブウィンドウの平行移動・拡大縮小を行ない、提案する拡張 DP マッチングを用いた。

実験結果の数値はフレーム単位で算出した。ただし、車両が停止している間は推定位置の変化が生じないため、車両が停止していたフレームを除いて評価した。

4.4.2 実験結果

進行方向の誤差が 2.0m 以内に収まる割合を図 4.10 に示す。また、提案手法と比較手法の位置推定誤差の累積分布を図 4.11 に示す。図 4.10, 4.11 より、提案手法、比較手法 2、比較手法 1 の順で良い結果が得られることが分かる。

比較手法 1 は、カメラの装着位置の違いや走行位置の違い、カメラの振動により発生する入力画像の見た目の変化を考慮しておらず、提案手法や比較手法 2 に比べ、位置推定精度が低かった (図 4.10, 4.11)。

実験環境は、CPU Xeon E5540 2.53GHz $\times 2$ 、メモリ 32G バイトであり、処理時間は入力画像 1 フレームあたり平均 0.78 秒であった。現状ではリアルタイム処理は困難であるが、計算機の性能向上や高速な DP マッチング手法の導入により、処理速度の向上が期待できる。

4.5 考察

提案手法が成功し、比較手法 1 が失敗した例を図 4.12 に示す。(b) と (c) 中の実線の矩形は (a) の入力画像に対して対応すると推定されたサブウィンドウである。(c) 中の点線の矩形は正しいサブウィンドウの位置を示している。提案手法では、(b) のように入力画像とデータベース画像が正しく対応している。一方、比較手法 1 では、(c) の 1 番目の画像のように入力画像とデータベース画像が一致する場合があるものの、2 番目の画像のように映像間の対応付けを誤る場合が多く見られた。このように、提案手法と比較手法 1 との比較から、全方位カメラ映像と通常カメラ映像間の照合のためには、画像の空間方向の探索も行う必要があることが確認できる。

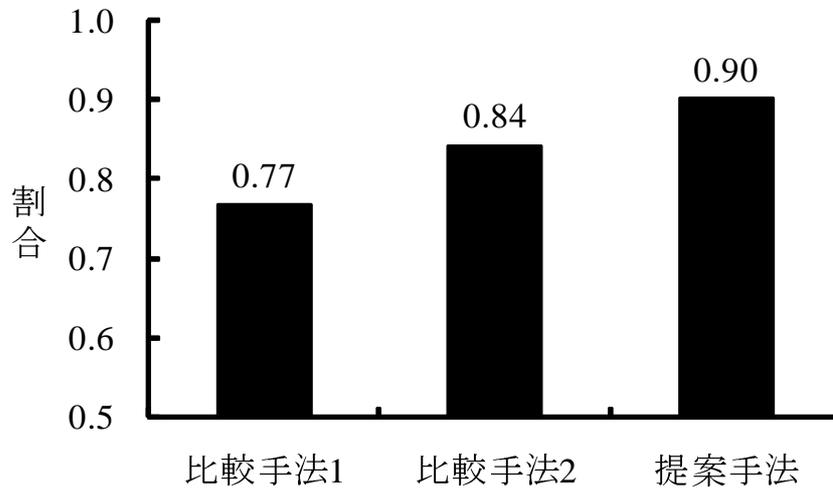


図 4.10 位置推定誤差が 2.0 m 以内に収まる割合 .

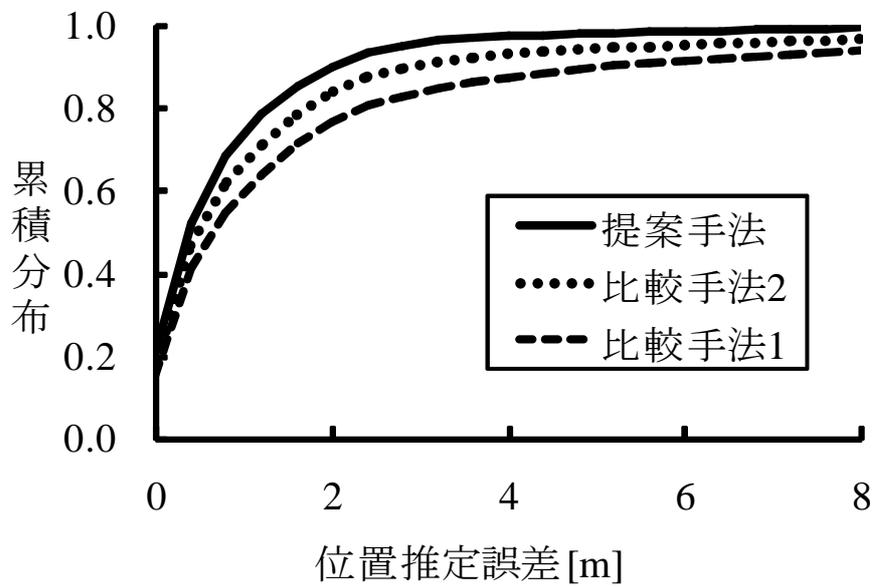


図 4.11 位置推定誤差の累積分布 .



1

2

(a) 入力映像



1

2

(b) データベース映像 (提案手法)



1

2

(c) データベース映像 (比較手法 1)

図 4.12 画像間の対応付け結果 (比較手法 1 との比較)

次に、提案手法が成功し、比較手法2が失敗した例を図4.13に示す。(a)は入力映像の連続する5フレームであり、(b)、(c)はそれに対応すると推定されたデータベース映像中のフレームを示している。(b)と(c)中の実線の矩形は(a)の入力画像に対して対応すると推定されたサブウィンドウであり、点線の矩形は正しいサブウィンドウのおおよその位置を示している。提案手法では(b)のように、サブウィンドウは入力画像と同じ部分を示している。このように、入力映像に対応するデータベース映像のフレームと、サブウィンドウの位置や大きさの推定に成功していることが分かる。一方、比較手法2では、フレーム単位でサブウィンドウの平行移動と拡大縮小を総当たりで行う。そのため、データベース画像上に類似する部分が複数存在する場合には、(c)の1, 2, 5番目の画像のように映像間の対応付けを誤ることがあった。(c)の1, 2番目の画像は、推定されたフレームはほぼ正しいものの、サブウィンドウの位置や大きさの推定に失敗した。5番目の画像は、サブウィンドウの位置や大きさの推定に加え、フレームの推定も失敗した。このように、提案手法と比較手法2の比較から、映像の時間軸方向だけではなく、空間軸方向の時系列情報も考慮する提案手法が有効であることが確認できる。

実験に用いた映像には、晴天と曇天に撮影したものが含まれており、撮影時刻はいずれも10:00~13:00であった。この条件では、位置推定失敗の原因となるような建築物の陰影等の強い影響は見られなかった。しかし、今後より多様な光源環境の下で実験を行う必要がある。その他に、通常カメラを車内に設置したため、天候や太陽の方向によっては、ガラスへの映りこみが発生し、位置推定を誤る例が見られた。また、全方位カメラの特性上、強い直射日光により、画像の一部分の明るさが大きく変化する例も見られた。そのため、今後カメラの設置方法等を検討する必要がある。

その他に、提案手法において、図4.14のように距離の異なる2つの建築物が同時に入力カメラに存在する際、位置推定誤差が増加する場合があった。これは、単純な平行移動と拡大縮小で、画像の見えの変化を近似し、単純な画素値に基づく画像特徴を用いたため、見えの変化への対処が不十分なためだと考えられる。このため、今後このような場合に対応できる画像特徴の改良が必要であろう。

実験では、通常カメラは車両の左前方に向けたが、提案手法はその他の方向も適用可能である。そこで、通常カメラを左後方に向け、前述の実験と同じ経路で、計7回実験を行った。その結果、位置推定誤差が2m以下の割合は90%であった。左前方の場合も90%であり、左前方と左後方で顕著な違いは見られなかった。

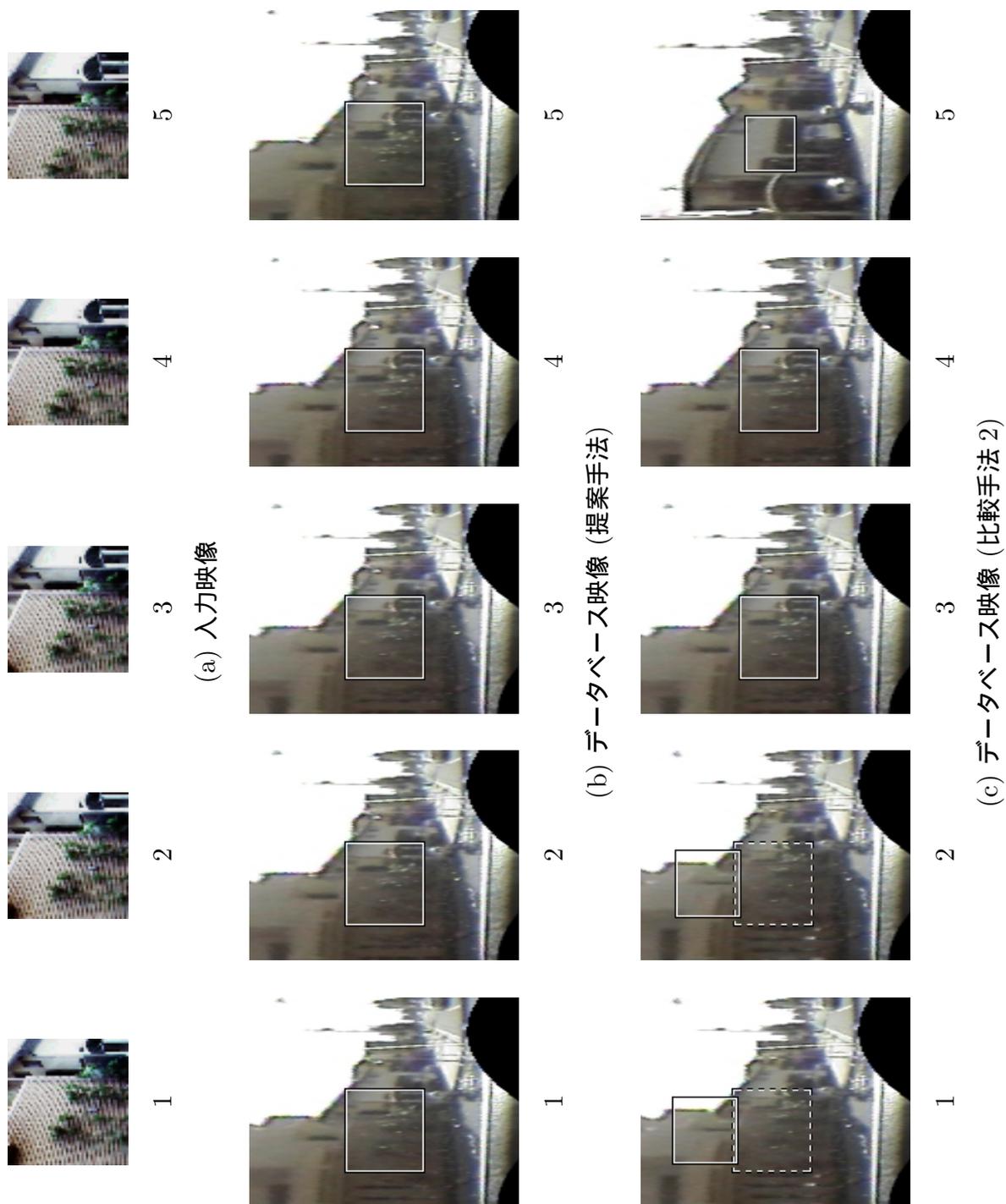


図 4.13 画像間の対応付け結果 (比較手法 2 との比較).

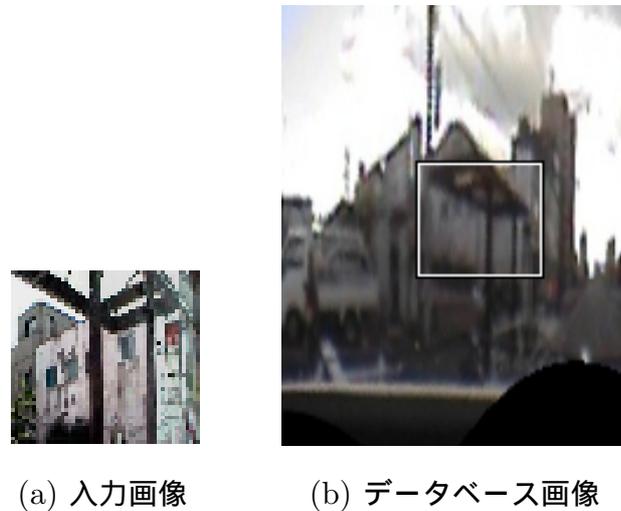


図 4.14 距離の異なる 2 つの建築物が同時に画像中に存在する場合，位置推定誤差が増加する．

実利用を考えた場合には，自車位置推定の開始時に，自車の初期位置が必要となる．おおよその初期位置を得るために，現在広く使われている普及型 GPS が利用可能である．しかし，普及型 GPS には 10～20 m の誤差が含まれるため，その誤差を考慮した一定長の範囲を参照映像として与え，提案手法により自車位置を絞り込む．このようにすることで，精度の低い初期値を基に，高精度な自車位置推定を実現する．また，一定長の範囲を参照映像として与えるため，照合計算量はデータベースの大きさに依存しない．

データベースの容量は，1 フレームの画像サイズを 4k バイト，フレームの撮影間隔を 0.4 m とすると，1 m あたり 10k バイトとなる．日本の道路の総延長が約 119 万 km であることから [87]，仮に日本全ての道路をデータベース化すれば約 12 T バイト，1 都道府県あたりにすれば平均約 253 G バイトである．実用上は，高精度な自車位置が必要となる都市部等，必要な個所だけをデータベース化すればよいとため，さらに少ない容量でよい．近年の記憶装置や通信技術の発達を考慮すれば，これは実用的な容量であると考えられる．

4.6 カメラ方向が未知の場合の異種カメラ間対応付け

4.6.1 手法概要

4.3 では，カメラの水平方向が既知である場合の手法について述べた．提案手法では，カメラの水平方向が未知の場合でも，入力映像に対応する領域をデータベース映像から探

索する際、探索方向を鉛直方向の平行移動・拡大縮小に加え、水平方向の平行移動も行うことで、自車位置推定が可能である。具体的には、拡張 DP マッチングの定式化の中に、鉛直方向の平行移動の場合と同様に、水平方向の平行移動の制約も追加する。本節では、このような手法によって、カメラ方向が未知の場合に自車位置推定を行った実験について述べる。

4.6.2 実験

4.6.2.1 実験条件

データベース構築に用いる全方位カメラとして VStone VS-C14N を使用し、車両の屋根に鉛直上向きに取り付けた。位置推定に用いる通常カメラは民生用のデジタルビデオカメラ Panasonic NV-GS150-S を使用し、車両のフロントガラスの左隅に水平に固定した。フレームレートはいずれも 30 fps である。

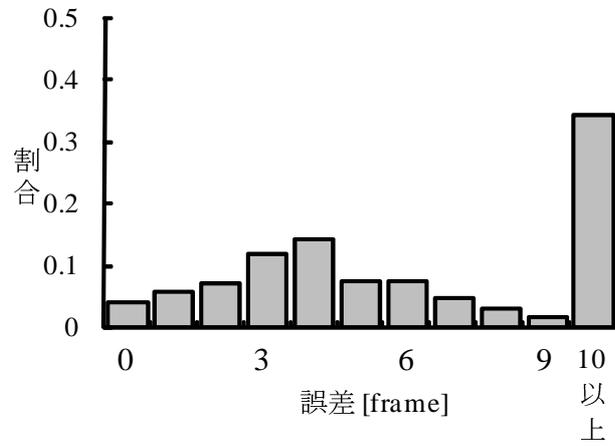
実験に用いた経路は約 400 m の片側 1 車線の直線道路である。全方位カメラ映像と通常カメラ映像をそれぞれ異なる走行で撮影した。通常カメラの撮影方向の違いによる影響を調査するために、通常カメラを次の 3 種類の方向に設置して比較した。

- (a) 前方（ほぼ正面）
- (b) 前方より左へ約 15°
- (c) 前方より右へ約 35°

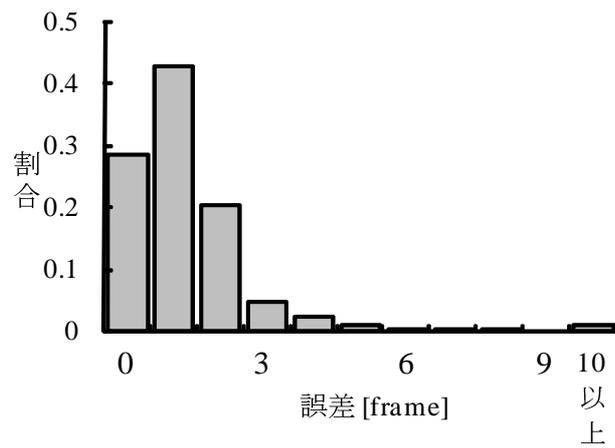
入力映像に対応する領域をデータベース映像から探索する際、理論的には探索方向について、鉛直・水平方向の平行移動、拡大縮小の 3 自由度を用いることが可能である。しかしながら、計算量が大きくなるため、本実験では拡大縮小を省略し、鉛直・水平方向の平行移動の 2 自由度のみとした。

4.6.2.2 実験結果

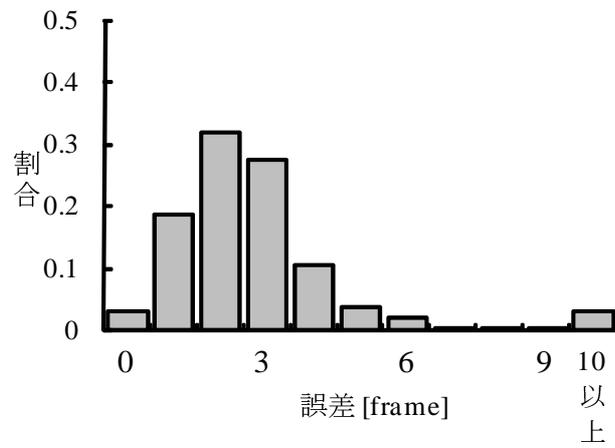
入力した通常カメラ映像のフレームに対し、提案手法で推定された全方位カメラ映像のフレームの真値との誤差（位置推定誤差）の平均は表 4.2 のようになった。図 4.15 に各カメラ方向における位置推定誤差の分布を示す。ただし、位置推定誤差を評価する際、車両が停止していたフレームは除外した。



(a) 前方



(b) 左前方



(c) 右前方

図 4.15 カメラ方向ごとの位置推定誤差の分布 .

表 4.2 位置推定誤差の平均

| カメラの方向 | 推定誤差 |
|---------|------------|
| (a) 前方 | 5.1 frames |
| (b) 左前方 | 1.3 frames |
| (c) 右前方 | 3.1 frames |

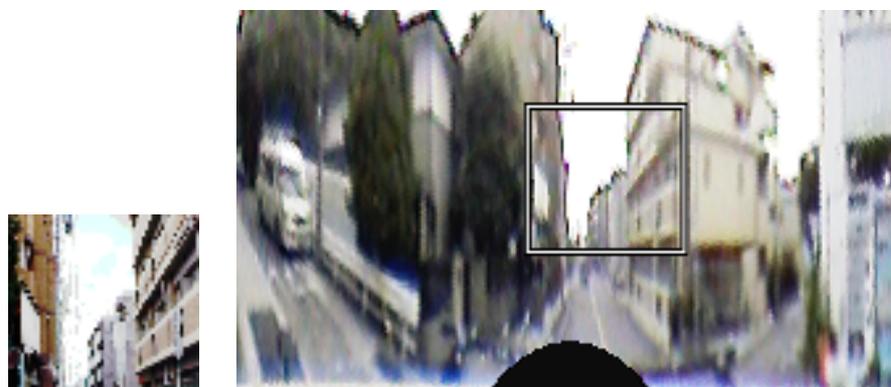
表 4.3 水平方向のカメラ角度の推定結果

| カメラ方向 | 平均 [°] |
|---------|--------|
| (a) 前方 | +10.1 |
| (b) 左前方 | -16.3 |
| (c) 右前方 | +35.8 |

ずらし窓の水平方向の座標から通常カメラの水平方向の角度が算出が可能である．参考として，表 4.3 に水平方向のカメラ角度の推定結果を示す．図 4.16 に正しく対応付けられた画像対の例をそれぞれのカメラ方向について示す．左が入力された通常カメラ画像で，右が対応する全方位カメラ画像である．全方位カメラ画像上に表示された矩形は推定された通常カメラ画像との対応部分である．入力画像と矩形内の画像は同じ領域を捉えており，正しく対応付けができたことが分かる．

4.6.2.3 考察

表 4.2 より，位置推定精度はカメラが (b) の左前方の場合が最もよく，続いて (c) の右前方，(a) の前方の順となった．(a) および (c) は (b) に比べて位置推定誤差の平均が大きくなった．図 4.17 に，通常カメラ映像内で 1~4 フレーム離れたフレーム間の距離（相違度）の平均をカメラ方向別に示す．(a) の前方の場合，通常カメラの視野角が小さいため，遠方のみが撮影される．図 4.17 から，(a) の前方は映像の見かけ上の変化が乏しいため，前後のフレームとの相違度が小さいく，区別が難しいことがわかる．そのため，位置推定精度が低下したものと考えられる．また，(c) が (b) に比べて推定誤差が大きいのも，同様の理由であると考えられる．



(a) 前方



(b) 左前方



(c) 右前方

図 4.16 対応付けの成功例．入力画像（左）とデータベース画像（右）で対応付けが正しく行われている．

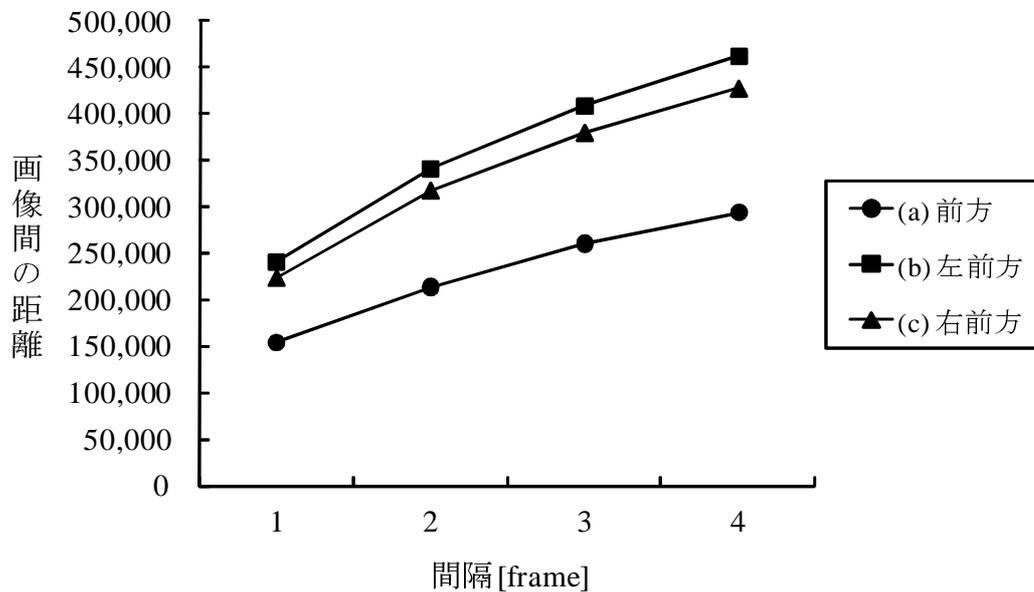


図 4.17 入力画像におけるフレーム間の距離の平均値

4.7 移動物体除去映像の自車位置推定への適用

4.7.1 概要

自車位置推定の際，データベース映像中に車両などの移動物体が存在する場合，背景が遮蔽されることにより位置推定精度が低下するという問題がある．しかしながら，第3章で提案した移動物体除去手法によりデータベース映像中の移動物体を除去した場合，この遮蔽の問題が改善されることが期待される．そこで，提案する移動物体除去手法で生成した映像を，自車位置推定の際のデータベース映像として利用する実験を行い，データベース映像中の移動物体除去の効果の評価した．具体的な評価項目として，移動物体除去前後の位置推定誤差の変化と，移動物体除去に使用する映像数と位置推定誤差の関係を調査した．以下，実験条件と実験結果について述べる．

4.7.2 実験条件

自車位置推定は4.3で説明した手法を用いた．自車位置推定時の車載カメラにも全方位カメラを使用し，映像の一部を用いることで仮想的に通常カメラ映像として扱った．フ

フレームレートは 15 fps であった。自車位置推定時の走行経路長は約 120 m であり、映像は 150 フレームであった。位置推定誤差の評価のために、真値として入力画像とデータベース画像の撮影位置が一致するフレームを人手で与えた。そして、位置推定誤差はフレーム単位で評価した。

自車位置推定のデータベースとして、比較のために以下の 2 種類を用いた。

- 移動物体除去前の映像
- 移動物体除去手法により、移動物体を除去した映像

移動物体除去に用いる映像の本数は、5～15 本で実験を行った。移動物体除去のためのパラメータは、サブウィンドウサイズが 30 pixel であり、式 (3.10) の重み λ は 0.6 であった。データベース映像のフレームレートは 15 fps であった。移動物体除去前後のデータベースで各映像の撮影位置が一致するようにするため、自車位置推定実験で用いる移動物体除去前の映像を移動物体除去の際の基準映像とした。

4.7.3 結果と考察

移動物体除去前後の映像をそれぞれデータベースとして用いた場合の位置推定誤差の違いを図 4.18 に示す。移動物体除去前は、データベース中の移動物体による遮蔽のため、位置推定誤差が 4 フレーム以上の区間の割合が大きい。移動物体除去後では位置推定誤差が 4 フレーム以上の区間はほとんど存在しなかった。また、走行区間全体としては、位置推定誤差の平均が移動物体除去前で 1.58 フレーム、移動物体除去後で 1.02 フレームであり、データベース中の移動物体除去を除去することで、大きな位置推定性能の向上が見られた。一方、移動物体除去前で遮蔽が存在する走行区間を評価から除いた場合、位置推定誤差の平均が移動物体除去前で 0.98 フレーム、移動物体除去後で 1.04 フレームであった。移動物体除去前後で遮蔽が発生していない区間の位置推定誤差にほとんど変化がなく、複数の映像の貼り合わせで生じる画質の劣化による画像間の照合精度の目立った低下は見られなかった。

移動物体除去で使用する映像数と位置推定誤差の関係を図 4.19 に示す。映像数と位置推定誤差の間で相関はなく、両者に関係性は見られなかった。これは、移動物体除去手法は実映像を貼り合わせることで映像を生成しているため、入力画像とデータベース画像間の照合には影響が少なかったものと考えられる。

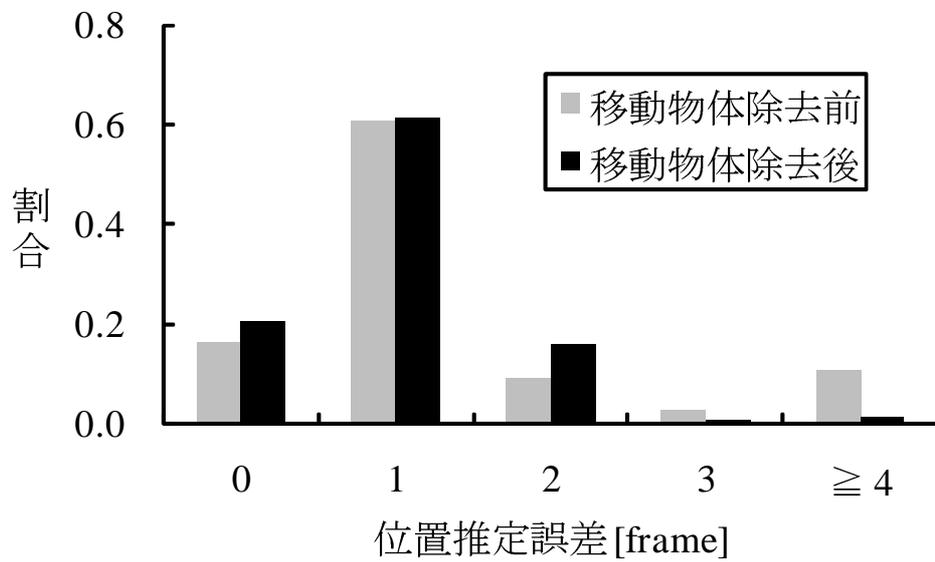


図 4.18 移動物体除去前後による位置推定誤差の違い。

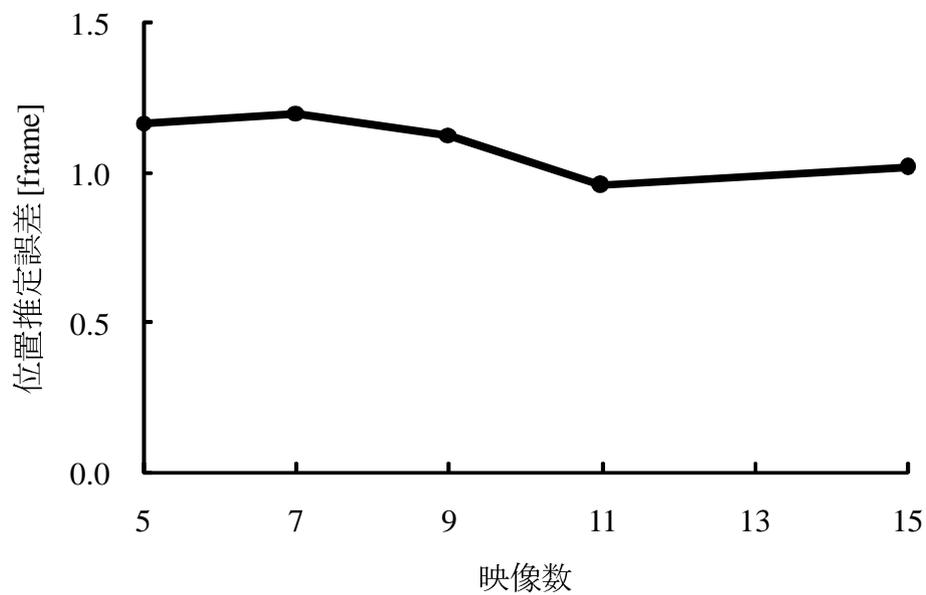


図 4.19 移動物体除去で使用する映像数と位置推定誤差の関係。

図 4.20 は、移動物体除去前のデータベースの駐車車両による遮蔽が発生している区間において、入力画像とそれに対応するデータベース画像を表したものである。データベース画像中の矩形は入力画像に対応する領域を示している。入力フレームは 5 フレーム間隔で並べてある。図 4.20 の 2, 3 行目のように、移動物体除去前では駐車車両による遮蔽によって画像間の対応付けが失敗しているが、移動物体除去後では正しく対応付けることが可能であった。

4.8 まとめ

アピアランススペースの手法による、自車位置推定のための全方位カメラ映像と通常カメラ映像のフレーム間の対応付け手法を提案した。視野角の違いに対処するため、全方位カメラ画像上にサブウィンドウを設置し、サブウィンドウの平行移動と拡大縮小を行うことで、カメラ間の見えの違いに対処した。また、平行移動と拡大縮小を制御するパラメータを拡張 DP マッチングに導入することで、映像の時間軸方向に加え、空間方向の時系列情報を利用した。実験の結果、空間方向の時系列情報を利用しない場合に比べ、位置推定精度が向上することを確認した。

提案手法では画像照合の際、単純に画像の輝度値を用いたため、距離の異なる 2 つの建築物が同時に画像中に存在する場合、位置推定誤差が増加するという問題があった。そのため、今後このような場合に対応できる画像特徴の開発が必要である。また、提案する自車位置推定手法において、第 3 章で提案した移動物体除去手法を適用した市街地映像データベースを用いることで、データベース中の遮蔽の問題が解決可能であることを確認した。しかしながら、自車位置推定時の車載カメラ映像中に移動物体が存在している場合、位置推定精度が低下する。そのため、自車位置推定時の車載カメラ映像中の移動物体による遮蔽に対処する手法の検討も必要である。

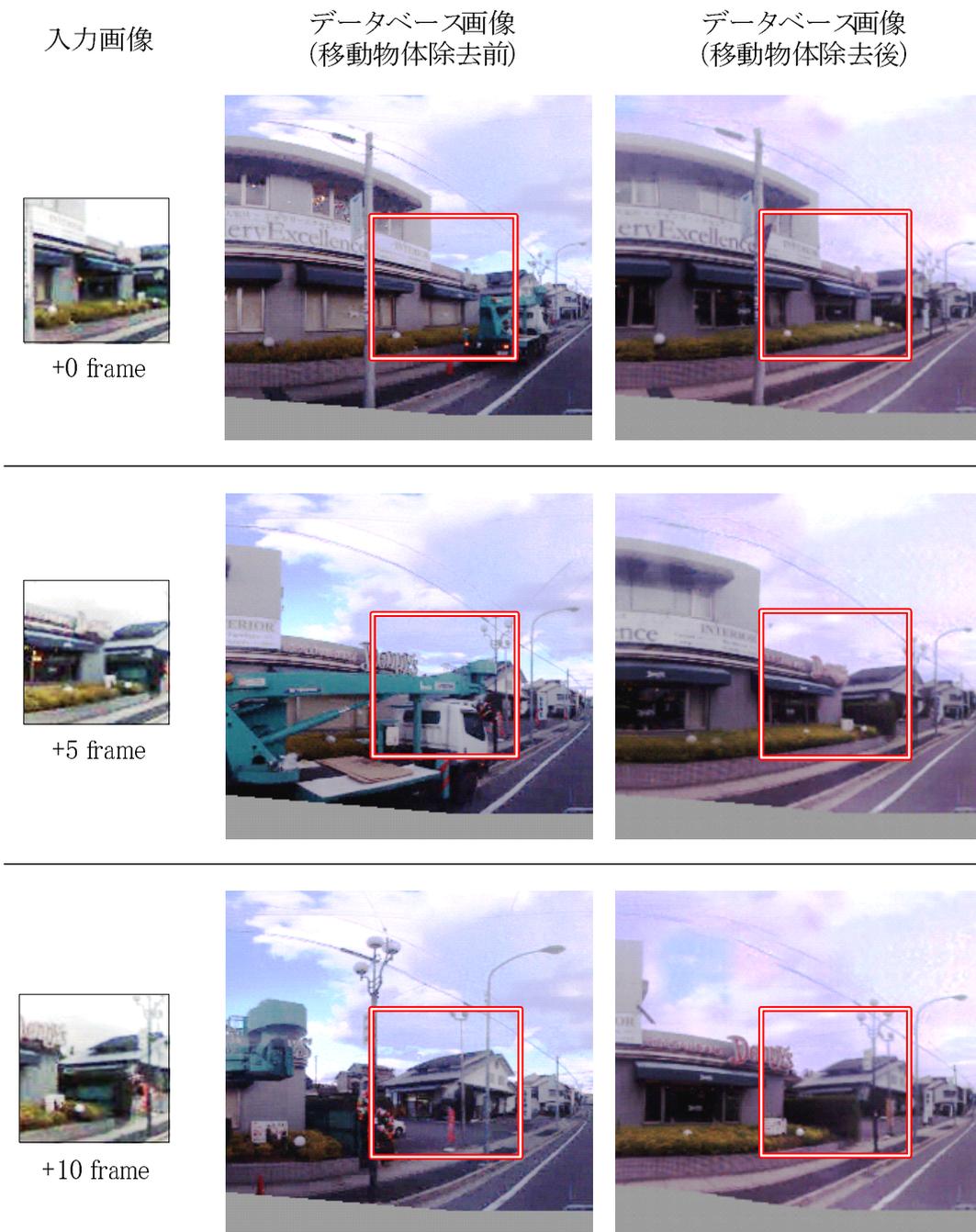


図 4.20 移動物体除去前後による画像間対応付け結果の違い。

第 5 章

2 台の車載カメラを用いた車線方向を含む走行位置推定

5.1 はじめに

第 4 章では、全方位市街地映像データベースと車載通常カメラ映像の対応付けによるアピランススペースの自車位置推定手法について述べた。しかしながら、第 4 章の手法を含め、アピランススペースの自車位置推定手法では位置推定を画像間の対応付けのみで行っている。このため 2.4.4 で述べたように、周囲環境と自車との幾何的な位置関係は考慮されないため、車両の進行方向の位置の推定は可能だが、車線方向の位置の推定は限定的である。

本章では、アピランススペースの位置推定手法で、車両の進行方向の位置だけでなく、車線方向の位置も推定可能な手法を提案する。1 台のカメラによるアピランススペースの位置推定手法では、環境の幾何的な関係を利用することが難しい。そこで、提案手法は自車位置推定時に 2 台の車載カメラを用いることで、この問題を解決する。また、2 台のカメラの照合結果を統合的に利用し、1 台のカメラをそれぞれ独立に扱うよりも高精度に自車位置推定を行う手法を提案する。1 台のみでは画像照合が不確かな区間でも、2 台のカメラの情報を統合的に利用し、補完しあうことで、位置推定精度が向上させることが可能である。

以降、5.2 で 2 台の車載カメラを用いた車線方向を含む走行位置推定手法を説明する。次に、5.3 で評価実験について述べる。5.4 で考察を行う。

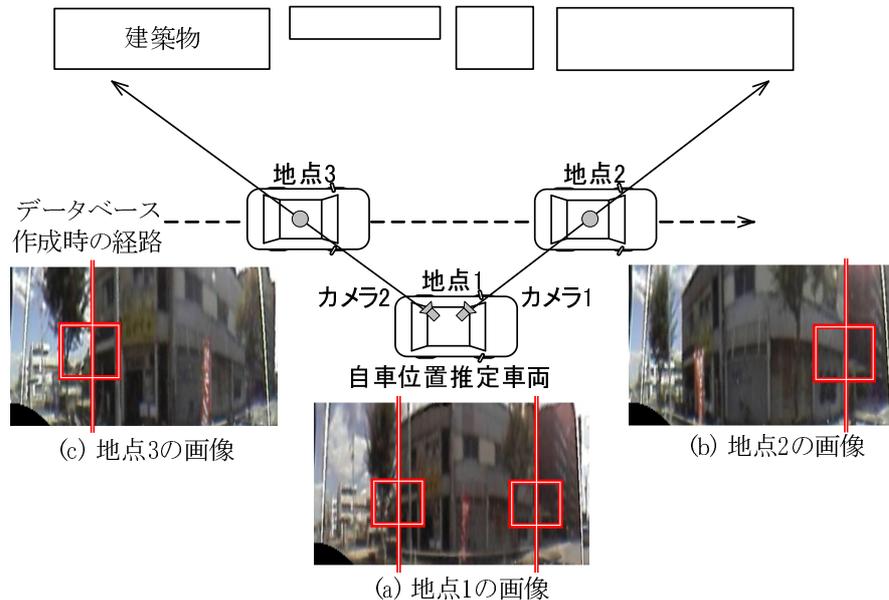


図 5.1 2 台のカメラを用いた自転車位置推定 .

5.2 2 台の車載カメラを用いた位置推定手法

5.2.1 基本戦略

2 台の車載カメラを用いた位置推定の概要を説明する．図 5.1 のように，自転車位置推定を行う車両が地点 1 にある場合，カメラ 1 から得られる画像はデータベース上の地点 2 から同じ方向を見た街並み画像の一部分に類似したものとなる．次に，カメラ 1 と異なる方向に，カメラ 2 を設置する．自転車位置推定を行う車両が地点 1 にある場合，カメラ 2 から得られる画像はデータベース上の地点 3 から同じ方向を見た街並み画像の一部分に類似したものとなる．(a)～(c) は，それぞれ全方位カメラによる地点 1～3 における画像の例である．画像中の矩形は通常カメラの視野とそれに対応する領域を模式的に図示しており，縦軸は光軸を通る鉛直な線を表している．地点 2，地点 3 はデータベース上にあるため，地点 2，地点 3 の座標がわかれば，三角測量を用いて，カメラ 1，2 の角度から地点 1 の位置を計算することができる．

提案手法では，さらに位置推定精度を向上させるため，画像の系列は経路に沿って連続に存在することを利用し，複数フレームを利用することで，2 台のカメラの画像照合結果

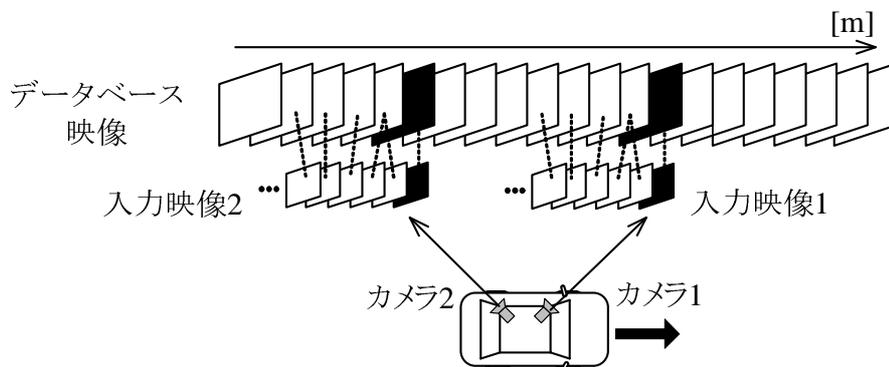


図 5.2 複数フレームを利用した 2 カメラの画像照合結果の統合 .

を統合する (図 5.2). 2 台のカメラの情報を統合的に利用し, 補完しあうことで, 2 台のカメラをそれぞれ独立に扱う場合に比べ, 位置推定精度を向上させることが可能となる. 道路を走行する車両の典型的な挙動から, 2 つの入力映像とデータベース映像のフレームの対応関係に制約を付加することができる. このため, 2 つの入力映像とデータベース映像の対応付けコストに関する目的関数を設定し, 車両の典型的な挙動を基にした制約の下で目的関数を最小化することで, 最適な 2 つの入力映像とデータベース映像の対応付けを決定する.

走行位置推定の流れを図 5.3 に示す. データベース構築時には, 第 4 章の手法と同様に, 自車位置推定時の処理の簡略化のため, データベース映像の正規化を行う. 自車位置推定時の手法は, (i) 画像間の照合コストの計算, (ii) 2 カメラの画像照合結果の統合, (iii) 三角測量の 3 つの処理に分けられる. (i) では, カメラ 1, 2 で撮影される画像と, データベース画像間の照合コストを計算する. (ii) では, (i) で得られたカメラ 1, 2 による照合コストを基にした目的関数を最小化することによって, 入力画像とデータベース画像の最適な対応付けを決定する. 最後に, (iii) では, 三角測量を用いて, 自車の位置を計算する. 以降, 提案手法の詳細を説明する.

5.2.2 画像間の照合

データベース画像の系列 $\{A[1], \dots, A[t], \dots, A[T]\}$ と, カメラ 1 から得られる入力画像の系列 $\{B^{(1)}[1], \dots, B^{(1)}[u], \dots, B^{(1)}[U]\}$, カメラ 2 から得られる入力画像の系列 $\{B^{(2)}[1], \dots, B^{(2)}[u], \dots, B^{(2)}[U]\}$ が与えられるとする. ここで, カメラ 1 とカメラ 2 の

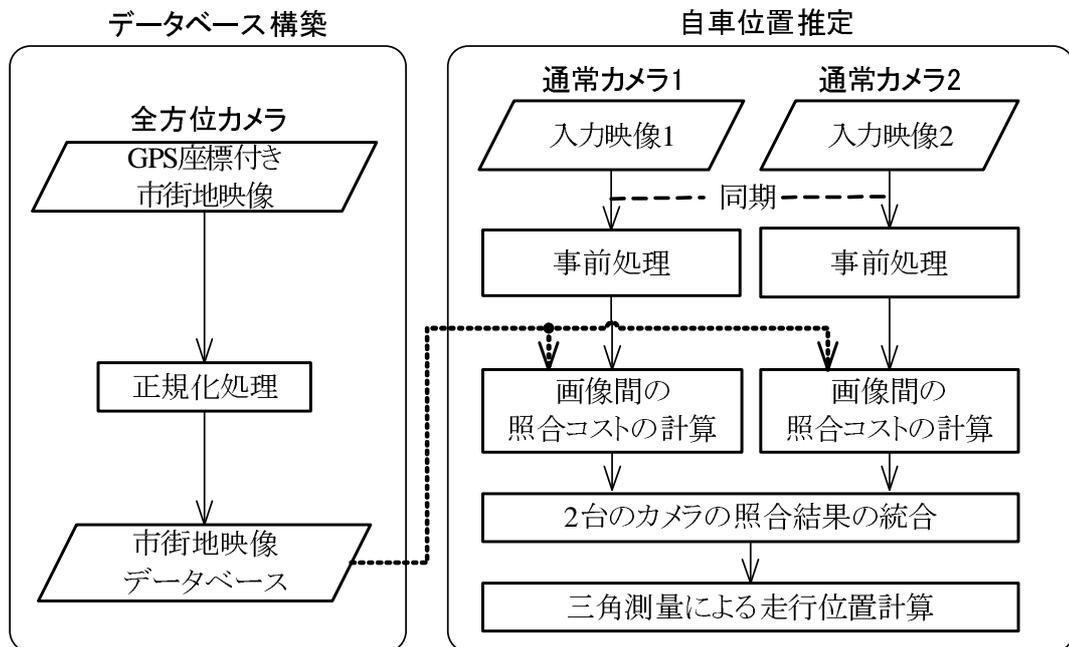


図 5.3 走行位置推定の流れ.

映像フレームは同期しているとする.

データベース映像の要素 $A[t]$ とそれぞれのカメラから得られる入力映像間の要素 $B^{(1)}[u]$, $B^{(2)}[u]$ の照合コストを計算する. u 番目のカメラ 1 の入力画像と t 番目のデータベース画像の照合コストを $f_1(t, u)$, u 番目のカメラ 2 の入力画像と t 番目のデータベース画像の照合コストを $f_2(t, u)$ と表記する.

$$f_1(t, u) = \text{dist}(A[t], B^{(1)}[u]). \quad (5.1)$$

$$f_2(t, u) = \text{dist}(A[t], B^{(2)}[u]). \quad (5.2)$$

ここで, $\text{dist}(\cdot)$ は以下で説明する画像間の距離尺度である.

本研究では全方位カメラ映像と通常カメラ映像間の照合を対象としている. そのため, $\text{dist}(\cdot)$ の最も単純な算出方法として, 入力画像 $B^{(1)}[u]$ に対応する領域をデータベース画像 $A[t]$ 中からテンプレートマッチングにより探索し, 最も類似する領域の相違度を距離尺度として利用する方法がある. しかしながら, 本手法では画像間の照合にも時系列情報を活用したい. そのため, 第 4 章の DP マッチングを用いた映像間の照合の際に算出される DP 距離を照合フレーム長で正規化したものを距離尺度とする. 式 (4.8) で計算さ

れるデータベース映像とそれぞれの入力映像を照合したときの DP 距離を $g_1(v, s, t, u)$, $g_2(v, s, t, u)$ とし, その照合フレーム長を N とすると, 式 (5.1) , (5.2) の画像間の照合コスト $f_1(t, u)$, $f_2(t, u)$ は次のようになる .

$$f_1(t, u) = \min_{v,s} \frac{g_1(v, s, t, u)}{N} \quad (5.3)$$

$$f_2(t, u) = \min_{v,s} \frac{g_2(v, s, t, u)}{N} \quad (5.4)$$

式中で平行移動, 拡大縮小のパラメータ v , s について最小をとるのは, 最も類似する領域を対応する領域として選択するためである .

5.2.3 カメラ間の拘束条件を用いた照合結果の統合

データベース映像の要素と 2 台のカメラから得られる入力映像間の要素の照合コスト $f_1(t, u)$, $f_2(t, u)$ から, 最適な入力画像とデータベース画像の対応付けを決定する . この際, 車両の典型的な挙動を基にした拘束条件から, 入力画像とデータベース画像の対応付けに制約を加える . 以降, カメラ 1, 2 の入力画像と照合されるデータベース映像のフレーム番号をそれぞれ $t^{(1)}$, $t^{(2)}$ と表記する . また, この 2 つをまとめて $\mathbf{t} = (t^{(1)}, t^{(2)})$ と表記する .

まず, 時刻 u でカメラ 1, 2 からそれぞれ入力画像が入力された場合, データベースの画像として $\mathbf{t} = (t^{(1)}, t^{(2)})$ が選択されるコスト $d(\mathbf{t}, u)$ を, $f_1(t^{(1)}, u)$ と $f_2(t^{(2)}, u)$ の和として計算する .

$$d(\mathbf{t}, u) = f_1(t^{(1)}, u) + f_2(t^{(2)}, u) \quad (5.5)$$

自車位置推定を行うためには, ある入力映像の時刻に対応するデータベース映像の時刻を, 2 つのカメラそれぞれで一意に決定する必要がある . すなわち, ある u に対する $\mathbf{t} = (t^{(1)}, t^{(2)})$ が 1 つだけ存在する必要がある . そこで, 図 5.4(a) のように, 入力映像の時刻 $u = 1, 2, \dots, U$ に対応する最適な \mathbf{t} の軌跡 $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_U$ を決定する . ある 1 つの u に 1 つの \mathbf{t} が対応するため, 次のような目的関数を設定する .

$$G(\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_U) = \sum_{i=1}^U d(\mathbf{t}_i, i) \quad (5.6)$$

目的関数は, 各時刻 u における入力画像の照合コスト $d(\mathbf{t}, u)$ の総和としている . このような目的関数を, $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_U$ について最小化することで, 入力映像 $u = 1, 2, \dots, U$ に

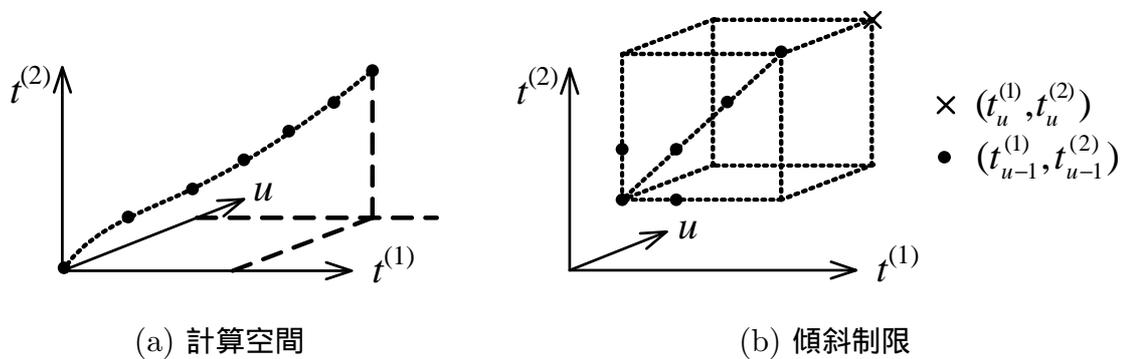


図 5.4 2カメラの統合のための計算空間。

対応する最適な t の系列 t_1, t_2, \dots, t_U を決定する。この式は、後の 5.2.5 で述べるように動的計画法で解くことができる。ただし、 $t_i = (t_i^{(1)}, t_i^{(2)})$ の選び方に対し、道路を走行する車両の典型的な挙動から、次節で述べる制約を加える。

5.2.4 車両の挙動による対応付けの制限

道路を走行する車両の典型的な挙動から、次の 2 つの条件を仮定する。

仮定 1 速度は一定値以下である

仮定 2 車線方向の位置の変化量は、進行方向の位置の変化量に比べて小さい

以上の仮定を満たす $t = (t^{(1)}, t^{(2)})$ の組を探索することにより、誤対応を削減する。 $(t^{(1)}, u)$ 、 $(t^{(2)}, u)$ の軌跡を模式的に図 5.5(a) の実線で示す。また、 u を媒介変数とした $t_u = (t_u^{(1)}, t_u^{(2)})$ の軌跡を図 5.5(b) の実線で示す。詳細は以降で説明するが、2 つの仮定に従う場合、軌跡は実線のようなになる。一方、仮定に従わない場合には軌跡は破線のようなになるため、このような経路を除外する。以降、仮定 1、2 に基づき、 $t_i = (t_i^{(1)}, t_i^{(2)})$ の選択方法に付加する具体的な制約を導出する。

仮定 1 の“速度は一定値以下である”という条件により、入力映像の時刻 u と $u-1$ に対応するデータベース画像の番号はそれぞれ近接したものとなる。これは、しきい値 θ_1

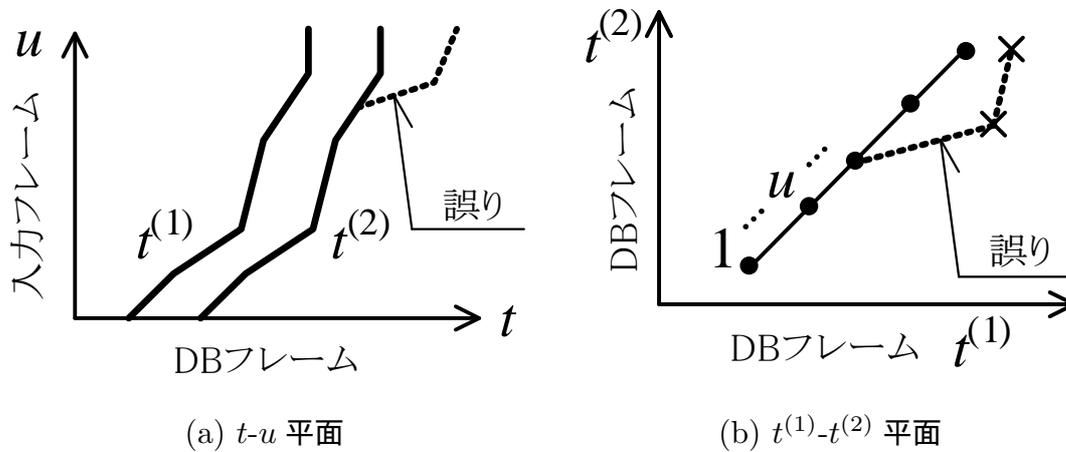


図 5.5 2カメラ間の拘束条件を用いた照合結果の統合。

を用いて、次のような条件として表すことができる。

$$0 \leq t_i^{(1)} - t_{i-1}^{(1)} \leq \theta_1 \quad (5.7)$$

$$0 \leq t_i^{(2)} - t_{i-1}^{(2)} \leq \theta_1 \quad (5.8)$$

この条件より、図 5.5(b) のサンプル点間の長さに制約が生じる。この条件は、具体的には自車位置推定時の速度がデータベース構築時の速度の θ_1 倍まで許容されることに相当する。

仮定 2 の“車線方向の位置の変化量は、進行方向の位置の変化量に比べて小さい”という条件により、図 5.5(a) のように、 u の変化量に対する $t^{(1)}$ と $t^{(2)}$ の間隔の変化量は小さい。これは、しきい値 θ_2 を用いて、次のような条件として表すことができる。

$$|(t_i^{(1)} - t_i^{(2)}) - (t_{i-1}^{(1)} - t_{i-1}^{(2)})| \leq \theta_2 \quad (5.9)$$

ここで、 $|\cdot|$ は絶対値記号である。また、 θ_2 の値が大きいほど車線方向の位置の変化を許容する。この条件より、図 5.5(b) の u を媒介変数とした $t_u = (t_u^{(1)}, t_u^{(2)})$ の軌跡は、局所的に観察した場合、傾きが 1 の直線となる。

式 (5.6) の目的関数の最小化を行う際、式 (5.7)、(5.8)、(5.9) の条件を満たすように t_1, t_2, \dots, t_U を決定する。式 (5.6) の t_1, t_2, \dots, t_U の制約の一例として、以下のような

ものが可能である．

$$(t_i^{(1)}, t_i^{(2)}) \in \left\{ (t_{i-1}^{(1)} + 0, t_{i-1}^{(2)} + 0), (t_{i-1}^{(1)} + 1, t_{i-1}^{(2)} + 1), \right. \\ (t_{i-1}^{(1)} + 2, t_{i-1}^{(2)} + 2), (t_{i-1}^{(1)} + 3, t_{i-1}^{(2)} + 3), \\ \left. (t_{i-1}^{(1)} + 2, t_{i-1}^{(2)} + 3), (t_{i-1}^{(1)} + 3, t_{i-1}^{(2)} + 2) \right\} \quad (5.10)$$

これは， θ_1 を 3， θ_2 を 1 とした例である．実際には，車両の速度や進行方向に対する車線方向の変化の許容量を考慮して θ_1 ， θ_2 を決定する．

5.2.5 照合結果の統合のアルゴリズム

式 (5.6) の目的関数を動的計画法 [88] を利用することにより， t_1, t_2, \dots, t_U について最小化する．この最小化問題は付録 B の式 (B.10) と同形であり，以下の漸化式を計算することで解くことができる．

$$g(\mathbf{t}_u, u) = d(\mathbf{t}_u, u) + \min_{\mathbf{t}_{u-1}} \{g(\mathbf{t}_{u-1}, u-1)\} \quad (5.11)$$

ただし， $u = 1$ のとき， $g(\mathbf{t}_u, u) = d(\mathbf{t}, u)$ であり， $t_i = (t_i^{(1)}, t_i^{(2)})$ の選び方は式 (5.10) に従う．具体的な導出方法を付録 B に示す．

5.2.6 三角測量による走行位置の計算

ここまでの処理により，カメラ 1，2 に対応するデータベース上の地点 A，B の座標が得られる（図 5.6）．自車の位置 (x_1, y_1) は，地点 A (x_A, y_A) ，B (x_B, y_B) ，カメラの角度 α_1 ， α_2 ，カメラ間の距離 d_1 を用いて，次の連立方程式を解くことで得られる．

$$\begin{cases} y_1 + w_1 - y_A &= (x_1 - x_A) \tan \alpha_1 \\ y_1 + w_1 - y_B &= -(x_1 - x_B - d_1) \tan \alpha_2 \end{cases} \quad (5.12)$$

5.2.7 データベース映像の正規化

5.2.2 の入力画像とデータベース画像の照合コストの計算と 5.2.4 の 2 台のカメラの統合処理で映像の時間軸の伸縮の範囲が制限されるため，データベース車両と位置推定車両の速度が極端に大きい場合，データベース映像と入力映像の対応付け処理の際に問題となる．そこで，第 4 章の手法と同様に，自車位置推定時の処理の簡略化のため，データベース映像の正規化を行う．正規化方法としては，4.3.5 で述べたように，データベース映像

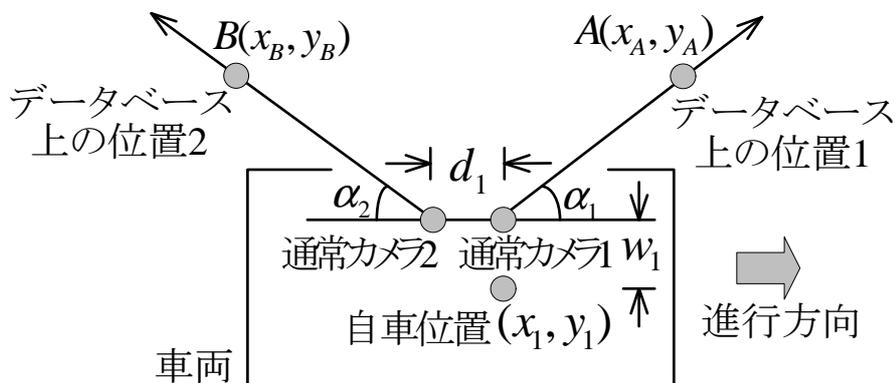


図 5.6 三角測量による走行位置の計算 .

のフレーム間の長さが一定になるように，フレーム間の間引きと挿入によるリサンプリングを行う．サンプリング間隔は式 (5.3)，(5.4) の入力画像とデータベース画像の照合コストの計算と，式 (5.6) の 2 台のカメラの統合処理の制限を考慮し，制限が厳しい方を採用する．

5.3 実験

5.3.1 実験条件

提案手法の有効性を確認するために，走行車線の認識率，進行方向の位置推定誤差，データベース構築時と位置推定時で同じ車線を走行した場合の車線方向の位置推定誤差を以下の実験により評価した．

実験に用いるカメラの車両への装着位置を図 5.7 に示す．データベース作成に用いる全方位カメラとして VStone 社製 VS-C14N を使用し，路面から高さ $h_2 = 1.5 \text{ m}$ の位置に設置した．撮影した映像のフレームレートは 30 fps であった．全方位カメラ映像の元の解像度は $640 \times 480 \text{ pixel}$ であった．パノラマ展開は球面投影によって行い，水平方向の解像度は 180 度あたり 250 pixel とした．自車位置推定用の通常カメラは 2 台とも民生用の DV カメラで，車内の助手席側サイドガラスに固定した．焦点距離は 35 mm 判換算で前向きカメラが 44.7 mm，後ろ向きカメラが 48.0 mm であった．また，フレームレートはいずれも 30 fps であった．入力画像の元の解像度は $640 \times 480 \text{ pixel}$ であり，パノラマ展開後の全方位カメラの解像度と通常カメラの画角を考慮し，拡大縮小の倍率が 1 のときの

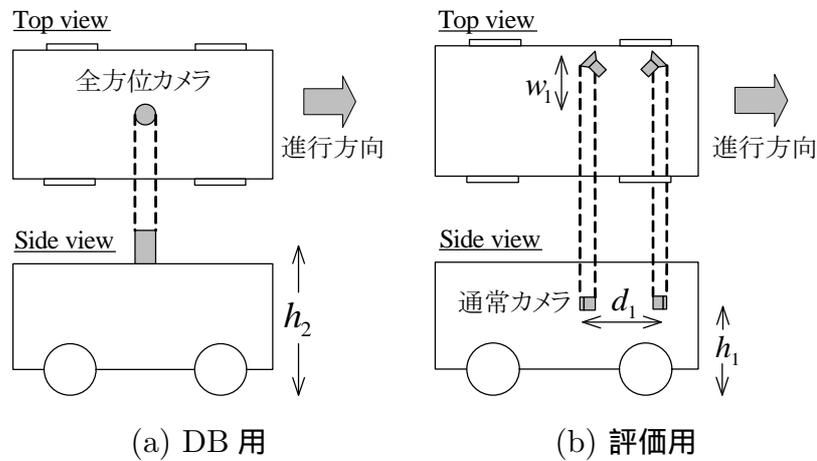


図 5.7 カメラの装着位置 .

解像度を 61×56 pixel として位置推定を行った．設置位置はいずれも車両の中心線から進行方向に対して左側に $w_1 = 0.5$ m 離れており，2 つの通常カメラの間隔は $d_1 = 0.2$ m とした．また，設置高さは $h_1 = 1$ m であった．経路長が 300 m と 400 m の 2 つの区間を用い，合計で 12 回の実験を行った．経路はほぼ直線の片側 2 車線の道路であり，走行中に車両の加減速，停車，車線変更を含む．

DP マッチングの各パラメータは予備実験により求めたものを使用した．データベース正規化後のフレーム間の位置間隔は 0.4 m であり，DP マッチングにおける画像の平行移動は 1.44 度刻みで 35 段階，拡大縮小は 0.03 倍刻みで 20 段階とした．式 (4.5) の θ_t は 3，式 (4.6) の θ_v は 1，(4.6) の θ_s は 1 とした．また，式 (5.7)，(5.8)，(5.9) の θ_1 は 3， θ_2 は 1 とした．これにより，位置推定車両の速度は 0 ~ 130 km/h に対応可能である．各評価用カメラの撮影方向は目視で求めた．入力画像の下半分は道路領域であるため，上半分のみを画像照合に使用した．また，実験結果の評価に用いる車両進行方向の位置の真値は目視で与えた．

本実験では，提案手法を 2 つの比較手法と比較した．比較手法 1 では，カメラ 1 台で位置推定を行った．比較手法 2 では，2 つのカメラ間の統合処理を行わず，2 台のカメラを独立で扱った．すなわち，2 つの映像に対してそれぞれ第 4 章の 1 台のカメラを用いた自車位置推定手法を適用した後，5.2.3 の処理を適用せずに三角測量を行って走行位置を求めた．

評価指標は進行方向の位置推定誤差と走行車線認識率である．映像の取得時に自車の位

置が車線内で不明なため，車線方向の位置推定精度として走行車線認識率を評価した．走行車線認識は，車線幅が 3.0 m かつデータベース構築時に車両は走行車線の中心を走行していると仮定し，車線方向の推定位置が $-4.5 \sim -1.5$ m の場合はデータベース中の車線に対して右隣の車線， $-1.5 \sim 1.5$ m の場合は同じ車線， $1.5 \sim 4.5$ m の場合は左隣の車線と判別した．

5.3.2 実験結果

以下の実験結果の数値は，車両が停止していたフレームを除き，フレーム単位で集計した．また，走行中に車線変更が含まれるものは，車線変更の前後で分け，それぞれ異なる車線の状態として集計した．

本手法および比較手法 1, 2 による進行方向の位置推定誤差が一定範囲に収まる割合を図 5.8 に示す．また，走行車線認識率を図 5.9 に示す．

図 5.8, 5.9 の結果から，提案手法は良好に自車位置推定が行えたことがわかる．図 5.8 のように，進行方向の位置推定誤差は提案手法，比較手法 2，比較手法 1 の順でよい結果が得られた．また，図 5.9 のように，走行車線認識率は比較手法のほうが比較手法 2 よりも良好な結果が得られた．

CPU Core i7 870 2.93GHz，メモリ 8GB の計算環境で，提案手法の処理時間は入力画像 1 フレームあたり平均 4.4 秒であった．また，比較手法 2 では入力画像 1 フレームあたり平均 4.2 秒であり，大部分の計算時間は画像間の照合コストの計算であった．

5.4 考察

本節では，以下について考察する．

- 5.3 の実験結果に対する考察
- 車線方向の位置推定誤差に関する考察
- 時間帯による影響

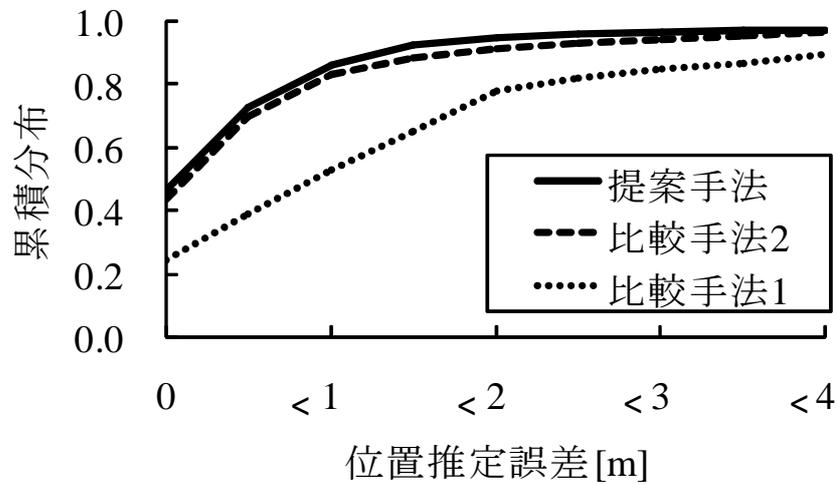


図 5.8 進行方向の位置推定誤差が一定範囲に収まる割合。

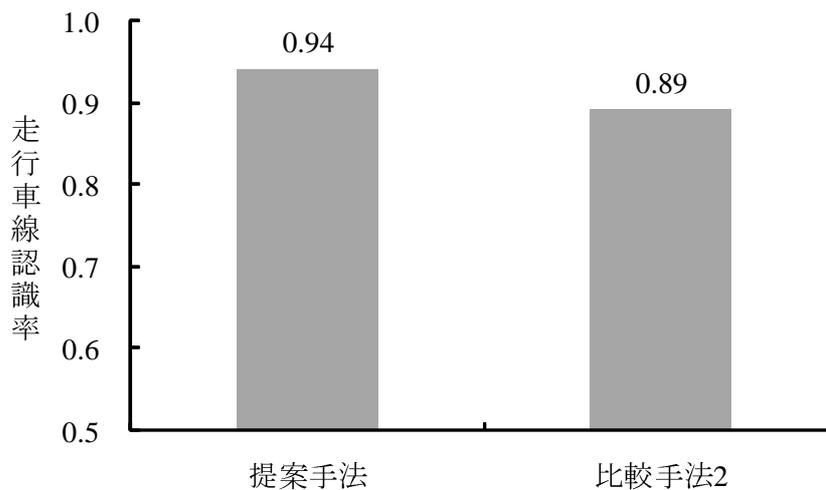


図 5.9 走行車線認識率。

5.4.1 実験結果に対する考察

図 5.8, 5.9 のように, 比較手法 2 に比べ, 提案手法の位置推定精度が向上した。これは, 1 台のみでは画像照合が不確かな区間においても, 2 台のカメラによる照合結果を統合的に利用することで, 安定して位置推定が可能となったためである。図 5.10 に統合処理の有無による入力映像とデータベース映像の対応付け結果の違いの例を示す。図 5.10(a),

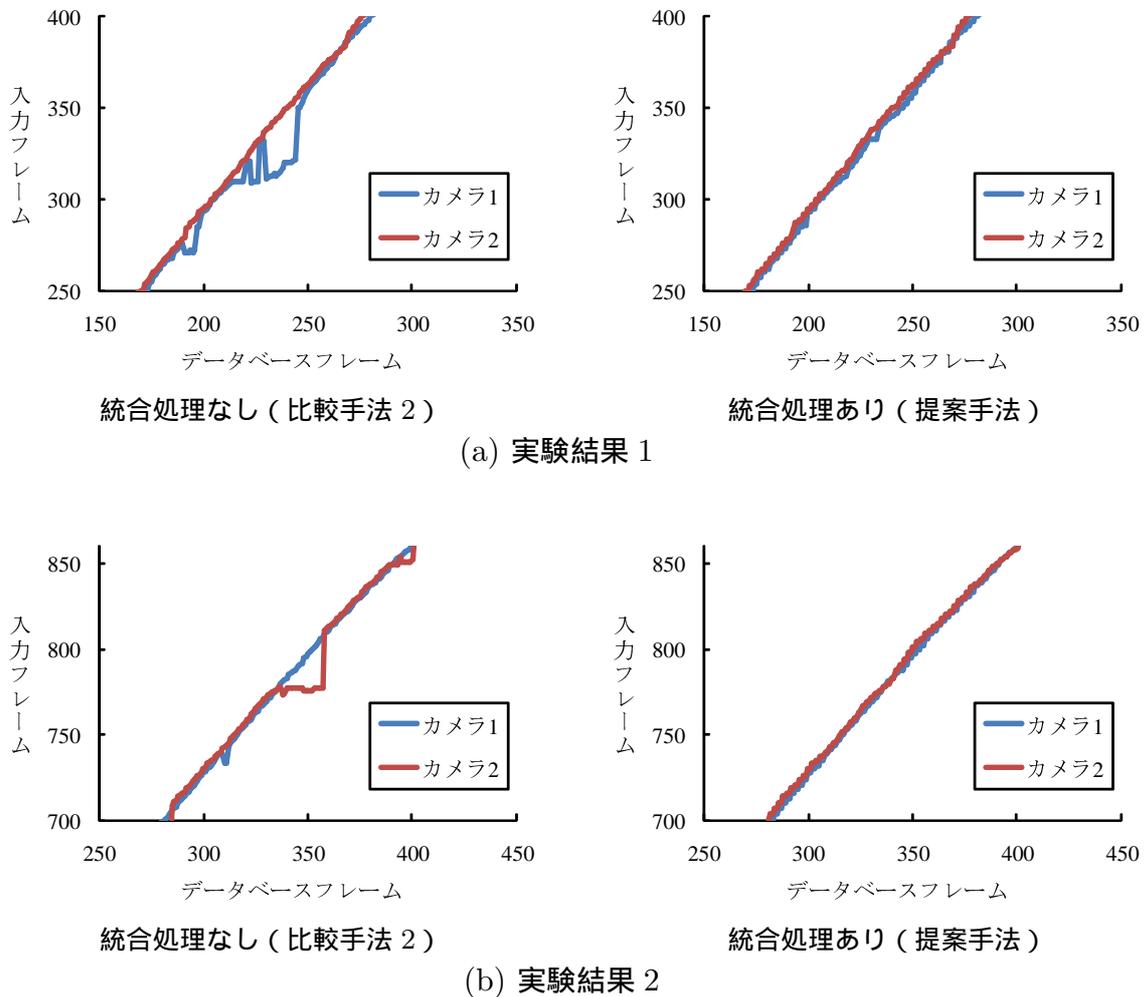


図 5.10 統合処理の有無による入力映像とデータベース映像の対応付け結果の違い。

(b) はそれぞれ異なる走行の実験結果である。左の列が統合処理なし（比較手法 2）の結果であり右の列が統合処理あり（提案手法）の結果である。グラフ中のパスは入力映像とデータベース映像の対応関係を表している。統合処理なしでは画像間の対応付けが失敗している区間でも、統合処理ありでは 5.2.4 の対応付けの制限に従うように、正しい対応付けに修正されていることが確認できる。しかしながら、カメラ 1 と 2 の両方で画像照合に失敗している場合、統合処理後も正しい対応付けに修正ができない事例が見られた。そのため、今後より多くのカメラを導入する等、安定性をより向上させる手法の開発が必要である。

また、カメラ 2 台を用いた提案手法と比較手法 2 は、カメラ 1 台を用いた比較手法 1 よ

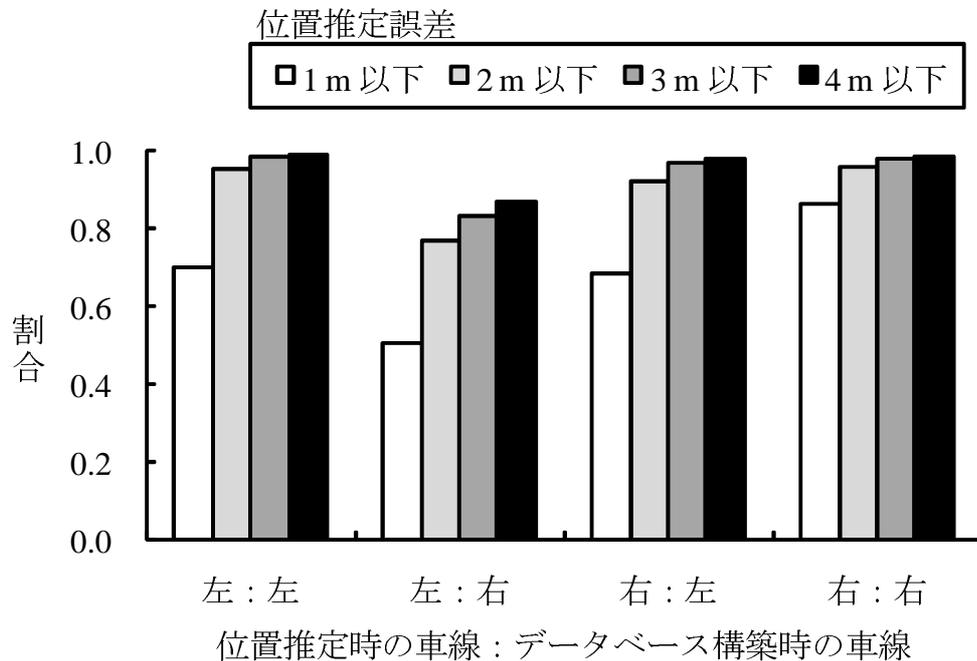


図 5.11 進行方向の位置推定誤差が一定範囲に収まる割合（車線別）

りも進行方向の位置推定誤差が小さい。これは、比較手法 1 では類似するフレームを探索する手法のため、自車位置推定時の車線方向の位置がデータベース構築時とずれていた場合、進行方向にも位置ずれが生じるためである。提案手法と比較手法では、2 台のカメラによって三角測量を行うため、進行方向の位置推定誤差は抑制される。

提案手法の進行方向の位置推定誤差と走行車線認識率の車線ごとの内訳を、それぞれ図 5.11, 5.12 に示す。位置推定時の車線とデータベース構築時の車線が異なる場合、進行方向の位置推定誤差と走行車線認識率のいずれも、精度の低下が見られた。これは、提案手法では位置の違いによる見えの違いを画像の平行移動と拡大縮小で近似しているが、車線方向の位置の違いが大きくなるほど歪みが大きくなるためである。そのため、精度をより向上するためには、画像間の照合手法について、改良の余地がある。

5.4.2 車線方向の位置推定誤差に関する考察

5.3 の実験では、映像の取得時の位置が車線内で不明なため、車線方向の位置推定精度を走行車線認識率で評価した。本節では、データベース構築用と位置推定用の実験映像を

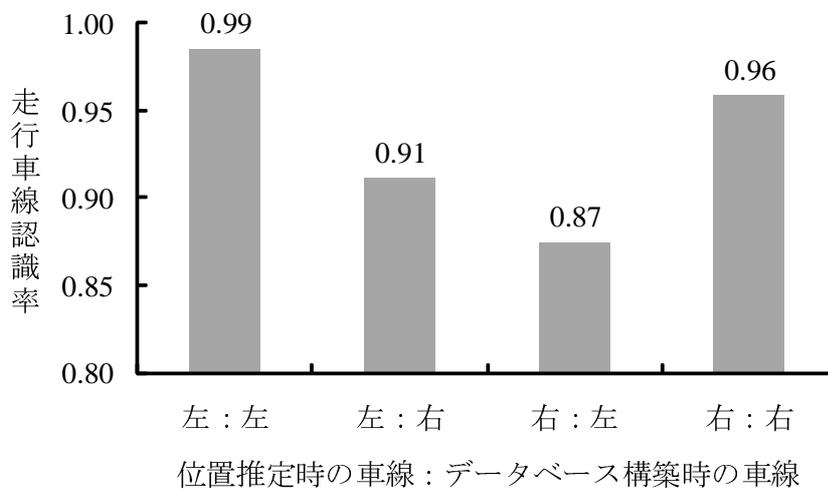


図 5.12 走行車線認識率 (車線別).

同じ走行で撮影することで、データベース構築時と位置推定時に同じ車線を走行した場合の車線方向の位置推定誤差を評価した実験について述べる。データベース構築時と位置推定時で全く同じ条件となるため、現実的な状況ではないものの、理想的な状況時での車線方向の位置推定誤差の参考値として評価することができる。

実験用の映像は、1つの実験車両にデータベース構築用の全方位カメラと位置推定用の2台の通常カメラを同時に装着して撮影した。実験経路や実験パラメータ等は5.3の実験と同じであり、合計4回実験を行った。

車線方向の位置推定誤差が一定範囲に収まる割合を車線別に図5.13に示す。車線方向の位置推定誤差は大部分の走行区間で1.0m以下に収まった。左車線と右車線では、大きな差は見られなかった。

車線方向の位置推定誤差が1.5m以下の割合が左車線走行時で0.99、右車線走行時で0.98であった。一方、5.3の実験では、データベース構築時に車両は走行車線の中心を走行していると仮定し、車線方向の推定位置が $-1.5 \sim 1.5$ mの場合はデータベース構築時と位置推定時で同じ走行車線とした。この実験では、データベース構築時と位置推定時で両方とも左車線を走行した場合には走行車線認識率が0.99、右車線を走行した場合には0.96であった(図5.12)。右車線走行時には多少の差が生じたものの、本節と5.3の実験はおおよそ同じ値であった。

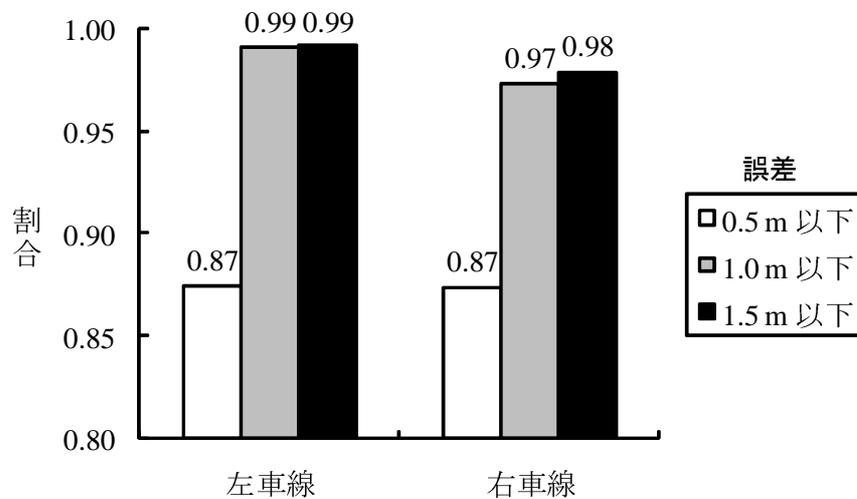


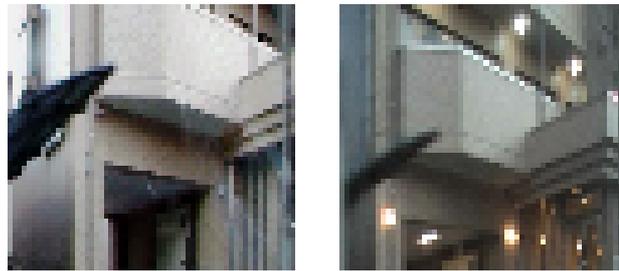
図 5.13 車線方向の位置推定誤差が一定範囲に収まる割合。

5.4.3 時間帯による影響

本手法は可視光カメラを用いるため、走行する時間帯によっては、照度の不足により本手法の適用が困難になる場合が考えられる。そこで、日没直後に測定した自車位置推定用の映像を用い、時間帯の違いによる影響を調査した。

本実験では、1回走行して位置推定を行った。なお、データベース用の映像は日中に撮影したものを使用した。実験の結果、進行方向の位置推定誤差が2.0m以下に収まった割合が95%であり、日没直後の時間帯でも提案手法が適用できることを確認した。これは、周囲が薄暗いため画像のコントラストが低下するが、建物の形状は画像上で確認できるためであると考えられる。図 5.14 に日中と日没直後の同じ建物の画像の例を示す。灯火によって画素値が飽和する部分が存在するなど、日中とは見た目が変化する構造物があったものの、変化する部分の面積は小さく、自車位置推定に影響は見られなかった。

本実験では、日没直後の薄暗い環境で撮影した映像と日中に撮影した映像の照合を行ったが、深夜に撮影した映像と日中に撮影した映像では、より照合が困難になると考えられる。このため、深夜に自車位置推定を行う場合、日中に撮影したデータベースとは別に、深夜に撮影したデータベースも用意する等の対策が必要である。



(a) 日中

(b) 日没直後

図 5.14 日中と日没直後の市街地の変化 .

5.5 まとめ

自車位置推定時に 2 台の車載カメラを用いることで、車両の進行方向の位置だけでなく、車線方向の位置も推定する手法を提案した。第 4 章のアピアランススペースの位置推定手法では、車線方向の位置を推定することは困難であったが、2 台目のカメラを導入することで解決した。具体的には、2 台の車載カメラから得られる入力画像に対し、類似する画像が得られる 2 地点をデータベース上から探索し、三角測量による走行位置推定手法を行った。また、位置推定精度を向上させるため、映像は経路に沿って連続に存在することを利用し、複数フレームを利用することで、2 カメラの画像照合結果を統合する処理を行った。その際、道路を走行する車両の典型的な挙動から、2 つの入力映像とデータベース映像の対応関係に制約を付加した。

実験により、2 台の車載カメラを用いることで、車線方向の位置も推定可能であることを確認した。また、2 台のカメラを用いた手法は、1 台のカメラを用いた手法より進行方向の位置推定精度が向上した。さらに、2 台のカメラの照合結果を統合したほうが、2 台のカメラをそれぞれ独立に扱うよりも高精度に自車位置推定が可能であることを確認した。

今後の課題として、より多くの車載カメラを用いた手法の開発や、時刻・天候の影響に関するより詳細な調査が挙げられる。

第6章

むすび

6.1 総括

高精度な自転車位置推定を実現するためには、それに適した市街地映像データベースの構築が必要であるため、高品質な市街地映像データベースの構築と、自転車位置推定手法の開発の両面から研究を行った。高精度な運転者支援を実現するためには、1 m 程度の位置推定精度が必要である。自転車位置推定手法として、全方位市街地映像データベースと車載通常カメラ映像の照合による手法を提案し、実験によって運転者支援のために十分な精度を得られることを確認した。高品質な市街地映像データベースとして、映像中に移動物体が存在しないデータベースの構築を行った。映像中に移動物体が存在しないデータベースを自転車位置推定に用いることによって、市街地映像データベース中の背景の遮蔽の問題を解決し、自転車位置推定性能が向上することを確認した。また、移動物体を除去することで、市街地映像データベースの利用の障害となるプライバシーの問題を解決した。これらにより、高精度な自転車位置推定および自転車位置推定に適する高品質な市街地映像データベースの構築を行うという本研究の趣旨が達成できたと考える。

第2章では、市街地映像データベースと自転車位置推定に関する周辺技術を紹介した。特に、市街地映像データベース中の移動物体の問題と、カメラによる自転車位置推定については詳しい説明を加えた。

第3章では、映像中に移動物体が存在しないデータベースの構築について説明した。映像中の背景の遮蔽やプライバシーの問題を解決するために、市街地映像中の移動物体を除去する手法を提案した。従来、映像中の不要な物体を除去する手法が提案されているものの、市街地映像データベースのような移動カメラ映像への適用には適していなかった。そ

ここで本研究では，同一経路を複数回走行して得られた映像を統合することで，市街地映像データベース中の移動物体を効率的に除去する手法を提案した．実験によって，歩行者や路上の車両を除去できることを確認した．

第 4 章では，全方位市街地映像データベースと車載通常カメラ映像の照合による自転車位置推定手法を提案した．従来のアピアランススペースの自転車位置推定は，市街地映像データベース構築時と位置推定時で同一種類のカメラを用いた手法は提案されていたが，異種カメラを用いた自転車位置推定にこれらの手法の枠組みを適用することは困難だった．そこで，映像間の時空間位置合わせによって，カメラの視野角が異なっても，高精度に映像間の照合が可能な手法を提案した．さらに，評価実験により，第 3 章で提案した移動物体除去手法を適用した市街地映像データベースを用いることによって，位置推定精度が向上することを確認した．これによって，本研究の意義の 1 つである市街地映像データベースを自転車位置推定に適するように加工・編集することの有用性を確認した．

第 5 章では，2 台の車載カメラを用いた車線方向を含めた走行位置推定手法を提案した．第 4 章の自転車位置推定手法は進行方向の位置情報のみ取得可能であり，車線方向の位置情報は取得することができなかった．そこで，自転車位置推定時に 2 台の車載カメラを用い，三角測量を行うことで，車線方向の位置情報も取得可能となるように拡張を行った．また，2 台のカメラの照合結果を統合的に利用し，2 台のカメラをそれぞれ独立に扱うよりも高精度に自転車位置推定を行う手法を提案した．

さらに，市街地映像データベースの応用として，市街地構造物の平面に画像情報を提示する手法を，付録 A に収録した．

6.2 今後の課題と展望

第 3 章では，複数映像統合による移動物体除去手法について述べた．ここで提案した手法では，同一経路の大量の映像が必要であるため，映像の収集方法について今後検討する必要がある．大量の映像を収集する方法の 1 つに，カメラを搭載した複数の車両が協調するという枠組みが考えられる．一方，少ない映像数で高い移動物体除去性能が得られる手法の開発を行うという方向性もある．例えば，移動物体らしさの計算の際に，映像の連続性を活かし，時間方向・空間方向の周辺の画像領域を使用する方法が考えられる．

第 4 章では，アピアランススペースの手法による全方位カメラ映像と通常カメラ映像の

照合による自車位置推定について述べた。提案する自車位置推定手法において、第3章で提案した移動物体除去手法を適用した市街地映像データベースを用いることで、データベース中の遮蔽の問題が解決可能であることを確認した。しかしながら、自車位置推定時の車載カメラ映像中に移動物体が存在している場合、位置推定精度が低下する。そのため、自車位置推定時の車載カメラ映像中の移動物体による遮蔽に対処する手法の検討も必要である。また、本論文では映像間照合に焦点を当てたが、実際の多くのカーナビゲーションシステムではGPSと自律航法を組み合わせている。そのため今後、このような複数種類のセンサを組み合わせた手法も検討していきたい。

第5章では、自車位置推定時に2台の車載カメラを用いる手法を提案した。2台のカメラを用いることで、1台の場合よりも位置推定精度が向上することを確認したが、今後より多くのカメラを導入することで、さらなる精度向上を目指すことが今後の課題として挙げられる。また、日没直後に実験を行うことで時間帯による影響の評価を行ったが、より環境の変化に頑健な手法を目指すためには、時刻・天候の影響のより詳細な調査を行う必要がある。自車位置推定は一般的に可視光カメラが用いられるが、赤外線カメラやマルチバンドカメラを用いることによって、夜間に対応する方法も考えられる。ナイトビュー(TOYOTA)やインテリジェント・ナイトビジョンシステム(HONDA)等、赤外線カメラを搭載した市販車両が発表されており、今後の普及が期待できる。

付録Aでは、市街地画像データベース中の構造物の平面に画像情報を提示する手法を述べた。ユーザにより違和感なく情報を提示するためには、重畳される画像の状態をより現実に近づけることが考えられる。例えば、光源特性の再現や、提示画像の前に前景物が存在した場合、提示画像の前景物による遮蔽を再現することが挙げられる。また、ユーザに適切な情報を提示するためには、地図情報と3次元的な提示位置の情報を対応付ける必要がある。地図情報と提示位置が対応付いていれば、例えばナビゲーションの用途において、現在地に応じた提示画像の選択が可能となる。今後、このような対応付け手法のより詳しい検討が必要である。

また、本論文では移動物体のない市街地映像データベースを構築したが、市街地映像データベースにより多様な情報を格納していくことも検討課題として挙げられる。このような情報として、映像フレームの画像照合の信頼度や市街地中の物体の属性、3次元構造などが挙げられる。多様な情報を格納することで、新たな自車位置推定手法の開発も可能になると期待できる。例えば、自車位置推定を行う際、画像照合を行う際の信頼度をデー

データベース中の画像に付加することができれば，信頼度の低い画像は位置推定に利用しない等の処理が可能になる．

市街地映像データベースは人間の「風景の記憶」に相当する．人間は過去に様々な場所を通過し，その地点で見た風景を記憶しており，これらの風景の記憶を様々な活動に利用している．人間の認識能力は非常に高く，風景の内容を容易に理解することができるが，現在の計算機では風景画像を解析し，理解することは容易ではない．一方，計算機のほうが人間よりも優れた点が存在する．例えば，人間は細部を正確に記憶することは困難であるが，カメラ画像は高解像度撮影によって細部の記録も容易である．また，市街地映像データベースは人間では困難な量のデータを，収集・記録することができる．そのため今後，人間に匹敵する市街地映像の解析性能を目指すとともに，このような市街地映像データベースの利点を活かすことによって，環境の変化に柔軟に適応可能かつ正確な自転車位置推定手法を目指したい．

最後に，本論文の成果が自転車位置推定および市街地映像を利用した研究の発展に貢献することを願って，本論文のむすびとする．

謝辞

本論文で述べた研究は、筆者が名古屋大学大学院情報科学研究科博士前期課程ならびに後期課程在学中に行ったものである。

まず最初に、博士前期課程・後期課程の5年間にご指導とご鞭撻を頂いた本学村瀬洋教授に深く感謝いたします。指導に当たっては、時に厳しく、時にやさしく接していただきました。この研究を博士論文として形にすることができたのは、先生あってのことと心から感謝します。本学井手一郎准教授に深く感謝いたします。研究に関する深い議論から些細なことまで幅広く助言を頂きました。特に、論文執筆時の注意点など、事細かに指導していただき、感謝いたします。本学出口大輔助教に深く感謝いたします。先生には、なかなか研究の進まない筆者を暖かく見守っていただき、事細かに指導していただきました。心から感謝いたします。また、岐阜聖徳学園大学高橋友和准教授に感謝いたします。昼夜を問わず熱心に研究の指導をしていただきました。深くお礼申し上げます。深い助言を頂いた中京大学目加田慶人教授に感謝いたします。先生には、未熟な筆者を温かく見守っていただきました。

本論文に目を通され、ご指導とご鞭撻をいただいた本学武田一哉教授、本学村瀬洋教授、本学井手一郎准教授に感謝いたします。多数の有用なご意見をいただき、ありがとうございました。特に、武田一哉教授には論文内容の不明瞭な点のみならず、細かい表現や言い回しまでご指摘していただき、感謝いたします。

本研究の一部を進めるにあたり、有用なご意見をいただきました日本電信電話株式会社コミュニケーション科学基礎研究所柏野邦夫博士ならびに川西隆仁博士に深く感謝いたします。共同研究という貴重な機会を提供していただき、ありがとうございました。

筆者が名古屋工業大学工学部在籍中に卒業研究の指導をしていただいた名古屋工業大学梅崎太造教授に感謝いたします。画像処理という分野に携わるきっかけを与えていただき、画像処理の面白さを教えていただきました。

未熟な筆者を励ましてくれた多くの友人や先輩，そして後輩に感謝します．特に，筆者が博士前期課程・後期課程在学中，村瀬研究室で学んだ野田雅文氏，道満恵介氏に感謝いたします．共に切磋琢磨することで，たいへん励みになりました．また，博士後期課程の先輩としてお手本を示していただいた石田皓之博士，Lina 博士，神谷保徳博士に感謝いたします．先輩方がいらっしゃらなかったら，筆者が博士後期課程に進むことはありませんでした．

また，研究を進めるにあたり，ご支援を頂きながら，ここにお名前を記すことができなかった多くの方々に心より感謝いたします．

最後に，私生活の面から 27 年間もの長い間，筆者を支えていただいた家族に感謝いたします．

付録 A

市街地映像データベースの拡張現実型画像情報提示への応用

市街地映像データベースは、画像中の不要な物体を除去するだけでなく、ユーザにとって有用な情報を新たに市街地画像中に付加することもできる。このような画像情報の重畳は、ナビゲーションのための道案内情報提示や広告提示等への応用が可能である。付録 A では、市街地映像データベース構築の応用として、市街地映像データベース中の建造物の壁面に画像情報を重畳表示する手法を示す(図 A.1)。

ユーザに正確かつ違和感なく情報を提示するためには、提示位置が正確である必要がある。そこで、市街地構造の 3 次元復元を行い、重畳対象である壁面を自動抽出し、そこに画像情報の重畳を行う。

拡張現実 (AR) を利用し、市街地画像中に文字や画像情報を提示する研究やサービスが行われている。例えば、Snavely らは複数の視点で撮影された投稿写真を対応付け、様々な視点から撮影対象物を閲覧可能な手法を提案している [26]。また、Google ストリートビュー [5] 等、インターネット上で市街地画像が閲覧できるサービスが提供されている。これらの手法やサービスでは、空間中に文字情報が重畳可能である。その他、車載カメラ映像に道案内情報を重畳表示するカーナビゲーションシステムが発表されている [89][90]。

AR を利用した広告提示のサービスの提供や研究も行われている。例えば、スポーツ映像中の背景物体上に広告を重畳表示する手法が提案されている [91][92]。セカイカメラ [93] は、文字・画像情報を空間中に重畳表示するサービスを提供しており、カメラ画像への物件情報の重畳が行われている。しかし、セカイカメラは GPS で位置推定を行ってい

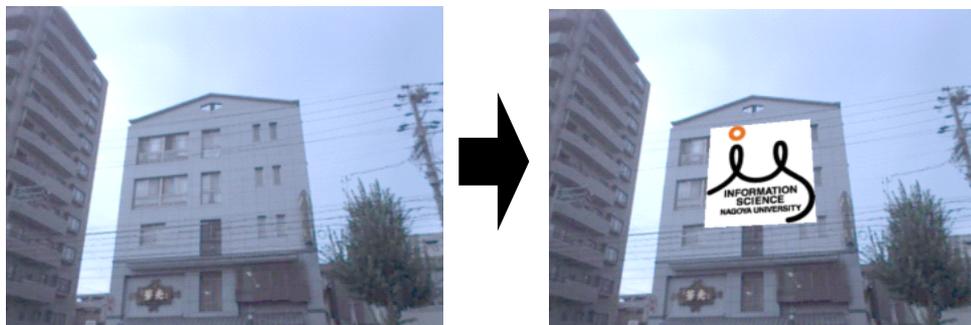


図 A.1 市街地映像データベースへの画像情報提示（本手法の適用例）.

るため、正確に対象物体に情報を重畳させることが難しい。

提示したい情報を正しく市街地画像に重畳するためには、現実環境と提示位置の位置関係を知る必要がある。現実環境と提示位置の位置関係を推定する方法として、主に以下の3つがある。

- (a) GPS などによる位置情報から提示位置を計算する手法 [92][93]
- (b) 提示位置を示すマーカを利用する手法 [91][92][94]
- (c) マーカを使用せず画像から適切な重畳対象物体を検出し、それに重畳する手法 [95]

(a) のように GPS 等を利用する場合には、位置推定精度が低いため、正しい位置に正確に重畳を行うことが難しいという問題がある。隅田らは GPS の位置推定誤差と情報提示の正確性の関係性をシミュレーションにより評価しており、GPS の誤差が建物幅の半分以上の場合、目標とする建物への情報提示が困難になると報告している [96]。また、GPS 等を利用する場合には、重畳の際に周囲の建築物等の 3 次元構造との整合性を考慮しておらず、自然な見た目となるように提示画像と市街地画像を合成することが難しい。(b) のようなマーカを利用する手法では、市街地のような広範囲での画像情報提示を目的とした場合、提示位置を手で設置することは現実的ではない。(c) のマーカを使用せず画像から適切な重畳対象物体を検出する手法として、Lee らは掌を検出し、その上に情報を重畳する手法を提案している [95]。このようなマーカを使用しない方針は、市街地映像データベースへの画像情報提示に適すると考えられる。そこで本研究では、(c) の方針を採用する。本手法では、マーカを設置する代わりに、周囲環境の 3 次元復元を行い、適切な提示位置を 3 次元構造から自動選択を行う。

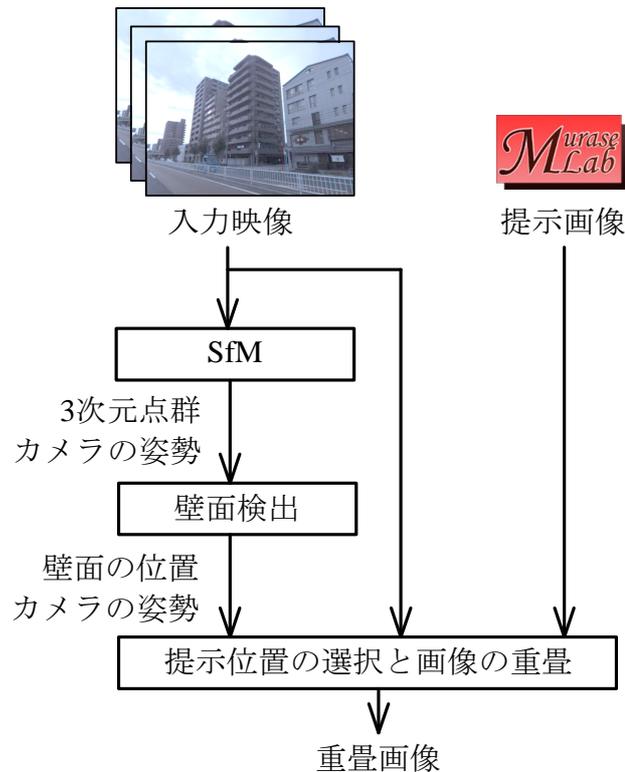


図 A.2 拡張現実型情報提示手法の流れ.

A.1 拡張現実型画像情報提示手法

本研究で示す市街地構造物の平面への重畳表示手法は、(1) Structure from Motion (SfM) による 3 次元点群の取得、(2) 3 次元点群からの壁面の検出、(3) 提示位置選択、(4) 提示画像の重畳の 4 つのステップに分けられる。本手法の流れを図 A.2 に示す。以降、本手法の詳細を説明する。

A.1.1 SfM による 3 次元点群の取得

まず、画像中の特徴点を Harris オペレータ [97] により検出する。次に、映像の連続するフレームに対して、KLT トラッカ [98] により特徴点を追跡する。最後に、追跡した特徴点から、SfM [99] により特徴点の 3 次元位置とカメラの姿勢を推定する。

事後処理として、以降で行う壁面検出に不要な点群の除去を行う。まず、路上の柵等の

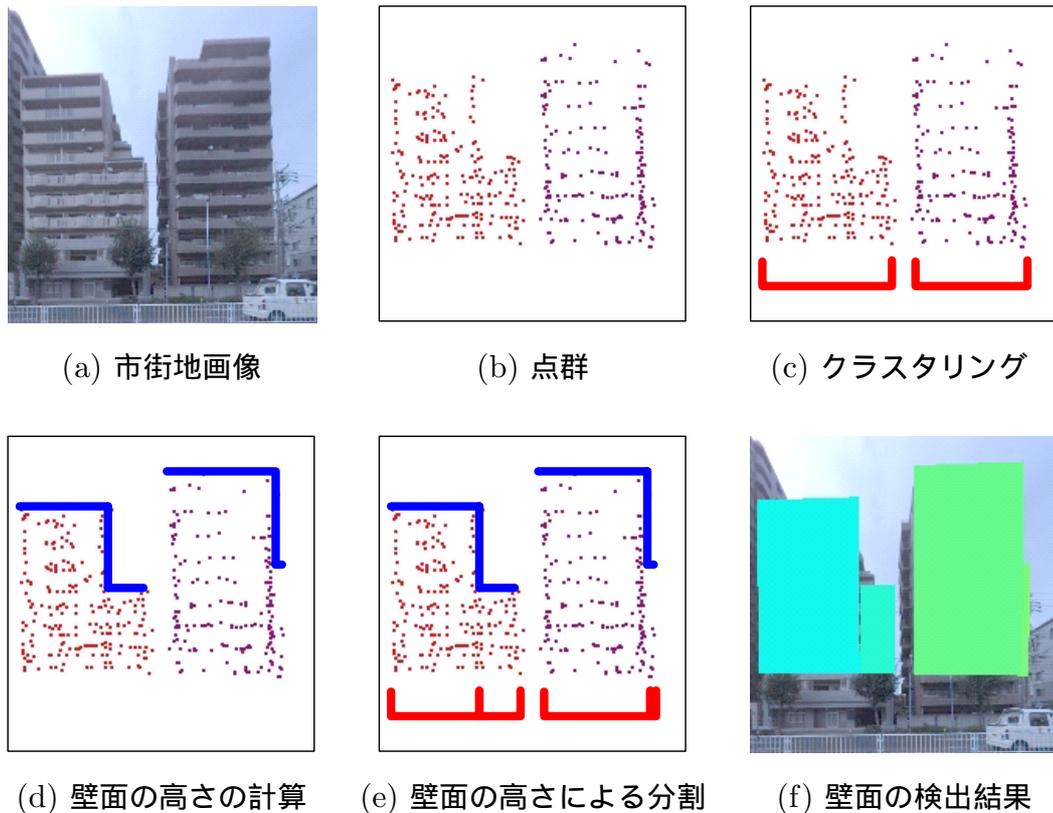


図 A.3 壁面の検出の模式図．

点群を除去するため、しきい値以下の高さの点群を除去する．また、SfM で誤って検出された点群を除去するために、孤立点の除去を行う．孤立点の除去では、ある点から一定の範囲内に存在する点がしきい値以下場合、その点を除去する．

A.1.2 壁面の検出

3次元点群から広告提示の対象である壁面群の検出を行う．壁面の検出は、(a) 点群のクラスタリング、(b) Hough 変換による平面パラメータの算出、(c) 壁面の高さの計算による矩形領域への分割の3つの処理に分けられる．以下、この処理の詳細を説明する．

(a) 点群のクラスタリング

まず、図 A.3(b) のような点群を、(c) のように異なる平面ごとに分離するため、3次元点群のクラスタリングを行う．クラスタリングは以下で説明する逐次的手法により行う．クラスタリングを行うために、初期条件として、すべての点に対し1つのクラスタを割

り当てる．そして，変化がなくなるまで，以下のような規則で各クラスを逐次的に併合していく．点群 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_N$ で構成されるクラス \mathbf{a} と，点群 $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_M$ で構成されるクラス \mathbf{b} を併合する場合を考える．任意に選択した $\mathbf{a}_i, \mathbf{b}_j$ が次の条件を満たせば，クラス \mathbf{a}, \mathbf{b} を併合する．

$$|(\mathbf{a}_i)_x - (\mathbf{b}_j)_x| < \theta_w \quad (\text{A.1})$$

$$|(\mathbf{a}_i)_y - (\mathbf{b}_j)_y| < \theta_h \quad (\text{A.2})$$

$$|(\mathbf{a}_i)_z - (\mathbf{b}_j)_z| < \theta_d \quad (\text{A.3})$$

ここで， $(\mathbf{a}_i)_x$ は点 \mathbf{a}_i の x 座標を示しており， $\theta_w, \theta_h, \theta_d$ はそれぞれしきい値である． x, y, z はそれぞれ道路方向，鉛直方向，車線方向であるとする．市街地では壁面は道路に面している場合が多いため，このような壁面を検出したい場合には， θ_w や θ_h は θ_d に比べて大きな値を設定する．

(b) Hough 変換による平面パラメータの算出

各クラスに含まれる特徴点群に対し，3次元空間における平面のパラメータを求める．3次元空間上の平面を構成する点群の配置には雑音や歪みが存在するため，Hough 変換 [100] を用いることで，投票により平面パラメータを決定する．

3次元空間における平面は，極座標表現 (ρ, θ, ϕ) を用いて次式で与えられる．

$$\rho = x \sin \theta \cos \phi + y \sin \theta \sin \phi + z \cos \theta \quad (\text{A.4})$$

ただし， ρ は原点と原点から平面への最近点との距離， ϕ は x - y 平面上の x 軸とのなす角， θ は z 軸とのなす角である (図 A.4(a))．3次元空間中の平面のパラメータを求めるためには，図 A.4(b) のように，クラスに含まれる点 (x, y, z) に対応する ρ - θ - ϕ 空間上の曲面を求める．クラスに含まれる全ての点による曲面が交わる点が平面のパラメータとなる．

この曲面が交わる点は， ρ - θ - ϕ 空間に頻度テーブルを用意し，投票によって決定する．具体的には，クラスに含まれるある点 (x, y, z) に対応する ρ - θ - ϕ 空間上の曲面を求め，これに対応する頻度テーブルの要素に投票を行う．クラスに含まれる全ての点による投票の結果，最も頻度の高い頻度テーブルの要素に対応する (ρ, θ, ϕ) が平面のパラメータとなる．

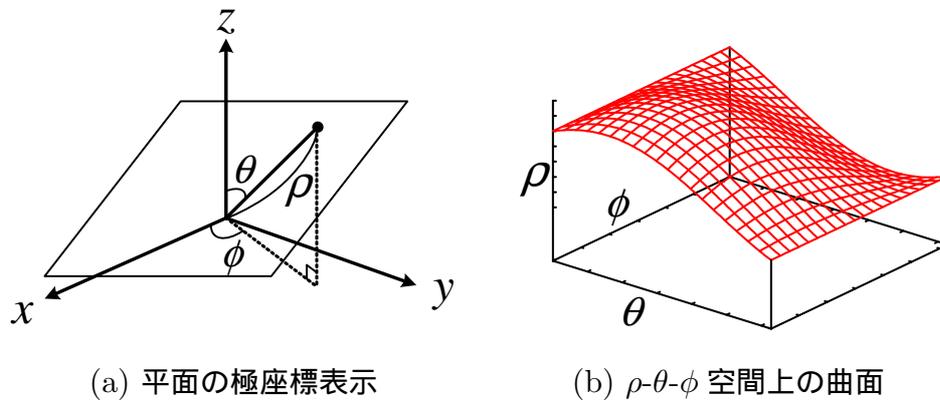


図 A.4 Hough 変換による平面パラメータの算出 .

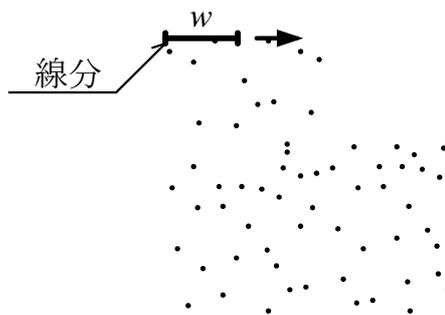


図 A.5 壁面の高さの計算 .

(c) 壁面の高さの計算による矩形領域への分割

検出された平面を矩形領域に分割するために，壁面の高さ情報を用いる．まず，図 A.3(d) のように，平面に投影した点群から壁面の高さを求める．次に，壁面の高さが大きく変化する部分を検出することで，図 A.3(e) のように (c) のクラスタリング結果をさらに分割する．

具体的には，図 A.5 のように水平で長さが w であり，点群の左端で上部に接する線分を用意する．そして，点群の上部に接するように，この線分を左から右に走査することによって，壁面の高さを求める．次に，壁面の高さの変化がしきい値以上の位置を探索し，この位置で矩形に分割する．

最後に，分割された点群を囲む矩形を求めることで，矩形の範囲を決定する(図 A.3(f))．

A.1.3 提示位置の選択と提示画像の重畳

推定した平面中で，提示画像の表示に適する領域を選択する．選択の基準として，広告提示の対象とする構造物であることや，面積が十分大きいこと，ユーザが視認するために適切な高さであること等を用いる．

選択した提示位置に提示画像を重畳表示する．提示画像として，次の 2 つの種類がある．

- 平面図形
- 立体図形

平面図形を重畳する場合，提示対象の壁面に平行となるように，2次元の提示画像を壁面上に設置する．また，立体図形を重畳する場合，3次元 CG で生成した広告物を提示対象の壁面上に設置する．立体図形には，袖看板のようなものがある．平面図形の重畳を行う際，SfM および壁面の検出で求めた提示画像の形状や提示位置・方向，カメラの姿勢のパラメータに基づいて，提示画像を射影変換する．立体図形も同様に，これらのパラメータを基に生成した CG 画像を市街地画像に重畳する．

A.2 画像情報提示実験

車載全方位カメラで撮影した市街地映像に，提示画像を重畳表示する実験を行った．経路長は約 300m であり，道路面および対向車線側の建築物の壁面に提示画像の重畳を行った．SfM の際，全方位カメラを透視投影変換した画像を用いた．また，SfM は，3次元復元の誤差を映像全体で最小化するバンドル調整 [101] も行った．

図 A.6 に本手法の処理過程の図を示す．1 段目が入力画像，2 段目が復元した 3次元点群，3 段目が矩形領域の検出結果，4 段目が広告画像を重畳した市街地画像である．また，図 A.7 に広告画像を壁面に平行となるように重畳した例を示す．これは，本手法により検出した壁面の中で，一定の大きさ以上を持つ全ての矩形領域に，無作為に選んだ広告画像を重畳したものである．図 A.8 は立体的な袖看板を壁面に垂直となるように重畳した例である．図 A.7，A.8 より，市街地構造物に良好に提示画像を重畳できたことがわかる．

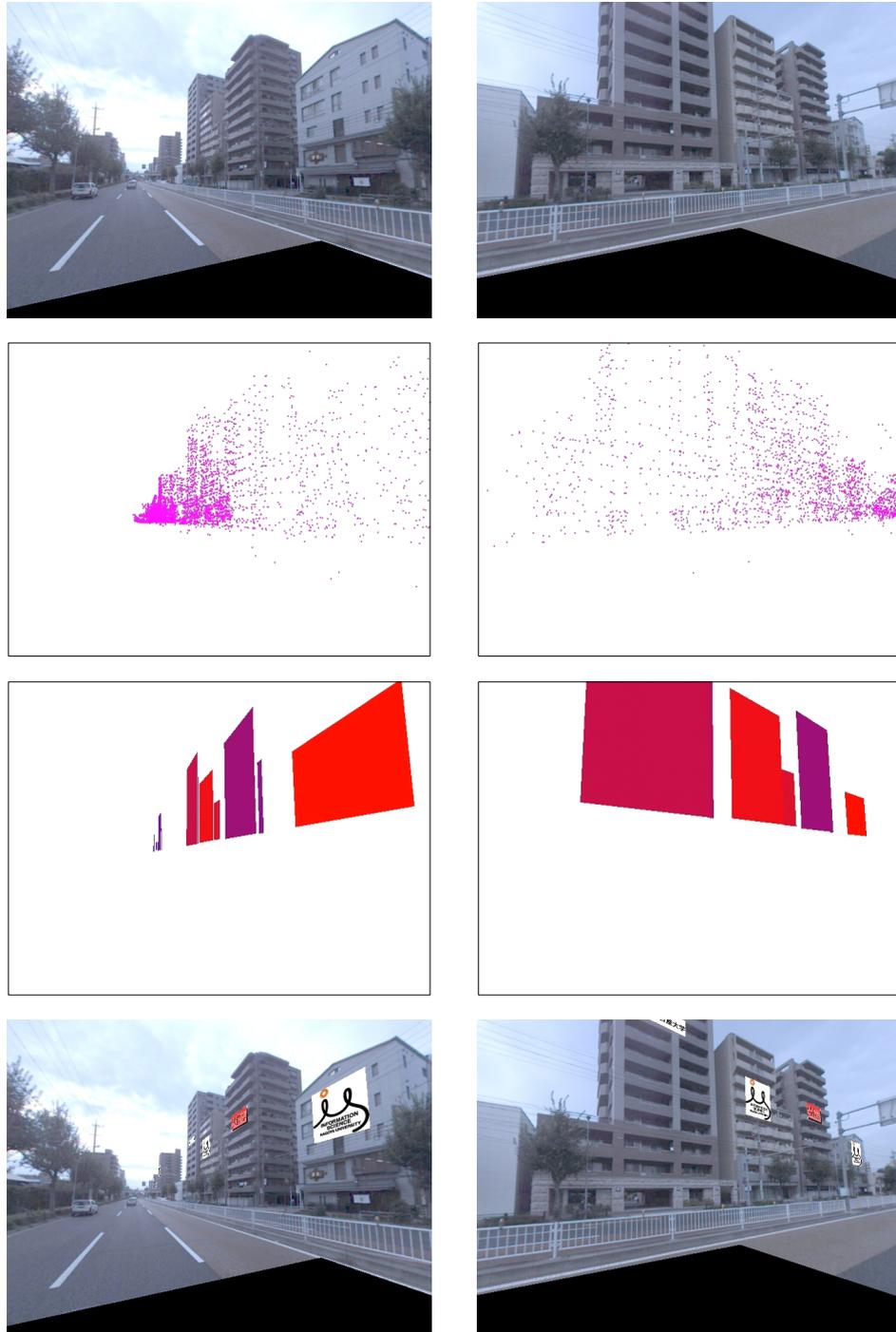


図 A.6 本手法の処理過程．1 段目：入力画像，2 段目：復元した 3 次元点群，3 段目：矩形領域の検出結果，4 段目：広告画像を重畳した市街地画像．



図 A.7 広告画像を壁面に平行に重畳した画像 .

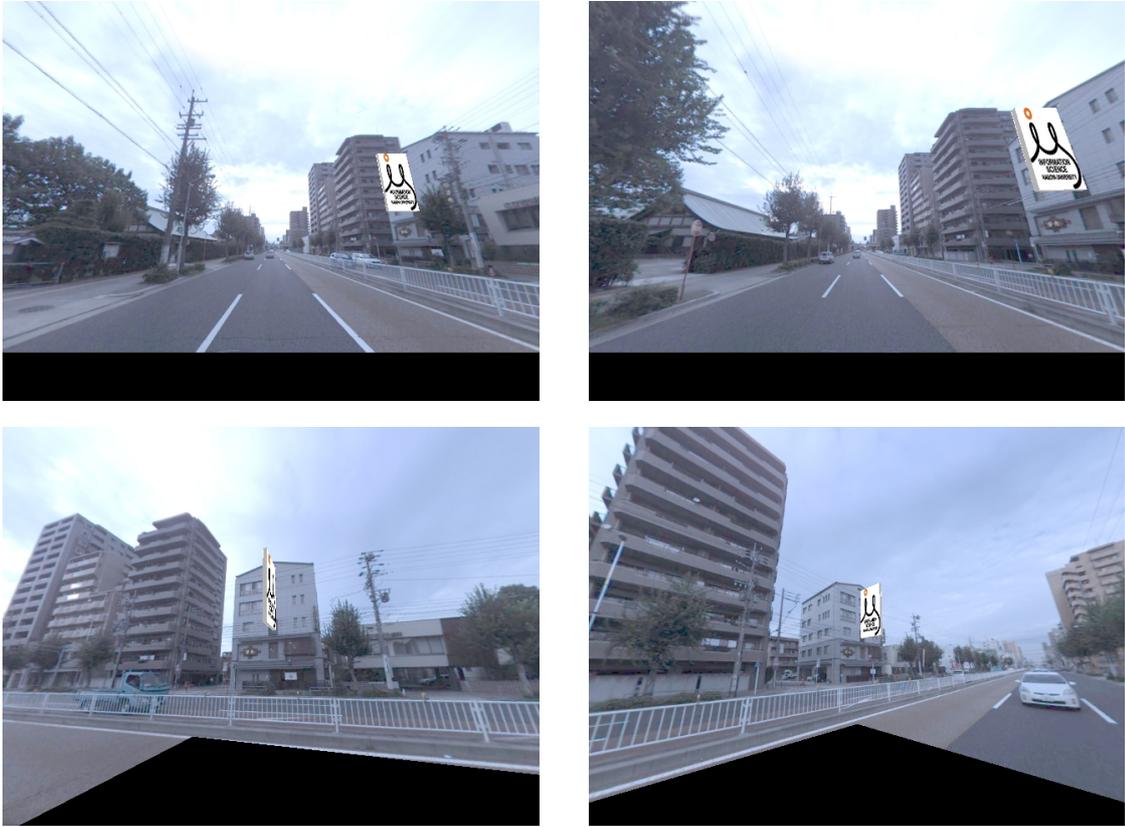


図 A.8 広告画像を壁面に垂直に重畳した画像 .

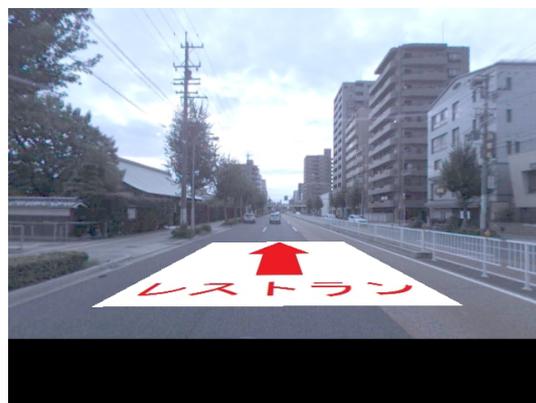


図 A.9 広告画像を路面に重畳した画像 .

本手法は建築物の壁面だけでなく，路面にも広告画像を重畳可能である．図 A.9 に道案内情報を路面に重畳した例を示す．

処理時間は，1 フレームあたり 1 秒程度であった．処理時間の大部分は SfM であり，事前に SfM を行うことができるインターネット上の市街地画像閲覧サービスのような用途には十分利用可能である．しかしながら，リアルタイム処理を必要とする用途では，SfM の高速化について検討の余地がある．

A.3 まとめ

市街地映像データベースの応用として，ナビゲーションのための道案内情報提示や広告提示等を目的とし，映像中の市街地構造物の平面に画像情報を提示する手法を示した．正確に提示対象の平面に重畳するために，SfM により壁面を構成する点群を復元し，点群から平面の検出を行った．そして，検出した平面に画像情報の重畳を行った．実験により，市街地映像データベース中に画像情報が提示可能であることを確認した．今後の課題として，環境中の光源特性の再現等，より現実に近い画像の重畳手法の開発や，視認しやすい画像情報提示位置・提示方法の検討が挙げられる．

付録 B

動的計画法による最適化

「軌道問題」の動的計画法による解法の導出する．以下の結果は，第 4 章の拡張 DP マッチングの導出および第 5 章の画像照合結果の統合処理に使用される．まず，図 B.1 に示されるような平面回路網で，常に上方向に動き， $i = 1$ のエッジ上のある点から $i = I$ のエッジ上のある点に動く間に通過する数 $d_i(j)$ の和を最小にする経路を探索する問題を考える．これは次式を最小化する u_1, u_2, \dots, u_I を探索する問題に帰着する．

$$F = \sum_{i=1}^I d_i(u_i) \quad (\text{B.1})$$

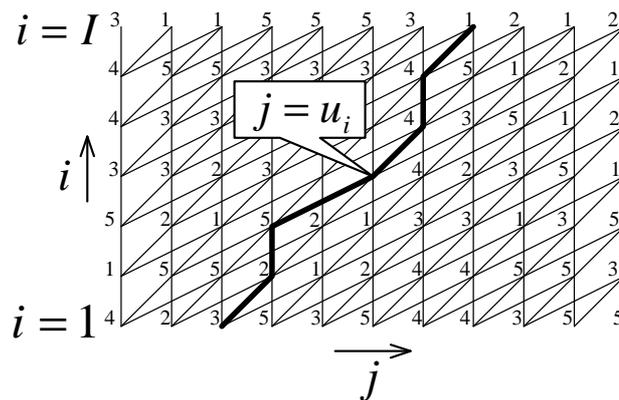


図 B.1 平面回路網

この最適化問題を以下のようにして解く .

$$\begin{aligned} \min F &= \min_{u_1, \dots, u_I} \sum_{i=1}^I d_i(u_i) \\ &= \min_{u_2, \dots, u_I} \left[\sum_{i=3}^I d_i(u_i) + \left\{ d_2(u_2) + \min_{u_1} d_1(u_1) \right\} \right] \end{aligned} \quad (\text{B.2})$$

ここで , 右辺の $\{ \}$ 内を

$$g_2(u_2) = d_2(u_2) + \min_{u_1} d_1(u_1) \quad (\text{B.3})$$

と置くと ,

$$\begin{aligned} \min F &= \min_{u_2, \dots, u_I} \left[\sum_{i=3}^I d_i(u_i) + g_2(u_2) \right] \\ &= \min_{u_3, \dots, u_I} \left[\sum_{i=4}^I d_i(u_i) + \left\{ d_3(u_3) + \min_{u_2} g_2(u_2) \right\} \right] \end{aligned} \quad (\text{B.4})$$

となる . ただし , 境界条件を ,

$$g_1(u_1) = d_1(u_1) \quad (\text{B.5})$$

とする . さらに ,

$$g_3(u_3) = d_3(u_3) + \min_{u_2} g_2(u_2) \quad (\text{B.6})$$

と置くと ,

$$\min F = \min_{u_3, \dots, u_I} \left[\sum_{i=4}^I d_i(u_i) + g_3(u_3) \right] \quad (\text{B.7})$$

となる . 結局 ,

$$g_i(u_i) = d_i(u_i) + \min_{u_{i-1}} g_{i-1}(u_{i-1}) \quad (\text{B.8})$$

の漸化式を計算することで ,

$$\min F = \min_{u_I} g_I(u_I) \quad (\text{B.9})$$

が得られる . ただし , 式中の \min の u_i の選び方は , 回路形状により拘束されるものとする .

次に , 図 B.2 のような立体回路網で , 常に上方向に動き , $i = 1$ の面上のある点から $i = I$ の面上のある点に動く間に通過する数 $d_i(j, k)$ の和を最小にする経路

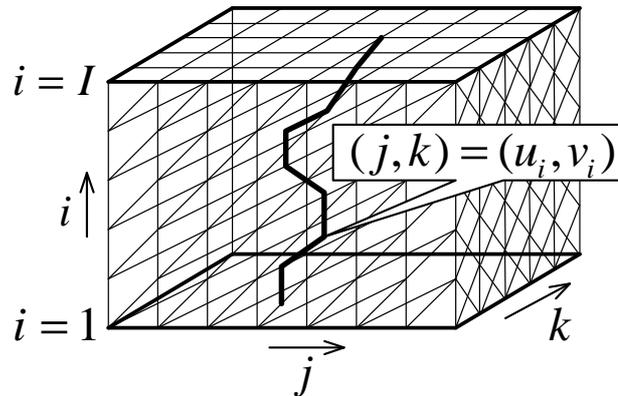


図 B.2 立体回路網

$(u_1, v_1), (u_2, v_2), \dots, (u_I, v_I)$ を探索する問題を考える. この場合, 問題は次式で置き換えられる.

$$\begin{aligned} \min F &= \min_{(u_1, v_1), (u_2, v_2), \dots, (u_I, v_I)} \sum_{i=1}^I d_i(u_i, v_i) \\ &= \min_{(u_2, v_2), \dots, (u_I, v_I)} \left[\sum_{i=3}^I d_i(u_i, v_i) + \left\{ d_2(u_2, v_2) + \min_{(u_1, v_1)} d_1(u_1, v_1) \right\} \right] \end{aligned} \quad (\text{B.10})$$

式 (B.2) ~ (B.9) と同様の式展開により, 以下の漸化式を計算することで解ける.

$$g_i(u_i, v_i) = d_i(u_i, v_i) + \min_{(u_{i-1}, v_{i-1})} g_{i-1}(u_{i-1}, v_{i-1}) \quad (\text{B.11})$$

ただし, 境界条件を,

$$g_1(u_1, v_1) = d_1(u_1, v_1) \quad (\text{B.12})$$

とし, \min の選び方は回路形状に拘束される. 第 4 章で紹介する拡張 DP マッチングはさらに 4 次元の回路網における問題の結果を使用するが, この問題も同様の定式化で解ける.

付録 C

メディアンの定義

メディアン（中央値）は代表値の一つである．ここでは，第 3 章で用いる 1 次元のメディアンとその多次元拡張である多次元のメディアンの定義について説明する．

C.1 1 次元のメディアン

実数値 x_1, x_2, \dots, x_N のメディアンは以下の 2 つのように定義することができる [79][102]．なお，実数値の数が奇数の場合，2 つの定義によるそれぞれのメディアンは同じ数値となる．

(a) 定義 1

1 次元のメディアンは，実数値を昇順に並べた場合の中央に位置する値と定義される（図 C.1）．ただし，実数値の数が偶数の場合，メディアンは中央に位置する 2 つの値の平均値とする．

(b) 定義 2

1 次元のメディアンは，絶対偏差の和を最小とする要素として定義することもできる（図 C.2）．

$$m = \arg \min_{x \in \{x_1, \dots, x_N\}} \sum_{i=1}^N |x - x_i| \quad (\text{C.1})$$

式 (C.1) 中の $|\cdot|$ は絶対値記号である．ただし，実数値の数が偶数の場合，メディアンは一意に定まらない．

統計的な解釈としては，1 次元の双指数関数の確率密度分布 $f(x)$ を用いて，最尤推定

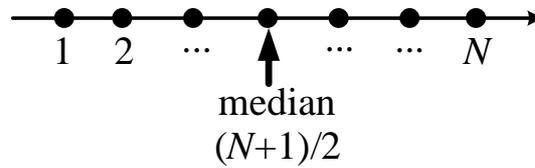


図 C.1 1次元のメディアン（定義1）.

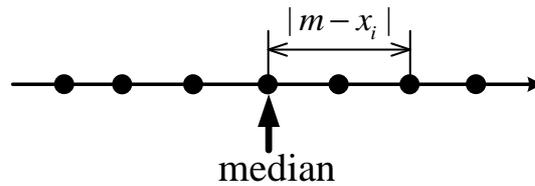


図 C.2 1次元のメディアン（定義2）.

によって分布の代表値 β を決定することに相当する． α をスケーリング係数とすると，確率密度分布 $f(x)$ と尤度関数 $L(\beta)$ は次式のようにになる．

$$f(x) = \frac{\alpha}{2} e^{-\alpha|x-\beta|} \quad (\text{C.2})$$

$$L(\beta) = \prod_{i=1}^N \frac{\alpha}{2} e^{-\alpha|x_i-\beta|} \quad (\text{C.3})$$

C.2 多次元のメディアン

多次元のベクトル $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ のメディアンは他のベクトルとの距離の和を最小にするベクトルとして定義することができる（図 C.3） [79]．

$$\mathbf{m} = \arg \min_{\mathbf{x} \in \{\mathbf{x}_1, \dots, \mathbf{x}_N\}} \sum_{i=1}^N \|\mathbf{x} - \mathbf{x}_i\|_{L_2} \quad (\text{C.4})$$

ここで， $\|\cdot\|_{L_2}$ は L_2 ノルムである．ただし，ベクトルの数が偶数の場合，メディアンは一意に定まらない．

統計的な解釈としては，1次元のメディアンの場合と同様に，多次元の実指数関数の確率密度分布 $f(\mathbf{x})$ を用いて，最尤推定によって分布の代表値 β を決定することに相当する． α と γ をスケーリング係数とすると，確率密度分布 $f(\mathbf{x})$ と尤度関数 $L(\beta)$ は次式の

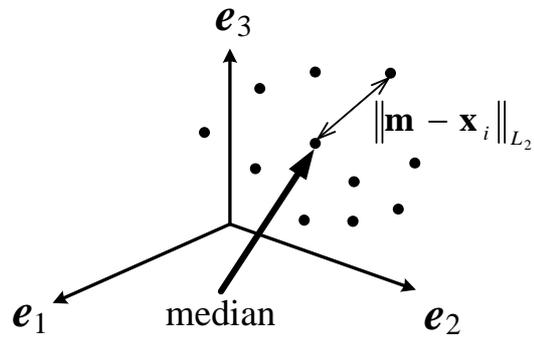


図 C.3 1次元のメディアン .

ようになる .

$$f(\mathbf{x}) = \gamma e^{-\alpha \|\mathbf{x} - \beta\|_{L_2}} \quad (\text{C.5})$$

$$L(\beta) = \prod_{i=1}^N \gamma e^{-\alpha \|\mathbf{x}_i - \beta\|_{L_2}} \quad (\text{C.6})$$

多次元のメディアンは1次元のメディアンを一般化したものである . そのため , 多次元のメディアンは1次元のメディアンと同様に , はずれ値を除外する性質がある . また , 多次元のメディアンの次元を1とした場合 , 1次元のメディアンの定義2と一致する .

参考文献

- [1] 片桐雅二, 杉村利明, 櫻井保志, 池田 武史, “実写ライブ動画映像を用いた移動体ナビゲーションの試み,” 信学技報, PRMU98-169, Dec. 1998 .
- [2] 阿部友紀, “ある一漁民の祈願と生業 : 山形県鶴岡市由良地区にみる,” 東北宗教学, Vol. 3, pp. 17–35, Dec. 2007 .
- [3] 小川琢治, “数理地理學,” 宇宙物理學研究會, 1948 .
- [4] 岩佐英彦, 粟飯原述宏, 横矢直和, 竹村治雄, “全方位画像を用いた記憶に基づく位置推定,” 信学論 (D-II), Vol. J84-D-II, No. 2, pp. 310–320, Feb. 2001 .
- [5] Google Maps, <http://maps.google.co.jp/>.
- [6] MapJack, <http://www.mapjack.com/>.
- [7] H. Kyutoku, D. Deguchi, T. Takahashi, Y. Mekada, I. Ide, and H. Murase, “On-road Obstacle Detection by Comparing Present and Past In-vehicle Camera Images,” Proc. IAPR Conf. on Machine Vision Applications (MVA) 2011, pp. 357–360, Jun. 2011.
- [8] A. Torii, M. Havlena, and T. Pajdla, “From Google Street View to 3D City Models,” Proc. IEEE 12th Int. Conf. on Computer Vision Workshops, pp. 2188–2195, Oct. 2009.
- [9] J. Böhm, “Multi-image Fusion for Occlusion-free Façade Texturing,” International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 35, Part B5, pp. 867–872, Jul. 2004.
- [10] A. Levin, A. Zomet, and Y. Weiss, “Learning How to Inpaint from Global Image

- Statistics,” Proc. IEEE 9th Int. Conf. on Computer Vision, pp. 305–312, Oct. 2003.
- [11] Y. Wexler, E. Shechtman, and M. Irani, “Space-Time Video Completion,” Proc. 2004 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, Vol. 1, pp. 120–127, Jun. 2004.
- [12] 松本吉央, 稲葉雅幸, 井上博允, “ビューベーストアプローチに基づく移動ロボットナビゲーション,” 日本ロボット学会誌, Vol. 20, No. 5, pp. 506–514, Jul. 2002.
- [13] Flickr, <http://www.flickr.com/>.
- [14] J. Zheng, “Digital Route Panoramas,” IEEE MultiMedia, Vol. 10, Issue 3, pp. 57–67, Jul. 2003.
- [15] Photosynth, <http://photosynth.net/>.
- [16] 尾池治郎, 川崎洋, 大澤裕, 小野晋太郎, 池内克史, “固有画像を用いたイメージベースによる仮想広域空間の実時間レンダリング手法,” 画像の認識・理解シンポジウム (MIRU) 2005 講演予稿集, pp. 1547–1553, Jul. 2005.
- [17] 川崎洋, 谷田部智之, 池内克史, 坂内正夫, “Omni ビデオカメラからの3次元都市の自動生成,” 情報処理学会研究報告, 99-CVIM-119-4, Nov. 1999.
- [18] S. Agarwal, N. Snavely, I. Simon, S. M. Seitz, and R. Szeliski, “Building Rome in a Day,” Proc. IEEE 12th Int. Conf. on Computer Vision, pp. 72–79, Sep. 2009.
- [19] B. Micusik, and J. Kosecka, “Piecewise Planar City 3D Modeling from Street View Panoramic Sequences,” Proc. 2009 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp. 2906–2912, Jun. 2009.
- [20] Panoramio, <http://www.panoramio.com/>.
- [21] 石川貴一郎, 村石隆介, 天野嘉春, 橋詰匠, 島嘉宏, 瀧口純一, 清水聡, “モバイルマッピングシステムと三次元空間情報管理システム (G-VIZTM) を用いた実都市空間モデリング&ウォークスルーシステム,” 第13回ロボティクスシンポジウム, Mar. 2008.

- [22] 山野芳樹, “モバイルマッピングシステムによる道路空間データ収集手法の開発,” 測量調査技術, No. 77, pp. 1–10, Nov. 2000.
- [23] 瀧口純一, 梶原尚幸, 島嘉宏, 廣川類, 黒崎隆二郎, 目黒淳一, 石川貴一郎, 天野嘉春, 橋詰匠, “高精度 GPS 移動測位技術を利用した移動監視システムとモバイルマッピングシステムの開発,” 第 25 回日本ロボット学会学術講演会, Sep. 2007.
- [24] 佐藤准嗣, 高橋友和, 井手一郎, 村瀬洋, “GPS 座標付き全方位映像群からの市街地映像マップの構築と街並変化の検出,” 信学論 (D-II), Vol. J90-D, No. 4, pp. 1085–1095, Apr. 2007.
- [25] N. Enami, N. Ukita, and M. Kidode, “Image Matching with a Car-mounted Camera Robust to Change in Imaging Conditions,” Int. J. of Pattern Recognition and Artificial Intelligence, Vol. 23, No. 7, pp. 1369–1395, Nov. 2009.
- [26] N. Snavely, S. Seitz, and R. Szeliski, “Photo Tourism: Exploring Photo Collections in 3D,” ACM Transactions on Graphics, Vol. 25, Issue 3, pp. 835–846, Jul. 2006.
- [27] T. Ishikawa, T. Yamasaki, and K. Aizawa, “Car Location Matching Using Front & Side-View and GPS Data,” 信学技報, IE2005-245, Jan. 2006.
- [28] 七里慶, 新井イスマイル, 西尾信彦, “異なる時間の写真をパッチング可能とする全方位パノラマビューアの開発,” 情報処理学会第 72 回全国大会論文集, Vol. 3, pp. 337–338, Mar. 2010.
- [29] 小島祥子, 山田啓一, 二宮芳樹, “画像と地図情報の融合による道路形状推定,” 信学技報, PRMU2000-143, Dec. 2000.
- [30] 三菱電機報道資料, “街並み映像を用いた「携帯電話向け道案内システム」を開発”, <http://www.mitsubishielectric.co.jp/news/2003/0925.htm>, Jul. 2003.
- [31] J. Blackman, “Omniveillance, Google, Privacy in Public, and the Right to Your Digital Identity: A Tort for Recording and Disseminating an Individual’s Image over the Internet,” Santa Clara Law Review, Vol. 49, pp. 313–392, Nov. 2008.
- [32] A. Frome, G. Cheung, A. Abdulkader, M. Zennaro, B. Wu, A. Bissacco, H.

- Adam, H. Neven, and L. Vincent, “Large-scale Privacy Protection in Google Street View,” Proc. IEEE 12th Int. Conf. on Computer Vision, pp. 2373–2380, Sept. 2009.
- [33] J. Segall, “Google Street View: Walking the Line of Privacy–Intrusion upon Seclusion and Publicity Given to Private Facts in the Digital Age,” University of Pittsburgh Journal of Technology Law & Policy, Vol. 10, No. 1, pp. 23–32, 2010.
- [34] 高橋英之, 堀磨伊也, 神原誠之, 横矢直和, “全天球画像データベース作成のための色調統一と動物体の除去,” 信学技報, MVE2009-40, Oct. 2009 .
- [35] 飯吉建彰, 三橋渉, “画像モザイク生成のための部分的な前景の自動除去,” 信学論 (D), Vol. J92-D, No. 8, pp. 1403–1413, Aug. 2009 .
- [36] 新井イスマイル, 堀磨伊也, 河合紀彦, 安部陽平, 市川昌宏, 里中祐輔, 新田竜規, 新田知之, 藤井陽光, 向井政貴, 堀見宗一郎, 牧田孝嗣, 神原誠之, 西尾信彦, 横矢直和, “人が消える地下街パノラマビューアの開発,” 情報処理学会, マルチメディア, 分散, 協調とモバイル (DICOMO2010) シンポジウム論文集, pp. 1183–1189, Jul. 2010 .
- [37] N. Kawai, T. Sato, and N. Yokoya, “Image Inpainting Based on Energy Minimization,” Proc. SPIE Electronic Imaging, Vol. 6498, pp. 649813.1–649813.9, Jan. 2007.
- [38] C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman, “PatchMatch: A Randomized Correspondence Algorithm for Structural Image Editing,” Proc. ACM Special Interest Group on Computer Graphics (SIGGRAPH) 2009, pp. 24:1–24:11, Jul. 2009.
- [39] 栗林宏輔, 川崎洋, 小野晋太郎, 池内克史, “移動カメラ映像の障害物除去のための時空間画像フィルタの提案,” 画像の認識・理解シンポジウム (MIRU) 2009 論文集, IS2-32, pp. 1065–1072, Jul. 2009 .
- [40] A. Yamashita, I. Fukuchi, T. Kaneko, and K. Miura, “Removal of Adherent Noises from Image Sequences by Spatio-Temporal Image Processing,” Proc.

- 2008 IEEE Int. Conf. on Robotics and Automation, pp.2386–2391, May 2008.
- [41] A. Flores, and S. Belongie, “Removing Pedestrians from Google Street View Images,” Proc. 2010 IEEE Int. Workshop on Mobile Vision, pp. 53–58, Jun. 2010.
- [42] C. Früh, and A. Zakhor, “An Automated Method for Large-Scale, Ground-Based City Model Acquisition,” Int. J. of Computer Vision, Vol. 60, No. 1, pp. 5–24, Oct. 2004.
- [43] 安田明生, “GPS 技術の展望,” 信学論 (B), Vol. J84-B, No. 12, pp. 2082–2091, Dec. 2001.
- [44] 臼井澄夫, “高度衛星測位技術の動向と応用,” 信学誌, Vol. 92, No. 9, pp. 768–774, Sep. 2009.
- [45] S. Brakatsoulas, D. Pfoser, R. Salas, and C. Wenk, “On Map-matching Vehicle Tracking Data,” Proc. 31st Very Large Data Base Conference (VLDB), pp. 853–864, Aug. 2005.
- [46] 那須野真人, 渋谷奈保, 高橋友和, 井手一郎, 村瀬洋, 小島祥子, “車載レーザレーダを用いたフレーム間差分及び L 字形状判別による移動車両検出,” 2008 年電子情報通信学会総合大会講演論文集, Vol. 情報・システム 2, p. 222, Mar. 2008.
- [47] 小野晋太郎, 川崎洋, 池内克史, “時空間距離画像の提案・考察と町のモデル化への応用,” 情報処理学会研究報告, CVIM2004-142, Jan. 2004.
- [48] 渋谷奈保, 高橋友和, 井手一郎, 村瀬洋, 小島祥子, 高橋新, “車載レーザスキャナによる距離データマップの構築と高精度自車位置推定,” 信学論 (D), Vol. J92-D, No. 2, pp. 215–225, Feb. 2009.
- [49] T. Weiss, N. Kaempchen, and K. Dietmayer, “Precise Ego-localization in Urban Areas Using Laserscanner and High Accuracy Feature Maps,” Proc. 2005 IEEE Intelligent Vehicles Symposium (IV2005), pp. 284–289, Jun. 2005.
- [50] M. Dissanayake, P. Newman, S. Clark, H. Durrant-Whyte, and M. Csorba, “A Solution to the Simultaneous Localization and Map Building (SLAM) Problem,”

- IEEE Trans. on Robotics and Automation, Vol. 17, Issue 3, pp. 229–241, Jun. 2001.
- [51] J. D. Tards, J. Neira, P. M. Newman, and J. J. Leonard, “Robust Mapping and Localization in Indoor Environments Using Sonar Data,” *Int. J. of Robotics Research*, Vol. 21, Issue 4, pp.311–330, Apr. 2002.
- [52] 道満恵介, 高橋友和, 目加田慶人, 井手一郎, 村瀬洋, “生成型学習とカスケード型識別器による道路標識検出・認識,” *画像の認識・理解シンポジウム (MIRU) 2008 講演予稿集*, IS1-7, pp. 376–381, Jul. 2008 .
- [53] 森屋崇, 石川直人, 中島真人, “市街地における車両衝突防止システム,” *信学技報*, IE2001-192, Jan. 2002 .
- [54] 山口直人, 田森信行, 塩見彰睦, “適応エッジ保存平滑化を用いた白線検出手法,” *信学論 (D-II)*, Vol. J88-D-II, No. 8, pp. 1421–1431, Sep. 2005 .
- [55] 中森卓馬, 石川直人, 中島真人, “動画像処理による車両前方監視システム,” *信学技報*, ITS2001-51, Jan. 2002 .
- [56] M. Betke, and L. Gurvits, “Mobile Robot Localization Using Landmarks,” *IEEE Trans. on Robotics and Automation*, Vol. 13, No. 2, pp. 251–263, Apr. 1997.
- [57] E. Royer, J. Bom, M. Dhome, B. Thuillot, M. Lhuillier, and F. Marmoiton, “Outdoor Autonomous Navigation Using Monocular Vision,” *Proc. Int. Conf. on Intelligent Robots and Systems (IROS) 2005*, pp. 3395–3400, Aug. 2005.
- [58] S. Se, D. Lowe, and J. Little, “Vision-based Global Localization and Mapping for Mobile Robots,” *IEEE Trans. on Robotics*, Vol. 21, No. 3, pp. 364–375, Jun. 2005.
- [59] 薄充孝, 中川知香, 佐藤智和, 横矢直和, “ランドマークデータベースに基づく静止画像からのカメラ位置・姿勢推定,” *日本バーチャルリアリティ学論*, Vol. 13, No. 2, pp. 161–170, Jun. 2008 .
- [60] M. Jabbour, V. Cherfaoui, and P. Bonnifait, “Management of Landmarks in a GIS for an Enhanced Localisation in Urban Areas,” *Proc. 2006 IEEE Intelligent*

- Vehicles Symposium, pp. 50–57, Sept. 2006.
- [61] 小島浩, 高橋勝彦, 関進, 岡隆一, “走行ロボット動画像からのモーションシーンのスポッティング認識,” 信学技報, PRU93-58, Sep. 1993 .
- [62] 小島浩, 伊藤慶明, 岡隆一, “Reference Interval-Free 連続 DP を利用した移動ロボットの時系列画像による位置同定システム,” 信学論 (D-II), Vol. J80-D-2, No. 3, pp. 724–733, Mar. 1997 .
- [63] 小野晋太郎, 松久亮太, 川崎洋, 池内克史, “車載カメラ映像の時空間マッチングによる自車位置推定,” 生産研究, Vol. 63, No. 2, pp. 209–215, Jun. 2011 .
- [64] 山澤一誠, 八木康史, 谷内田正彦, “移動ロボットのナビゲーションのための全方位視覚センサ HyperOmni Vision の提案,” 信学論 (D-II), Vol. J79-D-II, No. 5, pp. 698–707, May 1996 .
- [65] 前田佐嘉志, 白井良明, “固有空間法の位置推定能力を考慮した効率的な観測戦略による位置推定,” 情報処理学会研究報告, 2004-CVIM-144, Apr. 2004 .
- [66] 文仁赫, 三浦純, 白井良明, “不確かさを考慮した観測位置と移動のオンライン計画手法,” 日本ロボット学会誌, Vol. 17, No. 8, pp. 1107–1113, Nov. 1999 .
- [67] 前田佐嘉志, 久野義徳, 白井良明, “固有空間解析に基づく移動ロボットの位置認識,” 信学論 (D-II), Vol. J80-D-II, No. 6, pp.1502–1511, Jul. 1997 .
- [68] 石川敬之, 山崎俊彦, 相澤清晴, “GPS と画像特徴を用いた車載ライフログ映像処理,” 画像の認識・理解シンポジウム (MIRU) 2005 論文集, pp. 604–696, Jul. 2005 .
- [69] N. Yazawa, H. Uchiyama, and H. Saito, “Image Based View Localization System Retrieving from a Panorama Database by SURF,” Proc. IAPR Conf. on Machine Vision Applications (MVA) 2009, pp.118–121, May 2009.
- [70] 内田誠一, “DP マッチング概説 ~ 基本と様々な拡張 ~,” 信学技報, PRMU2006-166, Dec. 2006 .
- [71] J. Rongrong, S. Xiaoshui, Y. Hongxun, X. Pengfei, L. Tianqiang, and L. Xianning, “Attention-driven Action Retrieval with DTW-based 3D Descriptor

- Matching,” Proc. 16th ACM Int. Conf. on Multimedia, pp. 619–622, Oct. 2008.
- [72] T. Darrell, and A. Pentland, “Space-time Gestures,” Proc. 2004 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp. 335–340, Jun. 1993.
- [73] 岡隆一, “連続 DP を用いた連続音声認識,” 音響学会音声研資料, S78-20, Jun. 1978 .
- [74] 伊藤慶明, “時系列パターンの任意部分区間の高速マッチング手法 Shift CDP 法,” 信学論 (D-II), Vol. J86-D-2, No. 9, pp. 1267–1277, Sep. 2003 .
- [75] 西村拓一, 野崎俊輔, 岡隆一, “Non-monotonic 連続 DP によるスポットティングに基づく移動ロボットの時系列画像を用いた大局的位置の推定,” 信学論 (D-II), Vol. J80-D-2, No. 8, pp. 1876–1884, Aug. 1998 .
- [76] D. Rueckert, L.I. Sonoda, C. Hayes, D.L.G. Hill, M.O. Leach, and D.J. Hawkes, “Nonrigid Registration Using Free-form Deformations: Application to Breast MR Images,” IEEE Trans. on Medical Imaging, Vol. 18, No. 8, pp. 712–721, Aug. 1999.
- [77] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, “Multimodality Image Registration by Maximization of Mutual Information,” IEEE Trans. on Medical Imaging, Vol. 16, No. 2, pp. 187–198, Apr. 1997.
- [78] C. Studholme, D. Hill, D. Hawkes, “An Overlap Invariant Entropy Measure of 3D Medical Image Alignment,” Pattern Recognition, Vol. 32, Issue 1, pp. 71–86, Jan. 1999.
- [79] J. Astola, P. Haavisto, and Y. Neuvo, “Vector Median Filters,” Proc. of the IEEE, Vol. 78, No. 4, pp. 678–689, Apr. 1990.
- [80] M. Modata, G. Ridgwaya, Z. Taylora, M. Lehmannb, J. Barnesb, D. Hawkesa, N. Foxb, and S. Ourselina, “Fast Free-form Deformation using Graphics Processing Units,” Computer Methods and Programs in Biomedicine, Vol. 98, Issue 3, pp. 278–284, Jun. 2010.

- [81] B. Glocker, N. Komodakis, G. Tziritas, N. Navab, and N. Paragios, “Dense Image Registration through MRFs and Efficient Linear Programming,” *Medical Image Analysis*, Vol. 12, Issue 6, pp. 731–741, Dec. 2008.
- [82] Y. Yaguchi, K. Iseki, N.T. Viet, and R. Oka, “3D Object Reconstruction using Full Pixel Matching,” *Proc. 13th Int. Conf. on Computer Analysis of Images and Patterns*, pp.873–880, Sept. 2009.
- [83] M. Celebi, H. Kingravi, and Y. Aslandogan, “Nonlinear Vector Filtering for Impulsive Noise Removal from Color Images,” *J. of Electronic Imaging*, Vol. 16, Issue 3, 033008, Oct. 2007.
- [84] P. Trahanias, and A. Venetsanopoulos, “Vector Directional Filters: A New Class of Multichannel Image Processing Filters,” *IEEE Trans. on Image Processing*, Vol. 4, Issue 4, pp. 528–534, Oct. 1993.
- [85] D. Karakos, and P. Trahanias, “Combining Vector Median and Vector Directional Filters: The Directional Distance Filters,” *Proc. 1995 Int. Conf. on Image Processing*, pp. 171–174, Oct. 1995.
- [86] N. Plataniotis, D. Androutsos, and A. Venetsanopoulos, “Content-based Colour Image Filters,” *Electronics Letters*, Vol. 33, Issue 3, pp. 202–203, Jan. 1997.
- [87] 国土交通省道路局, “道路統計年報 2007,” Nov. 2007.
- [88] R. ベルマン, S. ドレイファス, “応用ダイナミック・プログラミング,” 日本科学技術連盟, 1962.
- [89] パイオニア報道資料, “カロツェリア サイバーナビ 4 機種を新発売,” <http://pioneer.jp/press/2011/0509-1.html>, May 2011.
- [90] 寺田智裕, 神原誠之, 横矢直和, “拡張現実感を用いた車載型アノテーションシステムの構築,” 信学技報, MVE2001-136, Feb. 2002.
- [91] F. Aldershoff, and T. Gevers, “Visual Tracking and Localization of Billboards in Streamed Soccer Matches,” *Proc. SPIE Electronic Imaging 2004*, pp. 408–416, Jan. 2004.

- [92] G. Medioni, G. Guy, H. Rom, and A. François, “Real-Time Billboard Substitution in a Video Stream,” Proc. 10th Tyrrhenian Int. Workshop on Digital Communications, pp. 71–84, Oct. 1998.
- [93] セカイカメラ , <http://sekaicamera.com/>.
- [94] BMW, http://www.bmw.co.uk/bmwuk/augmented_reality/homepage/.
- [95] T. Lee, and T. Hollerer, “Handy AR: Markerless Inspection of Augmented Reality Objects Using Fingertip Tracking,” Proc. 11th IEEE Int. Symp. on Wearable Computers, pp. 1–8, Oct. 2007.
- [96] 隅田知代, 増永良文, “拡張現実感を用いた建物名重畳システムにおけるセンサ誤差の扱い,” 信学技報, DE2005-124, Jul. 2005 .
- [97] C. Harris, and M. Stephens, “A combined corner and edge detector,” Proc. 4th Alvey Vision Conf., pp. 147–151, Aug. 1998.
- [98] J. Shi, and C. Tomasi, “Good Features to Track,” Proc. 1994 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp. 593–600, Jun. 1994.
- [99] R. Hartley, and A. Zisserman, “Multiple View Geometry in Computer Vision,” Cambridge University Press, 2004.
- [100] 田村秀行, “コンピュータ画像処理,” オーム社, 2002 .
- [101] 岡谷貴之, “バンドルアジャストメント,” 情報処理学会研究報告, 2009-CVIM-167, Jun. 2009 .
- [102] J. Burt, G. Barber, and D. Rigby, “Elementary Statistics for Geographers,” The Guilford Press, 2009.

研究業績

学術論文

- 内山寛之, 出口大輔, 高橋友和, 井手一郎, 村瀬洋, “拡張 DP マッチングを用いた視野角の異なるカメラ映像間の時空間対応付けによる自車位置推定,” 電子情報通信学会論文誌 (D), Vol. J93-D, No. 12, pp. 2659–2665, Dec. 2010.
- 内山寛之, 出口大輔, 高橋友和, 井手一郎, 村瀬洋, “複数画像系列の部分画像選択に基づく移動物体を含まない車載カメラ映像の生成,” 電子情報通信学会論文誌 (D), Vol. J94-D, No. 12, pp. 2093–2104, Dec. 2011.

国際会議

- H. Uchiyama, T. Takahashi, I. Ide, and H. Murase, “Frame Registration of In-Vehicle Normal Camera with Omni-Directional Camera for Self-Position Estimation,” Proc. of Third Int. Conf. on Innovative Computing, Information and Control, WS01-007, Jun. 2008.
- H. Uchiyama, D. Deguchi, T. Takahashi, I. Ide, and H. Murase, “Ego-localization using Streetscape Image Sequences from In-vehicle Cameras,” Proc. of 2009 IEEE Intelligent Vehicles Symposium, pp.185–190, Jun. 2009.
- H. Uchiyama, D. Deguchi, T. Takahashi, I. Ide, and H. Murase, “Removal of Moving Objects from a Street-View Image by Fusing Multiple Image Sequences,” Proc. of Int. Conf. on Pattern Recognition (ICPR 2010), pp. 3456–3459, Aug. 2010.
- H. Uchiyama, D. Deguchi, T. Takahashi, I. Ide, and H. Murase, “3-D Line

Segment Reconstruction using an In-Vehicle Camera for Free Space Detection,”
Proc. of 2011 IEEE Intelligent Vehicles Symposium, pp.290–295, Jun. 2011.

国内会議・シンポジウム等

- 内山寛之，高橋友和，井手一郎，村瀬洋，“自車位置推定のための車載全方位カメラ映像と通常カメラ映像のフレーム間対応付け，” 電子情報通信学会技術研究報告，PRMU2007-272，Mar. 2008 .
- 内山寛之，出口大輔，高橋友和，井手一郎，村瀬洋，“街並み画像系列と2台の車載カメラの画像系列との照合による自車の走行位置推定，” 電子情報通信学会技術研究報告，PRMU2008-159，Dec. 2008 .
- 内山寛之，出口大輔，高橋友和，井手一郎，村瀬洋，“市街地映像マップと2台の車載カメラを用いた自車位置推定，” 第12回 画像の認識・理解シンポジウム (MIRU) 2009 論文集，IS2-21，pp. 992–999，Jul. 2009 .
- 内山寛之，出口大輔，高橋友和，井手一郎，村瀬洋，“複数映像統合による車載カメラからの移動物体除去手法の検討，” 平成21年度電気関係学会東海支部連合大会，O-139，Sep. 2009 .
- 内山寛之，出口大輔，高橋友和，井手一郎，村瀬洋，“複数映像統合による車載カメラ映像からの移動物体除去，” 電子情報通信学会技術研究報告，PRMU2009-102，Nov. 2009 .
- 内山寛之，出口大輔，高橋友和，井手一郎，村瀬洋，“複数車載カメラ映像の部分画像選択に基づく移動物体のない映像の生成，” 第13回 画像の認識・理解シンポジウム (MIRU) 2010 論文集，IS3-62，pp. 1991–1998，Jul. 2010 .
- 内山寛之，出口大輔，高橋友和，井手一郎，村瀬洋，“車載カメラ映像中からの移動物体除去を目的とした複数映像統合手法の比較検討，” 平成22年度電気関係学会東海支部連合大会，P5-5，Aug. 2010 .
- 内山寛之，出口大輔，高橋友和，井手一郎，村瀬洋，“車載カメラ映像中の線分の3次元位置推定と走行可能領域検出への応用，” 第14回 画像の認識・理解シンポジウム (MIRU) 2011 論文集，IS4-9，pp. 1340–1346，Jul. 2011 .
- 内山寛之，出口大輔，高橋友和，井手一郎，村瀬洋，“市街地構造物への拡張現実

型画像情報提示手法の予備的検討，”平成 23 年度電気関係学会東海支部連合大会，
C1-5，Sep. 2011 .

- 内山寛之，出口大輔，井手一郎，村瀬洋，川西隆仁，柏野邦夫，“市街地構造物への
拡張現実型広告提示，”ビジョン技術の実利用ワークショップ (ViEW) 2011 講演
論文集，pp. 434-438，Dec. 2011 .

受賞

- IEEE Nagoya Section Student Paper Award，平成 21 年度電気関係学会東海支
部連合大会，Jan. 2009 .

その他の研究業績

- 林泰宏，内山寛之，出口大輔，井手一郎，村瀬洋，“位置によらない動き特徴を
用いた遠隔共同調理映像間の動作対応付け，”電子情報通信学会技術研究報告，
MVE2010-145，Mar. 2011 .