

博士論文

車載単眼カメラを用いた地図情報更新システム
のための画像解析要素技術

榎並 直子

2011年 3月 30日

奈良先端科学技術大学院大学
情報科学研究科 情報処理学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に
博士(工学) 授与の要件として提出した博士論文である。

榎並 直子

審査委員：

木戸出 正継 教授 (主指導教員)
千原 國宏 教授 (副指導教員)
浮田 宗伯 准教授 (副指導教員)
村瀬 洋 教授 (名古屋大学)

車載単眼カメラを用いた地図情報更新システム のための画像解析要素技術*

榎並 直子

内容梗概

本論文では、高頻度のデジタル地図情報の更新を可能とする、ドライブレコーダ用の単眼カメラと普及型 GPS を搭載した一般車両を利用するデジタル地図情報更新システムの構築を目指し、システムに必要な画像解析要素技術の提案を行う。

現在、正確なナビゲーションにとって重要な地図情報の更新は、少数の観測車両による現地調査などにより行われており、高いコストと時間が必要である。そのため、地図情報を最新の状態に維持することは困難であり、正確性が求められる地図にとって致命的な問題である。これに対し提案するシステムでは安価なセンサを搭載した多数の一般車両による観測と画像解析要素技術により、低コストかつ高頻度での地図情報更新の実現を目指す。

はじめに、デジタル地図情報更新システムの枠組みを提案する。システムでは、観測車両に多数の一般車両を用いる。また安価で一般に普及してきているドライブレコーダ用の単眼カメラと普及型 GPS を組み合わせて用いるなど、実用に耐えうるシステム構成を目指す。またシステムでは、常時観測により得られる大量画像集合を画像解析処理により、自動的にチェックすることで確認作業のコスト低減を目指し、そのために必要な画像解析要素技術について検討する。

次に、システム実現に必要な大量の観測情報を効率的に処理するための画像解析要素技術を提案する。

*奈良先端科学技術大学院大学 情報科学研究科 情報処理学専攻 博士論文, NAIST-IS-DD0861001, 2011年3月30日.

まず、最新の入力データの画像と同一地点を撮影した街並情報データベース上の過去データの画像を対応付ける必要がある。本研究では多様な車両・カメラにより様々な天候・走行条件で撮影された低解像度の画像を、GPS から得られた位置情報と撮影条件の変化に対して頑健な段階的な画像間のマッチングにより対応付ける手法を提案する。この処理は2段階の処理で構成される。第1段階では異なる撮影条件下で撮影された画像から、画像補正を行い連続数フレームを繋げた相似した見えのパノラマ画像間での対応付けを行なう。第2段階では、第1段階の結果をもとに特徴点のマッチングによる対応付けを行う。システム運用時と同じ低画質で撮影条件の異なる画像間の対応付け実験により、提案手法の有効性を確認した。

そして、対応付けられた過去画像と入力画像の比較により街並変化を検出する手法を提案する。提案手法では、3フレームの過去画像から抽出された変化抽出結果により多数決を行うことで、事前知識なく地図情報更新に必要な建物や看板の構造やテクスチャの変化といった街並の変化検出を実現する。まず、特徴点間のマッチングにより探索領域全体に渡る大きな街並の変化の検出と画像間の対応関係推定を行う。次に画像の対応関係を利用した効率的な画素間マッチングにより変化領域抽出を行う。変化領域抽出では、計算コストの削減と誤対応の削減のため画像間から拘束条件を推定し、その拘束をもとに最適化アルゴリズム(グラフカット)を用いた画素単位のマッチングを行う。3フレームの過去画像とのマッチングにより抽出された変化抽出結果から「変化が生じている時間」により街並変化の検出を行う。対応付けの際にも利用した実画像を用いて、提案手法の有効性を確認した。

最後に、本研究で得られた成果をまとめ、さらに今後、提案システムを実用化へと繋げるために必要となる画像解析要素技術の実現性と求められる精度について議論する。

キーワード

車載カメラ, 地図情報更新, 車載画像間対応付け, 街並変化検出, パノラマ画像, SIFT, グラフカット

Image Analysis Techniques for a Map Update System using a Car-Mounted Camera*

Naoko Enami

Abstract

In this thesis, we propose a novel concept of the map information update system using normal automobiles with off-the-shelf car-mounted cameras and GPS location systems.

A car navigation system is important for a smooth car-oriented society. For accurate navigation, the map should be up-to-date. However, most map information is updated on the basis of surveillance fieldwork by a small number of investigation vehicles, which is expensive and slow. Therefore, if a system could collect such streetscape information at low cost with automatic change detection, e.g., check for new buildings or, changed signs, then the map information could be more efficiently updated.

For an efficient update of the map information, we propose a map information update system. In order to achieve such a system in reality, solving several problems such as a image matching and scene change detection are indispensable.

First, we describe the details of the map update system framework. To realize such a system, we consider the system settings and processing. Moreover, we also consider the use of image analysis technology for such a system.

Second, we propose a matching method for images captured at different times and under different conditions. This method starts, by using position information

*Doctoral Dissertation, Department of Information Processing, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-DD0861001, March 30, 2011.

obtained from the GPS receiver to identify images observed at the same location from a number of images observed at different times and locations. However, complete positive identification using only position information is impossible because off-the-shelf GPS systems have a margin of error of about 15m. We propose a two-stage matching algorithm, for image matching. The first stage creates previous and current panoramic images from sequential rectified images. The second stage compares the current and previous images by feature points that has several appealing points such as robustness for changes in illumination and rotations, and scale invariance to identify as corresponding points. To confirm the effectiveness of the proposed method, we conducted experiments matching real images captured under various conditions.

Third, we propose a method to detect scene changes from images captured at the same location at different times and under different conditions by car-mounted cameras. Since car-mounted cameras may capture almost the same image in almost the same location, comparing images captured before and after the change can detect a scene change. If a region in a newly captured image does not match with any regions in an old image, this region can be regarded as a changed region in the scene, i.e., a scene change. A scene changes is detected by the majority from a change extraction result that are obtained by comparing the new image with three old images. To perform this comparison first SIFT keypoints are utilized to obtain the fundamental matrix from the images. Then, a dense region matching algorithm (i.e., Graph-cuts) with constraint efficiently detects changed regions between the two images. We evaluated this effectiveness with a large set of images captured by the car-mounted cameras in real environments.

Keywords:

Car-Mounted Camera , Map Update, Car-Mounted Camera Image Matching , Change Detection in Streetscapes , Panoramic Image, SIFT, Graph-cuts

目次

1. 序論	1
1.1 本論文の背景	1
1.2 本論文の目的	4
1.3 本論文の構成	6
2. デジタル地図情報更新システムの枠組み	9
2.1 システム概要	10
2.2 システム構成と処理の流れ	16
2.3 システム実現に必要な画像解析手法とその目標	22
3. 撮影条件の変化に頑健な2段階マッチングによる車載カメラ画像間の対応付け	25
3.1 はじめに	25
3.1.1 研究目的	26
3.1.2 関連研究と位置付け	28
3.1.3 提案手法の流れ	30
3.2 第1段階：パノラマ画像による撮影条件の変化に頑健な車載カメラ画像間の対応付け	32
3.2.1 撮影条件の変化によって生じる問題点	32
3.2.2 撮影条件の変化に頑健なパノラマ画像生成法	36
3.2.3 パノラマ画像を用いた画像間の対応付け	48
3.3 第2段階：SIFT特徴点による車載カメラ画像間の対応付け	52
3.3.1 短冊画像からのSIFT特徴点の抽出	53
3.3.2 SIFT特徴点によるマッチング	58
3.4 実験と評価	59
3.4.1 実験概要	59
3.4.2 実験結果	62
3.4.3 従来手法との比較	69

3.5	まとめ	74
4.	複数車載カメラ画像からの変化領域抽出による街並の変化検出	77
4.1	はじめに	77
4.1.1	研究目的	79
4.1.2	関連研究と位置付け	81
4.1.3	提案手法の流れ	83
4.2	画像の対応関係を用いた画素間マッチングによる車載カメラ画像からの変化領域抽出	87
4.2.1	SIFT 特徴点の対応付けによるエピポーラ拘束の推定	89
4.2.2	エピポーラ拘束を用いたグラフカットによる変化領域抽出	91
4.3	複数画像からの多数決原理に基づく街並変化の検出	104
4.4	実験と評価	106
4.4.1	実験概要	106
4.4.2	変化領域抽出結果	107
4.4.3	従来手法との比較	119
4.5	まとめ	124
5.	結論	127
5.1	本論文の主たる成果	127
5.2	今後の課題と展望	129
	謝辞	131
	参考文献	132
	研究業績一覧	140

目 次

1	高度道路交通システム (Intelligent Transport Systems:ITS)	2
2	本論文の構成	7
3	デジタル地図情報更新システム概要	10
4	(a) 本システムのカメラ設置位置とカメラ向き (b) 本システムの撮 影画像 (白い四角の領域が対象となる街並領域)	11
5	地図会社で行われているデジタル地図情報更新処理の流れ	16
6	提案するデジタル地図情報更新システムの処理の流れ	18
7	街並情報データベース更新の流れ	20
8	1対多の画像間マッチングとパノラマ画像間マッチング	27
9	提案手法の流れ	31
10	異なるカメラ向き, 車線から撮影された同一地点の画像	33
11	パノラマ画像生成の考え方	36
12	パノラマ画像生成の流れ	37
13	回転補正	38
14	(回転補正なしのパノラマ画像例:(a),(b),(c), 回転補正ありのパノ ラマ画像例:(d),(e),(f)	40
15	抽出線の決定方法	41
16	左右揺れに対応した抽出線を用いた結果	42
17	(a):パノラマ画像平面上の3次元投影 (b):投影画像の連結:各矢印 は余白や重なりがなく投影画像を連結するための変換を示している	44
18	生成されたパノラマ画像例	45
19	車載カメラの上下揺れ補正結果	47
20	パノラマ画像を用いたマッチングの流れ	48
21	エッジ抽出に用いるフィルタ	49
22	エッジ抽出とノイズ除去結果	49
23	第2段階:SIFT 特徴点による車載カメラ画像間の対応付けの流れ	52
24	Difference-of-Gaussian 処理	54
25	特徴点検出	55

26	特徴量の記述の範囲	57
27	特徴量の記述	58
28	SIFT 特徴点の対応付けによるマッチング結果	59
29	実験データを収集した走行区間	60
30	第1段階の対応付け結果の確認方法	62
31	撮影条件・街並・移動物体による変化が生じている画像に対しての 第1段階の対応付け結果例(成功例)	63
32	失敗した第1段階の対応付け結果例	65
33	第2段階の対応付け結果(成功例)	67
34	第2段階の対応付け結果(失敗例)	68
35	SIFT 特徴点のマッチングのみの対応付け結果	72
36	街並に変化はなく撮影条件・移動物体などの対象外の変化を含む 撮影画像例	78
37	街並の変化とその他の様々な変化を含む撮影画像例	79
38	提案手法の流れ	84
39	画像の対応関係を用いた画素間マッチングによる車載カメラ画像 からの変化領域抽出手法	87
40	エピポーラ幾何	89
41	SIFT 特徴点の対応付け結果	92
42	グラフカット例	94
43	隣接関係数の異なるグラフカットによるマッチング結果	98
44	変化マップ, マッチングエラーマップの出力画像	100
45	グラフカットによる変化領域抽出結果	102
46	多数決による街並変化検出処理の流れ	104
47	多数決による街並変化検出結果	105
48	変化領域抽出結果(成功)	108
49	変化領域抽出に影響を及ぼす照明条件の変化	109
50	エピポーラ拘束推定の失敗例	110

51	街路樹による遮蔽が生じている撮影画像からの SIFT 特徴点抽出結果例	111
52	エピポーラ拘束なしのマッチング結果	112
53	多数決による変化検出結果 (成功例) : 地点 G, Q	113
54	多数決による変化検出結果 (成功例) : 地点 D, G	114
55	多数決による変化検出結果 (成功例) : 地点 F, H	115
56	多数決による変化検出結果 (失敗例) : 地点 E, N	116
57	多数決による変化検出結果 (失敗例) : 地点 I, T	118
58	既存手法による変化領域抽出結果 (街並変化なし)	121
59	既存手法による変化領域抽出結果 (街並変化あり)	122

表 目 次

1	既存手法と提案システムの撮影画像比較	13
2	提案システムと既存手法の観測車両構成比較	14
3	第 1 段階による画像間対応付け結果	64
4	第 2 段階の画像間対応付け結果	69
5	動的計画法による画像間対応付け結果	73
6	複数画像からの街並変化の検出結果	117
7	街並変化の生じていない画像に対しての変化領域抽出結果	120
8	街並変化の生じている画像に対しての変化領域抽出結果	120

1. 序論

1.1 本論文の背景

現在、情報通信技術を用いて人・道路・車両をネットワーク化することにより、円滑な自動車交通の実現と交通事故や渋滞といった道路交通の問題解決を図る高度道路交通システム (Intelligent Transport Systems:ITS) の構築が進められている (図 1)。その研究分野の一環として高度化が進められているカーナビゲーションシステムは、円滑な自動車交通にとって重要な役割を担っている。また、その出荷台数は 2010 年度で 4129 万台を超え今後もさらなる普及が見込まれる。現在のカーナビゲーションシステムには様々な機能が付随しているが、基本機能となる正確なナビゲーションのためには、デジタル地図情報を常に最新の状態に保つ必要がある。このデジタル地図情報とは、交差点・標識・信号機の位置や車線数など道路の構造情報である「道路情報」と、固定された人工物である下記の建物・看板の位置・構造・テクスチャ(店舗のみ)である「街並情報」の 2 つの情報から構成される。

建物: 店舗, 一般建物 (マンション・一般家屋)

看板: 店舗に付随する看板 (一時的に設置された工事の看板・のぼり旗などは除く)

現在のカーナビゲーションシステムの地図情報更新周期は (1) 定期的な地図全体の一括更新 (2) 変更が生じた時に合わせての差分更新の 2 種類である。特に地図を常に最新の状態に保つというニーズを満たすため、更新が行われた地点の地図情報を通信技術を用いて配信・更新を行う地図差分配信システムは今後の地図更新方法の主流となるシステムである。しかしながら、現在道路情報の更新には有料道路で開通から最短で 7 日、一般道路で最高 3 ヶ月の時間を要する。一方、街並情報に至っては最短で半年から 1 年の時間を要する。現在の地図情報更新は、各地図会社が専用の観測車両により観測を行う実地調査や、役所や店舗などによる情報提供により行われている。そのため、地図情報の更新には高いコストと膨大な時間が必要である。加えて、地図更新の頻度は都市部と郊外など地域によって偏りがある。よって、まずは街並情報の高頻度な観測・更新が実現されなけれ

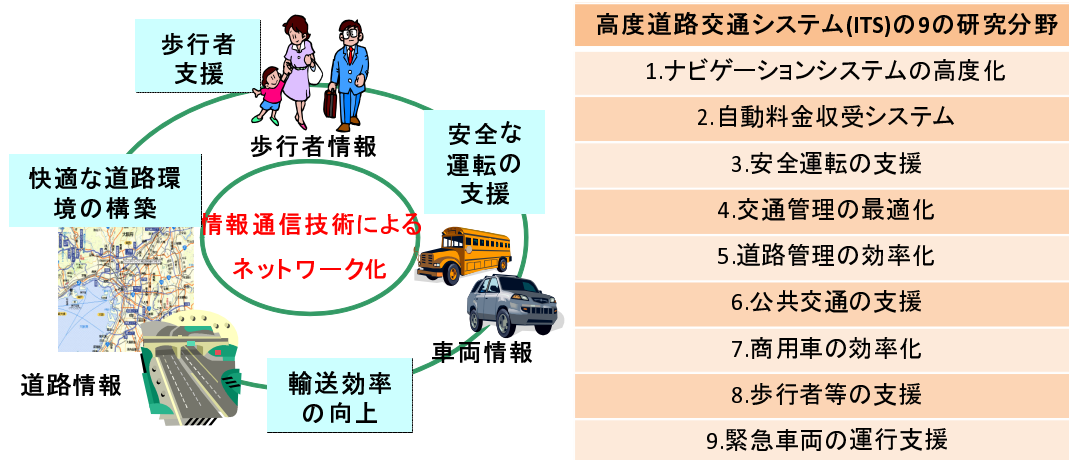


図 1 高度道路交通システム (Intelligent Transport Systems:ITS)

ば、地図を最新の状態に保つことは不可能であり、これは正確性が求められる地図情報にとって致命的な問題である。

道路情報は今後設置を行った警察や国土交通省によって、データベース化される予定である。そこで、地図情報更新の効率化を図るためには日本全域の一般道沿いに存在する街並情報を偏りなく高頻度に観測し、得られた情報を自動的に解析することで、街並情報の常時観測と効率的な更新処理を実現する必要がある。

そこで、街並を観測・解析する手法がいくつか提案されている、その1つが航空写真や衛星写真を利用する文献[1, 2, 3]である。文献[3]では、異なる時間に同一地点を撮影した衛星写真間を比較することで、道路・街並の変化を検出する。しかし上空から画像撮影を行っているため、広域の地域を一度に観測可能ではあるが、建物の側面や看板などの重要な街並情報を観測することができない。したがって、デジタル地図情報の更新に必要な情報を観測するには、街並の側面を観測する必要がある。

また、街並の3次元形状を計測(レーザレンジセンサ[4, 5]やステレオカメラ[6, 7])することでより正確に街並の変化を捉える研究も存在する。これらのセンサーを用いて生成された3次元モデルは、街並の側面を主として構築されている。文献[7]では、さらに3次元モデルにGlobal Positioning System(GPS)の測定値が

付随されており，その位置測定により同じ地点を異なる時間に観測した画像から作成した3次元モデル間の比較を可能にしている．しかしながら，正確な3次元モデルを毎回生成することは容易ではなく，加えて，不安定な撮影条件下では3次元復元は困難であり，これらの手法では安定した撮影条件のもと高価で特殊な設備を用いる必要がある．

他方では，より簡易な観測装置を利用した手法が提案されている．文献[8]では，車載全方位カメラと普及型GPSを用いた街並画像データベースの構築と街並変化検出手法を提案している．文献[8]は先述した[4, 5, 6, 7]の手法と同じく，ビルの側面の観測を目的としているため，より詳細な街並変化検出を可能としている．さらに文献[9]によって高性能なGPSと全方位カメラを用いた情報収集の手法も提案されている．

しかしながら，これらの手法[8]，[9]は文献[4, 5, 6, 7]と同じく，観測車両が低速であり撮影画像は高解像度かつ高フレームレートを要するなどの安定した撮影条件の下，専用の観測車両で撮影が行われる場合でのみ機能するという問題を持っている．更に，これらの従来システムでは車両中に高解像度画像の大量記録・高速処理可能な計算機が必要であり，このような高いコストと安定した撮影条件が必要な観測車両を多数用意することは現実的ではない．よって広域の地域を同時に観測することは不可能であり，各地点は不定期な少ない観測のみとなり，高頻度の地図情報の更新を実現することは困難である．これは同じ6台の高精細な撮影機器を搭載し，少数の観測車両によって街並画像データベースを構築しているGoogle streetview[10]が更新にかなりの時間(1-2年ごと)を要することからも明らかである．

さらに，この不定期で少ない観測の結果，他の走行車両や駐車車両による遮蔽や悪天候による著しい照明条件の変化などの障害のない状態で毎回撮影が行えるとは保障されない．よって街並変化検出のために適切な街並の撮影が困難となる．また，地図情報更新を目的とした画像解析手法は提案されておらず建物の構造の変化や特定の店舗のみを変化検出対象としており限定されている．つまり，更新に必要な街並変化を全て検出可能な手法は存在しない．これらの問題から，デジタル地図情報更新システム実現を目指すためには既存手法では困難である．

1.2 本論文の目的

本論文では、高頻度に広域のデジタル地図の街並情報更新を実現する画像解析要素技術の提案を行う。

また、高頻度の更新には広域を常時観測する必要があるが、従来の高価な装置と専用の観測車両による観測方法では実現が困難である。そこで、安価で汎用的な装置と一般車両による高頻度かつ広域観測が可能な地図更新システムの枠組みを提案することで、従来法の問題点だけでなく高頻度観測を実現するシステムにおける問題解決を図り、実用性の高い画像解析要素技術を実現する。

提案する街並情報更新のための画像解析要素技術と観測手法では、下記3つの目標の実現を目指す。

目標(1): 高頻度(1カ月ごと)のデジタル地図情報更新を行う。

目標(2): 現在のデジタル地図情報に含まれる街並情報を観測・更新対象とする。

目標(3): 地域に偏りなく日本全域の一般道沿いの街並情報の観測・更新を行う。

まず、目標(1)にあげたとおり、更新の頻度目標は1カ月ごととする。これは現在、道路情報の更新が最短一週間、最高3カ月で更新されているため、その平均的な更新頻度を目標とした。さらに、既存手法では検出不可能な現在のデジタル地図情報に含まれる全ての街並情報の観測・更新を行う。また、対象となる地域は日本全域となるため大量データの効率的な処理が必要となる。

目標(1)と(3)の高頻度かつ広域の観測を実現するため、提案システムでは観測車両に一般車両を利用し、観測機器には既に普及している安価な既存機器を用いる。つまり、安価なセンサを搭載した多数の一般車両が日常的に各地を走行するだけで、常時最新の街並の様子を画像として捉えることができる。これにより、現在行われている専用の観測車両を利用する従来法 [8, 9] に比べ、安価かつ容易に多数の観測車両を用意することが可能となる。

次に、目標(2)と(3)の更新処理の効率化と大量データの解析実現のため、収集される大量画像集合から事前に街並画像データベースとして記録しておき、日々新しく得られる撮影画像との比較により街並の変化(建物や看板の構造やテクス

チャの変化)を自動的に検出し、その変化情報をオペレーターに提示することで地図情報更新を行う。現在は人手により行われている画像解析を自動化することで、更新処理を格段に効率化し高頻度の地図情報更新を可能とする。

本論文で提案する地図情報更新のための画像解析要素技術は、下記の2つである。

画像解析要素技術(1): 撮影条件の変化に頑健な車載カメラ画像間の対応付け

画像解析要素技術(2): 地図情報更新のための車載カメラ画像間の比較による街並の変化検出

前者では普及型GPSで取得できる撮影時の位置情報には15m程度の誤差が生じるため、位置情報だけで入力画像と同一地点を撮影した街並情報データベース上の過去画像を選択することは不可能である。そこで、文献[8]と同様に、まずは位置情報により入力画像と同一区間を撮影した時系列画像をデータベースから選択した後、車載カメラ画像間の対応付けにより同一地点を最も近い位置から撮影した画像を抽出する必要がある。また、観測車両に一般車両を利用するため撮影画像間は撮影条件の違いが著しく、簡単なマッチングでは対応が困難である。そこで本論文では新たに撮影条件の変化に頑健な異なる日時に撮影された車載カメラ画像間の対応付け手法を提案する。

次に、対応付けられた同一地点を撮影した過去画像と入力画像を比較し、画像中から街並の変化を検出する。街並変化検出では画像中にどのような建物や看板が存在するのかという事前知識なく、異なる時刻に撮影された車載カメラ画像中から変化を検出しなければならない。そこで本論文では新たに事前知識なく異なる日時に撮影された車載カメラ画像から街並の変化検出する手法を提案する。

提案システムでは、本論文で提案する画像解析処理により得られた街並の変化情報をもとに、オペレーターが地図情報と街並情報データベースの更新を行う。このような街並情報更新の効率化により「街並変化に迅速に対応できるカーナビゲーションシステム」から「正確な街並の解析に基づいた道路・都市計画」などの社会応用まで幅広い技術の実現が期待される。

1.3 本論文の構成

本論文の構成は以下の通りである．最初に提案するシステムの枠組みについて述べ、後にシステムを実現するために必要となる画像解析処理について述べる(図2)．

第1章 序論

第2章 デジタル地図情報更新システムの枠組み

第3章 2段階マッチングによる車載カメラ画像間の対応付け

第4章 複数車載カメラ画像からの変化領域抽出による街並の変化検出

第5章 結論

第2章では、高頻度の地図情報更新を実現するため、単眼カメラ、普及型GPSという安価で一般的な装置を搭載した一般車両を利用する地図情報更新システムの枠組みを提案する．システムを実現するために必要なシステムの構成、収集した街並画像データベースの構成、地図更新の処理の流れについて述べる．さらに、処理の効率化に必要な画像解析要素技術に求められる精度と解決すべき問題点について議論する．

第3章では、地図情報更新システムの実現に必要な画像解析処理の1つ、過去画像と入力画像間の対応付け手法を提案する．普及型GPSには前後15m程度の誤差があるため位置情報だけで同一地点を撮影した画像を街並画像データベースから抽出することは不可能である．そのため過去画像と入力画像の対応付けが必要となるが、画像は低解像度で低フレームレートであり、さらに撮影条件の違いから簡単な画像のマッチングでは対応付けが困難である．提案手法では、特性の異なる段階的なマッチングによる撮影条件の違いに頑健な画像間の対応付けを手法を提案し、過去画像と入力画像の安定した対応付けを実現する．

第4章では、第3章で得られた最も近い位置から同一地点を撮影した過去画像と入力画像を比較することにより、街並の変化を検出する手法を提案する．街並は多種多様であり、画像中に存在する街並に関する事前知識は得られない．さら

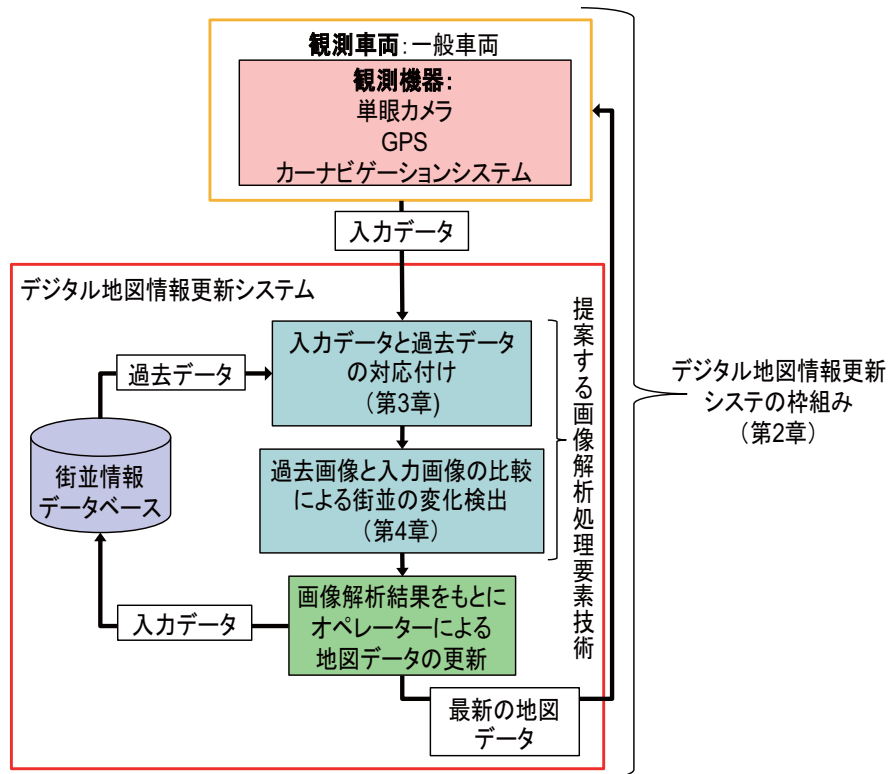


図 2 本論文の構成

異なる日時・撮影条件下で撮影された画像間は、同一物体であっても画像中の見えの違いが著しい。そこで本論文では、特徴点間マッチングと画素間マッチングにより、画像中の変化を抽出する。さらに、抽出された変化から街並の変化を識別するために、3 フレームの過去画像と入力画像の変化抽出結果をもとに多数決を行うことで街並の変化検出を実現する。

2. デジタル地図情報更新システムの枠組み

先述したとおり，これまで提案されてきた観測手法では高頻度かつ広域の街並観測は困難であり，さらに地図更新を目的とした画像解析手法ではないため，地図情報更新に必要な情報を得られないといった問題がある．そこで，本章ではまず高頻度かつ広域の街並観測を実現するデジタル地図情報更新システムの枠組みの検討を行う．

既存手法では現在の高価な装置と専用の観測車両が必要であるため，コストの問題から高頻度かつ広域の観測に必要な多数の観測車両を用意することが困難である．そこで本論文では下記の方法を用いて観測を行う．

1. 観測車両に多数の一般車両を用いることで，低コストに高頻度かつ広域の観測を実現する．
2. 観測装置はドライブレコーダーなどで普及しはじめている低解像度の単眼カメラとカーナビゲーション用の普及型 GPS とする．
3. 車両は情報収集のみを行い，観測情報は無線通信によりセンターである地図情報更新システムに送信し処理する．これにより，車両へのコスト・処理の負担を無くす．
4. 専用の観測車両ではないため，車両の走行状況には制限を設けない．

第 2.1 節ではこれらの詳細なシステム構成，地図情報更新処理の流れについて述べる．地図情報更新に必要な街並の変化情報を得るために，観測された情報と同一地点の過去に収集・蓄積された情報から構築した街並情報データベースの情報を比較する．本論文では現在は人手によって行われている，この画像解析の処理を自動化することで地図更新処理の効率化を目指す．

第 2.3 節では，処理の効率化に必要な画像解析要素技術の検討と，システム実現に必要な精度目標について考察する．

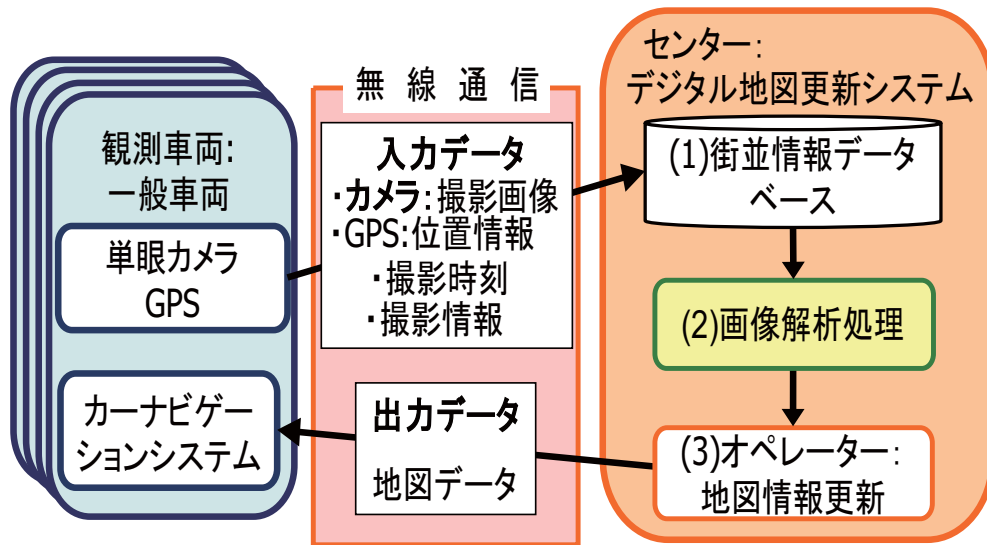


図 3 デジタル地図情報更新システム概要

2.1 システム概要

図 3 はシステム構成の簡潔な概要である。まず、観測車両の構成について述べる。

観測車両:

システムで利用する観測車両は一般車両の普通乗用車とし、その車種・車両サイズは問わない。ただし普通乗用車と大きく車高の異なる大型車両(バス・トラック)は使用しない。車両の走行条件は法定速度内とし、車両速度、運転状況に制限はなく、街中を自由に走行するものとする。さらに車両は観測のみを行い、収集されたデータは蓄積せず、無線通信により逐次センターへと転送される。

観測装置:

観測車両に搭載する装置は単眼カメラ、カーナビゲーション用の普及型 GPS、カーナビゲーションシステム、収集情報をサーバーへ転送する無線通信システム

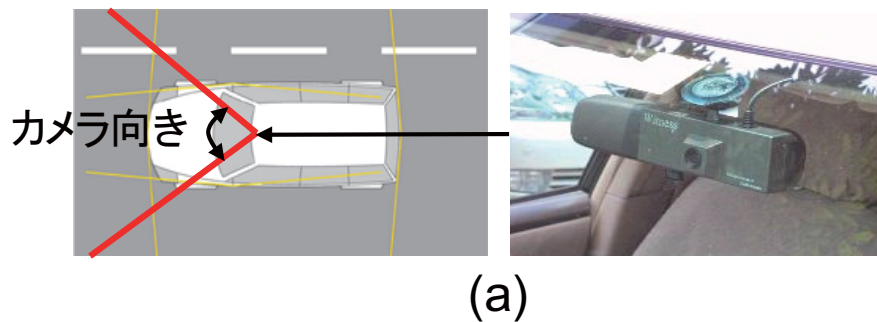


図 4 (a) 本システムのカメラ設置位置とカメラ向き (b) 本システムの撮影画像 (白い四角の領域が対象となる街並領域)

とする．下記にそれぞれの詳細について述べる．

単眼カメラ：

街並の観測を目的とし，ドライブレコーダーなどに利用されている低解像度の一般既製品の単眼カメラ1台とする．設置位置はドライブレコーダーを想定するため，バックミラー裏に設置され，そのカメラ向きはほぼ車両進行方向と同じであると限定する(図4(a))．よって，街並方向にカメラを向けて撮影した画像など上記以外のカメラ方向の画像は利用しない．ただし，車高などの影響によりカメラ設置位置や向きは完全に同一ではない．図4(b)が本システムで想定される撮

影画像の例である。カメラのパラメータは、通常のドライブレコーダーと同様とし、各カメラごとに異なり制限は設けない。ただし、そのフレームレートの下限は3-4fpsとする。カメラの内部パラメータは既製品であるため、メーカー側が測定し既知とする。内部パラメータに含まれる焦点距離は一般的なドライブレコーダーで利用されている4.3mm～8mmとし、この範囲を超える撮影画像は使用しない。

普及型 GPS：

撮影時の位置情報取得を目的とする。普及型GPSの位置情報の精度は水平方向に誤差15m程度とし、その更新は1秒間に1度とする。位置情報には緯度、経度、測位時刻、速度が含まれる。測位時刻と同じ撮影時刻の画像に位置情報が付与される。位置情報が取得できない場合はもっとも近い時刻の位置情報が付与される。

カーナビゲーションシステム：

すでに利用されている通常のHDDカーナビゲーションシステムとする。そのハードディスクの容量は現在のHDDナビの一般的な容量である60GB～80GB程度とする。収集された情報は一時的にこのHDDに蓄積され、ナビでも利用している無線通信機能を用いてセンターへ情報を送信する。その後、送信された情報は破棄する。また、サーバーで更新が行われた最新の地図情報は地図差分配信システムを利用して、自動的に更新が行われる。

無線通信機能：

観測車両は収集した情報を、地図差分配信技術でも利用されている携帯電話の無線通信を用いて、センターである地図情報更新システムに送信する。

もし撮影画像のファイルサイズ(JPEG圧縮のかかったVGAサイズ)をおよそ100KB-150KB、フレームレートを3fpsと仮定すると1秒間に最大450KBのデータ転送が発生する。現在の携帯電話の通信速度は、送信時最大384kbps実測値でおよそ200-260kbpsである。しかしながら、次世代モデルの通信速度は実験値で

提案システム		既存手法	
フレームレート	3-4fps	フレームレート	15-30fps
観測車両走行条件	制限なし	観測車両走行条件	制限あり
撮影条件	制限なし	撮影条件	制限あり
画像サイズ	制限あり (VGA サイズ)	画像サイズ	制限なし
解像度	制限あり	解像度	制限なし

表 1 既存手法と提案システムの撮影画像比較

20Mbps を記録している．そのため，上記の圧縮のかかった画像を各地点で選出された数台の車両からのみ受信するように制御すれば，センターへの情報をオンラインで転送し地図の更新を行うことは十分可能と考えられる．

入力データの構成:

次に，観測車両からセンターへ送信される入力データ (図 3) について述べる．入力データは下記とする．

- (1) 撮影画像．
- (2) 位置情報．
- (3) 撮影時刻．
- (4) 撮影情報．

(1) 撮影画像は，上記の装置の制限と無線通信を利用するため，データ転送量を考慮し画像は JPEG 圧縮のかかった低解像度画像 (最小サイズは VGA サイズ:640×480 画素) とする．画像間は非同期でそのフレームレートは無線通信を考慮した 3-4fps とする．既存手法で得られる撮影画像と比べ，本システムの撮影画像が低品質である点を表 1 にまとめる．

(2) 位置情報には緯度，経度，測位時刻，速度が含まれる．位置情報取得時刻と撮影時刻が同じものに位置情報が割り当てられる．

提案システム		既存手法	
使用車両	一般車両	使用車両	専用車両
車両台数	多数	車両台数	少数
車両速度	制限なし(法定速度内)	車両速度	低速(一定速度が望ましい)
運転状況	制限なし	運転状況	制限あり(車線変更は行わず、 同じ車線を走行する)
使用カメラ	低解像度の単眼カメラ1台	使用カメラ	高解像度の全方位カメラ または 複数の高解像度単眼カメラ
GPS	普及型GPS	GPS	高精度GPSまたは普及型GPS

表2 提案システムと既存手法の観測車両構成比較

(3) 撮影時刻は画像撮影時刻とする。

(4) 撮影情報は車高，撮影フレームレート，カメラパラメータの3つとする。

カメラパラメータはカメラの内部パラメータとし，カメラ設置時に得られた値とする。この内部パラメータには画像中心座標，焦点距離が含まれる。車高情報はおよそのカメラ設置位置を把握するため取得する。

表2に，既存手法と提案システムの情報収集を行う観測車両とその構成の比較を示す。

センターの構成：

センターである地図情報更新システムの構成について述べる。センターは過去データが蓄積されている街並情報データベース，観測車両からデータを受信し解析する画像解析処理部，地図情報と街並情報データベースの更新を行うオペレーターで構成される。下記に各構成について述べる。

街並情報データベース：

街並情報データベースは位置情報を元に記録され，1つの位置情報に対して街並を観測した撮影画像系列，撮影時に得られた入力情報，デジタル地図情報で構成される。1つの撮影画像系列は同一の車両により撮影されたもので，1地点に

つき3シーケンス以上の異なる日時に得られた入力データが蓄積される。(なお、街並情報データベース構築方法については第2.2節において述べる。)

画像解析処理部：

画像解析処理部では地図情報更新のために、大量に得られる入力データの解析を自動的に行う。画像解析は以下の手順で行われる。

- (1) 位置情報をもとに入力データと同一の位置情報を持つ過去データを街並情報データベースから抽出する。
- (2) 入力画像と過去画像の対応付けにより、同一地点を最も近い位置から撮影した画像を選択する。
- (3) 異なる時間に同一地点を撮影した画像列間を比較することで街並の変化検出を行う。
- (4) 検出した変化が生じている入力データをオペレーターに提示する。

上記の画像解析により、地図情報更新に必要な街並の変化を自動的に検出することで大量に得られる画像を効率的に処理し、高頻度の地図更新を実現する。

オペレーター：

オペレーターは画像解析処理により変化が検出された画像を確認し、変化情報に合わせて地図情報の編集と街並情報データベースの更新を行う。現在、地図更新には年間延べ約28万人¹1日に1000人強の調査員が携わっている。そこで、オペレーターは全国で1000人とし、日に8時間、週5日をデータ確認と地図更新に費やす時間とする。

システムでは日本全域の一般道沿いの街並情報の観測、解析、地図更新の全ての処理が1カ月ごとに行われることを目的としている。

¹株式会社ゼンリン <http://www.zenrin.co.jp/product/gis/znettown/znettown.html> 参照

2.2 システム構成と処理の流れ

本節では街並情報データベース構築方法と地図情報更新システムの処理の流れについて述べる。まず、現在の地図会社で行われている地図情報更新処理の流れを図5に示す。図5の街並情報データベースは街並を撮影した画像と撮影時の位置情報および、それらに対応付けられたデジタル地図情報により構築されている。まず、観測により新たにデータが得られた際、入力データと同一の位置情報をもつ過去画像群が街並情報データベースから抽出され、オペレーターに提示される。位置情報には誤差が含まれる場合があるため、入力データと同一地点をもつとも

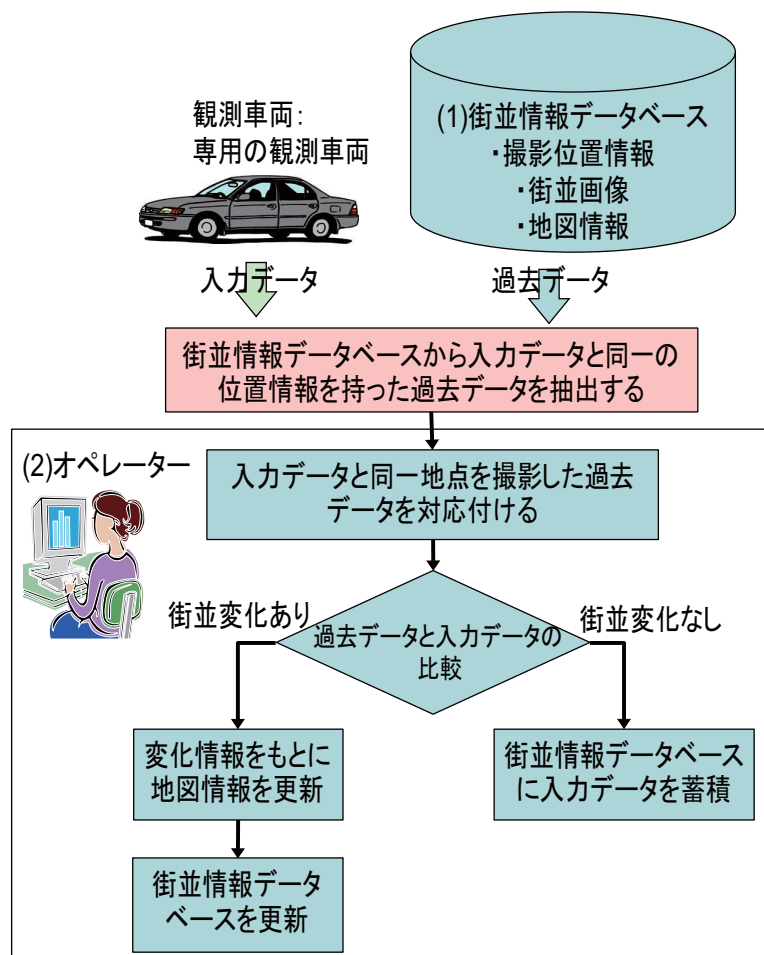


図5 地図会社で行われているデジタル地図情報更新処理の流れ

近い位置を撮影した過去データを選択する．そして、これらと比較することで街並情報の変化を検出する．次に、検出された変化情報をもとに地図情報の編集を行い地図情報の更新が行われる．地図の更新作業では地図情報は変化地点だけではなく、その周辺も画像上で確認する．さらに、駅や大型施設などの大きな変化が生じていた場合は再度調査を行い、不整合がでないように合わせて修正が行われる．街並情報データベースは変化が検出された場合は、過去データと入力データの入れ替えが行われる．変化が無い場合であっても、収集されたデータは地図情報と関連づけて街並情報データベースに蓄積される．このように現在の更新処理はほぼ全てを人手によって行われているが、対象範囲が広くとも、半年または1年ごとの観測・更新頻度であれば提案システムと同数のオペレーター（毎日1000人）での更新が可能である．

しかしながら、本研究が目指す1カ月ごとの更新では、データ量の問題から現在の方法では対応できない．また、地図情報の正確性を保つためには変化を確実に見つけることが重要であり、特に観測回数が少ない場合は見逃しの影響が大きい．処理の効率化のためにこれらの処理を全て画像解析技術で自動化するには、下記の問題を解決しなければならない．

- (1) 必ずしも比較に最適な画像が得られない．(他車両による遮蔽などが生じる)
- (2) 街並の形状やテクスチャは多種多様である．
- (3) どのような変化が生じているか不明である．

現状の画像解析手法ではこれらの問題を解決し、街並情報の変化を漏れなく検出することは不可能である．

そこで、本研究では1カ月以内の更新を目指し、観測頻度の増加と画像解析技術により地図情報処理の効率化を目指す．観測頻度の増加により見逃し率の影響を抑え、さらに画像解析技術により人手による更新処理コストを減少させる．

次に、本論文で想定する地図更新システムで行われる地図更新処理の流れ全体を図6に示す．サーバーである地図情報更新システムは、(1)街並情報データベース、(2)画像解析処理、(3)オペレーターの3つで構成される．本システムでは処理の効率化のため、街並の変化情報の検出を画像解析技術により自動化すること

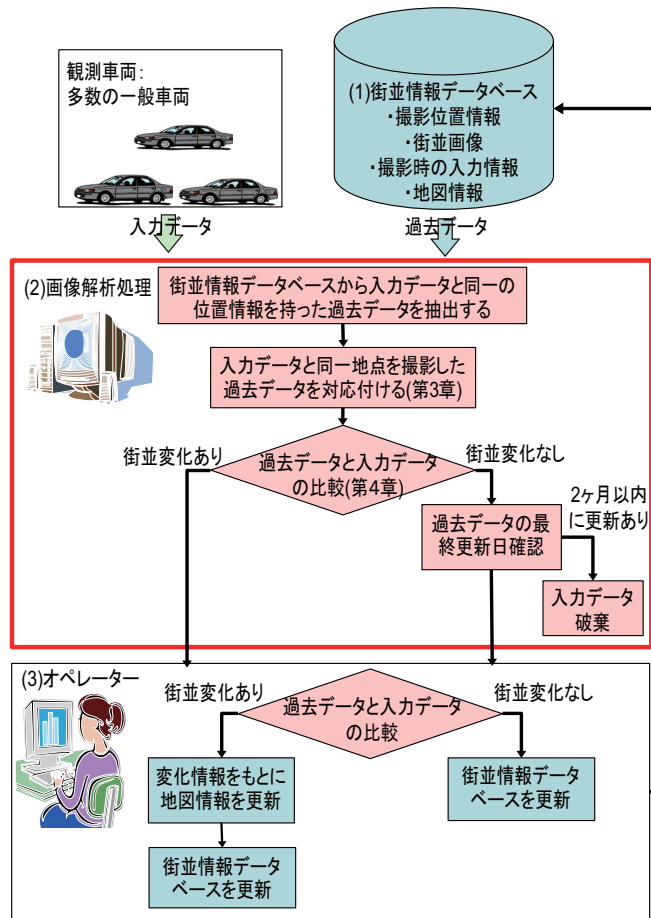


図 6 提案するデジタル地図情報更新システムの処理の流れ

で人手による画像解析のコストを減少させる。画像解析処理では、入力データと同一位置情報をもつ過去データをデータベースから抽出する。位置情報は誤差を持つため、入力データの街並画像と過去データの街並画像を対応付けることで、最も近い位置から同一地点を撮影した過去データを選択する。過去データと入力データの比較により、街並の変化を検出する。変化が検出された場合と2カ月間一度も更新が行われていない場合に次のオペレーターによる確認を行う。このように、オペレーターによる確認を行うデータ量を減少させることで更新処理の効率化を実現し、さらに確認を行うことで見逃しを無くし地図情報の正確性を保ちながら高頻度の地図情報更新を実現する。また、大量の画像を効率的に処理する

ため複数台のサーバーによる並列処理とする。日本全国から観測情報を受信し、効率的に処理するため約 330 台 (1 都道府県につき 6-7 台) のサーバーとする。また、サーバーはマルチスレッド処理が可能な 4 コア以上の CPU 搭載のものとし、1 台の値段はおよそ 4 コアで約 50 万円、最新の 8 コアで約 180 万円である。画像の全てを人手による処理で解析するコストと比較すれば、十分に実現可能である。次の (3) オペレーター処理では、画像解析処理の結果を確認し変化が検出された場合は現状と同じ地図情報と街並情報データベースの更新を行う。また変化が検出されなかった場合は、入力データを街並情報データベースに蓄積、またはもっとも古いデータとの入れ替えを行う。街並情報データベースの更新・蓄積は必ず人手による確認が行われたデータのみに限られ、データベース中の同一地点のデータ間には街並の変化は生じていないデータとする。また、オペレーターは現在の地図更新処理に携わっている同数の 1000 人とする。

次に図 7 に街並情報データベース更新の流れを示す。入力データ取得基準は撮影時刻と位置情報を基準とする。夜間のデータは使用せず、さらに逆光など照明条件を考慮し撮影時刻 11 時から 15 時の撮影画像とする。次に位置情報にはおおよその車両速度も含まれる。フレームレートが低いため、車両速度が早い観測車両から得られた撮影画像では時系列画像間の重なり合う領域が小さく、対応付けが困難となる。よってこの車両速度をもとに、公道における一般車両の最高時速 60km 以上の車両速度の車両から観測された画像も使用しない。さらに、位置情報と撮影時刻が把握できれば、天候情報を利用し照明条件の変化が激しくワイパーが画像中に写りこむ恐れのある雨天時の画像も省くことも可能となる。今回は、雨量が少量の場合は観測するが激しい降雨時は、対象画像から省く。

データベースに蓄積される同一地点の画像は、過去データにバリエーションを持たせることで画像間のマッチングを容易にする。さらに 3 フレームの画像間との比較結果を利用し多数決的に街並の変化を検出するため 1 地点につき 3 つの時系列画像とする。その数は 1 方向につき 3 シーケンス、計 6 シーケンス以上とする。1 シーケンスは同一車両で撮影された画像で構成され、位置情報の前後 15m の誤差を考慮し、60m 以上の区間を撮影した時系列画像とする。

次に入力データが過去データとして蓄積される基準について述べる。下記の条

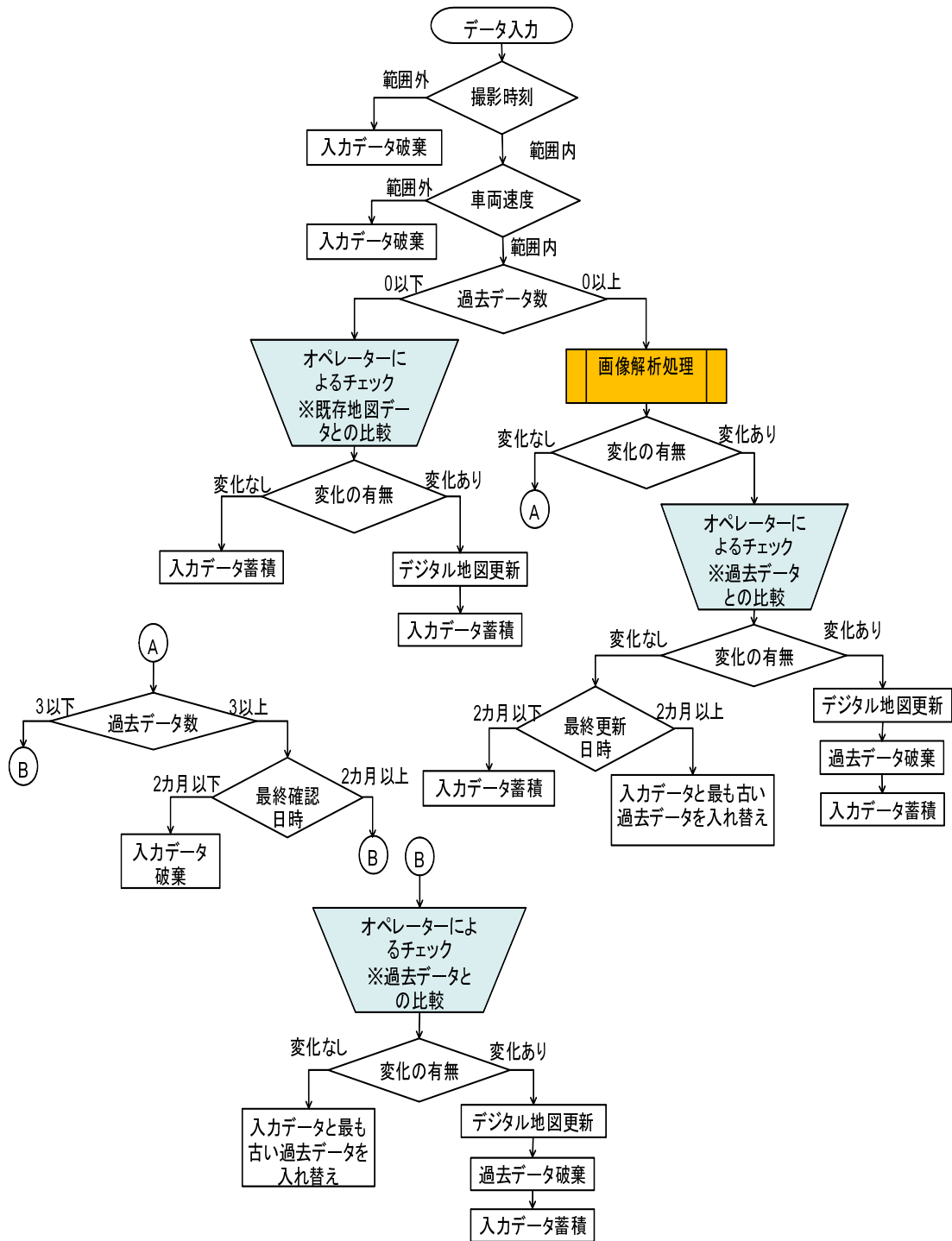


図 7 街並情報データベース更新の流れ

件時に入力データの蓄積または更新が発生する。

- (1) データベース中に、同一地点の画像が存在しない。
- (2) データベースに蓄積されている過去データが規定シーケンス数以下である。
- (3) 最後に過去データの更新および蓄積がされてから、2 カ月以上経過している。
- (4) 変化が検出され地図更新が行われた。

(1) では、データベース中にデータが存在しないためオペレーターによる、既存のデジタル地図情報との整合チェックの後に蓄積が行われる。(2) では規定のシーケンス数に達するまで変化検出処理後、変化の有無に関わらず蓄積が行われる。(3) では時間経過により形状が変化する街路樹などの問題から、撮影日時が近い画像間でマッチングを行うことが望ましい。そこで人手による最終確認日から2 カ月以上経過している場合は、もっとも古い過去データと入力データを入れ替えるデータベースの更新処理を行う。(4) では、地図更新が行われた際は過去データを全て破棄し入力データの蓄積を行う。

また、図7に示したとおり、地図更新およびデータベースの更新は必ずオペレーターによる確認後に行われる。さらに、2 カ月間データの確認が行われていない地点の確認を行うことで画像解析処理による変化検出漏れを防ぐ。

2.3 システム実現に必要な画像解析手法とその目標

本節ではシステム実現に必要な画像解析処理手法とその精度目標の検討を行う。現在の地図更新は、収集した画像を目視により確認することで街並に変化のある地点を探索し、得られた変化情報を元に地図の更新を行っている。

現在の日本の一般道の総延長は11,962,166km(平成20年4月1日)であり、1地点の平均フレーム数(20フレーム)から、一般道全域を一度観測すると最低約800,000,000フレーム(進行方向違いを含む)の大量の画像の確認が必要となる。

そのため、常時観測により大量に観測情報を収集することが可能となっても、人手による目視・確認作業が増加し、高頻度の更新は実現できない。そこで、システムでは画像解析処理により、同一地点を撮影したデータベース上に存在する過去画像と、新しく得られた入力画像を比較することで、画像中の変化を自動的に検出し、確認作業コストの大幅な低減を図る。

この画像解析には下記の2つの処理が必要となる。

(1) 撮影条件の変化に頑健な車載カメラ画像間の対応付け

画像間の比較を行うには、まず新しく得られた入力画像と同一地点をもっとも近い位置から撮影した過去画像をデータベースから抽出する必要がある。GPSの位置情報をもとにおおよそ同一区間を撮影した画像は抽出可能ではあるが、位置情報には前後15m程度の誤差があるため、完全に画像を対応付けることは不可能である。そのため、文献[8]と同様に過去画像と入力画像をマッチングすることで、もっとも近い位置から同一地点を撮影した画像を選択する。

(2) 地図情報更新のための車載カメラ画像間の比較による街並の変化検出

地図情報更新を行うために、異なる時刻に同一地点を撮影した画像間の比較による街並情報の変化を検出する必要がある。地図情報の更新に必要な街並情報の変化とは、固定された人工物である建物と店舗に付随する看板の下記の変化である。

店舗：構造・位置・テクスチャの変化

一般建物：構造・位置の変化

看板：構造・位置・テクスチャの変化

ただし，表示の変化する電光掲示板のテクスチャおよび看板のテクスチャの部分的な変化(例:ガソリンスタンドの値段表示)は必要としない．地図情報更新のためには，これらの全ての変化を検出する必要がある．しかしながら，観察対象となる人工物は多種多様であり，全てを網羅して事前に把握しておくことは不可能である．さらに，対象となる街並情報以外の変化や撮影条件の違いから生じる見えの変化も画像中には生じる．そこで，本論文では街並情報の変化を漏らさず検出するために画像中に生じる変化を下記のとおり「変化が生じる時間」により識別し検出を行う．

常時変化：歩行者や車両といった移動物体，形状が変化しやすい街路樹・のぼり旗

短期変化：一時的に存在する駐車車両

長期変化：固定された人工物の構造とテクスチャの変化

この長期変化には，画像中の大半を占めるような大きな建物の変化といった大域変化と看板やテクスチャや小さな建物の変化といった局所変化が含まれ，これらの変化を漏れなく検出することが，この変化検出の目的となる．

(3) 画像解析処理要素技術の精度目標の検討

1カ月ごとの地図情報更新を実現するために，提案する画像解析要素技術ではオペレーターによる確認処理が必要な画像数を低減させることが目標となる．そこで，1カ月の間でオペレーターによる確認(入力データと過去データの比較)が必要な画像数の目標を定める．日本全域のフレーム(最低数の約800,000,000フレームと仮定)を人手により全て確認した場合，1人が1日で28800フレーム(60secで60フレーム処理すると仮定)を確認できたと仮定しても，1000人で28日間の時間が必要となり，1カ月に1度の更新は不可能である．そこでシステムでは，1カ月で全画像中の30%に街並の変化が生じていると仮定し，人手による確認が必要

な画像数を全画像中の40%以下とする。40%以下であれば1000人で11日間での確認が可能となる。次に、画像解析処理の精度目標について述べる。システムの精度評価として街並変化検出漏れ率と街並変化誤検出率の2つを設定する。検出漏れ率とは、検出すべき街並の変化として定義した変化を「変化なし」と判断した場合とする。そして、誤検出率は定義した街並変化が生じていない画像に対して「変化あり」と判断した場合とする。地図情報の更新を目的としているため、街並変化検出漏れ率をできるだけ低下させることが重要となる。そこで、本論文では街並変化検出漏れ率の最低精度目標を10%以下とする。

よって全画像中の30%の画像に街並変化が生じていると仮定した時、誤検出率を50%、検出漏れ率が10%以下であれば確認が必要な画像を全画像中の40%とすることが可能となる。今回はこれを画像解析要素技術の最低精度目標とする。

3. 撮影条件の変化に頑健な2段階マッチングによる車載カメラ画像間の対応付け

3.1 はじめに

先述の通り，GPSの誤差(水平方向15m)のために位置情報だけでは同一地点を観測した画像間を対応付けることは不可能である．よってこのシステムでも文献[8]，[11]と同様に，まず初めに位置情報地点 P で観測した入力画像列と位置情報 $P \pm 15m$ で観測された過去画像列とを比較することで，過去画像列から入力画像と最も近い位置から同一地点を撮影した画像を抽出する必要がある．また位置情報の履歴から車両の進行方向が分かるため，同じ進行方向を走行している車両から撮影された入力画像と過去画像を比較する．

このとき，たとえ看板の変化など街並に変化が生じていても，過去画像と入力画像のマッチングによる抽出は成功しなければいけない．さらに，先行文献[8]に比べて提案システムの撮影条件は一般車両や安価なカメラを用いるため，非同期・低解像度・低フレームレートであり，さらに撮影条件も安定しない．そのため，画像間は同一地点を撮影した画像であっても，撮影位置は多く異なり，さらに撮影条件の違いから見えの変化も生じ，テンプレートマッチングのような簡単な画像間のマッチングを適用することは困難である．それゆえ，本章では地図更新システム実現のために，新たな車載画像間の対応付けの手法を提案する．以下に，対処が必要な問題を整理する．

- (a) 多数の一般車両によって画像を撮影しているため，画像間は非同期で車両速度も異なり，無線通信のためにフレームレートも低い．これらの特徴の結果，画像の撮影位置は大きく異なり，画像間の見えは大きく異なる．
- (b) たとえ対象となる街並(建物や看板など)に変化があっても，同一地点を観測した画像間の対応付けを成功させなければいけない．
- (c) 街路樹や歩行者，他の走行車両など変化検出対象ではない多数の変化や遮蔽が生じるが，画像間の対応付けを成功させなければいけない．

- (d) 異なる日時の撮影画像間には照明条件の変化が生じる。
- (e) 撮影画像は一般的なカメラで撮影され、またその設置位置は図4に示すように街並方向ではなく前方へ向けて設置されており対応付けおよび変化検出に適した画像は得られない。
- (f) 道路環境(道路の段差,カーブ走行)や運転状況(左・右折,ブレーキ)により,撮影画像に上下および左右方向の揺れの影響が生じる。
- (g) カメラの向きや走行車線は車両ごとに異なるため,画像中の街並の見えはたとえ同じ地点でも大きく異なる。

(b),(c),(d)が従来法でも取り上げられた問題で,(a),(e),(f),(g)が本研究が想定するシステムの枠組み独自の問題である。加えて,(e)の問題がなくマッチングに最適なカメラ方向の撮影画像が得られたとしても,解決しなければいけない問題点(b),(f),(g)は同じである。本研究ではこれらすべての問題を解決する画像対応付けとして,特性の異なる段階的なマッチングを提案する。

3.1.1 研究目的

想定するシステムでは撮影間隔が長く(3-4fps),さらに車両ごとに車速と撮影間隔が異なるため,同じ道路上でも異なる車両間で撮影タイミングを同期させることは不可能である。また撮影間隔が長くとも,図8(a)に示すように連続撮影画像中には同じ街並が写っている重なりがある。一方で,過去画像と入力画像を比較して「同じ地点で撮影された画像間」を対応付けるという目的のためには,また過去・入力画像列の間で最も近い場所で撮影された画像を一対一で比較しても(図8(a)),撮影間隔が長いためにそれらの重なりあう領域は小さく,重なり領域の単純な探索によるマッチングは困難である。また,大量の画像を処理する必要があるため,位置情報を考慮した前後30m以上の区間の過去画像系列と1フレームの入力画像をマッチングするような,1対多数の画像間のマッチング行うことは計算コストの問題から効率的ではない。また,対応付けは街並の変化が生じて

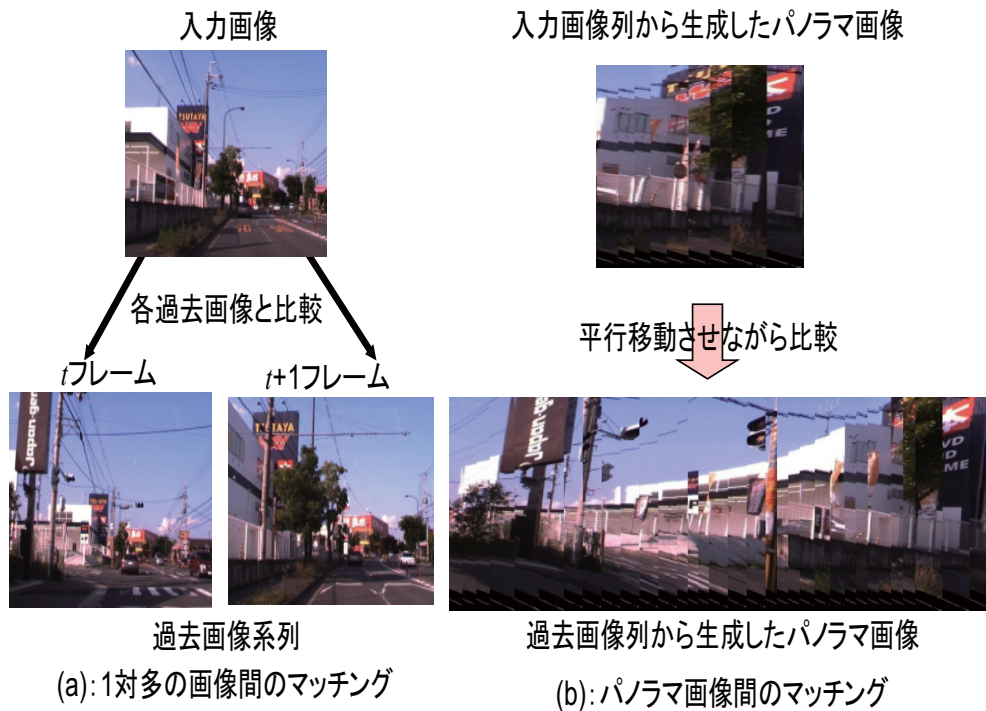


図 8 1 対多の画像間マッチングとパノラマ画像間マッチング

いても成功させる必要がある。さらに、街並以外の変化や遮蔽も生じるため、重なり領域が小さいという問題がマッチングにおよぼす影響は大きい。

そこで既存手法ではマッチングが困難な撮影条件の異なる画像間の安定したマッチングを目指し、提案手法では2段階のマッチングにより画像間の対応付けを行う。

第1段階では撮影位置の違いから生じる重なり領域が小さいという問題を解決するため、過去・入力画像ともに連続画像同士から重なりのない領域のみを切り出し連結させた幅の広いパノラマ画像(図8(b))を用意し、このパノラマ画像間で比較を行うことで、重なり合う領域を探索する。これにより重なり幅の大きい画像同士の安定したマッチングが可能になる。また、重なりあう領域増やすことでマッチングを行う領域が増やし、部分的な遮蔽や街並変化の影響を抑える

第2段階では、第1段階で対応付けられた入力画像1フレームと過去画像の前後数フレームから特徴点の対応付けにより、もっとも近い位置から撮影された画

像を選択する．第 1 段階のマッチングで探索範囲を絞り込み，第 2 段階でその探索範囲内においてより密なマッチングを行うことで，計算コストの低減と精度向上を目指す．

3.1.2 関連研究と位置付け

提案手法と同じパノラマ画像を用いたマッチングの手法は，パノラマ画像作成方法により大きく分けて 2 つに分類される．

- (1) 文献 [12] では街並と平行な直線上を前進しながら一定の密な間隔で停止し，撮影された時系列全方位画像からパノラマ画像を作成している．
- (2) 文献 [13], [14], [15], [16] では等尺のラインスキャン画像をつなぎ合わせてパノラマ画像を生成している．

文献 [13] ではロボットに搭載されたラインスキャンカメラにより撮影を行っている．また，文献 [14], [15], [16] では車載カメラにより撮影を行っている．文献 [16] ではパノラマ画像上の物体の奥行きを推定することで，画像のブレを補正する手法が提案されている．また，文献 [15] では，レーザレンジセンサによって作成された CAD モデル上にラインスキャン画像をテクスチャとしてその側面に投影することで，3 次元都市モデルを生成している．これらの手法であれば正確で詳細なパノラマ画像や都市モデルを生成することができる．しかし，これらの手法では本システムで想定している街並とは異なる方向に向けられた普及型カメラの利用や，異なる撮影条件，例えば異なる車速と異なるカメラ向きなどには対応できないという問題がある．

さらにパノラマ画像と同じように時系列画像をつなぎあわせて観測シーンを解析する手法に，EPI (Epipolar Plane Image) 解析 ([17], [18]) がある．

EPI 解析とは 1 または 2 次元の濃淡画像をを時間軸に沿って並べて作成する時空間画像を解析する手法である．カメラが等速に水平方向に移動したとき得られる時空間画像では，画像間の対応点は連続したエッジを描きの画像にはさまざまな傾き成分を持ったエッジが含まれる．この傾きの差は，カメラの水平位置の違いによる視差から生じ，EPI 中のエッジの傾きがそれを描いた点の奥行き位置に対

応することになる．よってEPIからのエッジ検出により，対象物体のカメラからの奥行き情報を取得することで3次元情報の獲得が可能になる．このEPI解析手法の中で，車載カメラ画像を用いた手法が提案されている．文献[19]ではEPI画像を利用した3次元復元と任意視点画像生成が行われている．また文献[20]では多数のカメラで撮影されたEPI画像のマッチングから，任意視点画像生成を行う手法が提案されている．しかし，提案手法で生じる撮影条件の変化はEPI解析の失敗の原因となり，その影響は文献[19]でも確認されている．

提案手法では街並に変化や遮蔽，さらに上記でも問題となる撮影条件の違いが生じていても安定して対応付けを行うことを目的とする．パノラマ画像を用いることで，探索を行う画像間の重なり領域を増やし街並の変化や遮蔽に対して頑健なマッチングを目指す．さらに撮影条件の変化に対応するために下記の手法を取り入れる．

- (1) 撮影画像をカメラ向きに基づいて補正する．
- (2) カメラの水平方向の移動量をもとに短冊画像の切り出し位置を決定することで，車両の左右揺れを補正する．
- (3) カメラの水平方向の移動量をもとに各撮影画像から，画像の一部を切り出し短冊状の画像（以下，短冊画像と呼ぶ）を抽出することでパノラマ画像の幅を動的に調節する．
- (4) 街並に対して並行のパノラマ画像平面に投影することで，マッチングに適したパノラマ画像を作成する．
- (5) カメラの垂直方向の移動量をもとに，車両の上下揺れを補正する．
- (6) 照明条件の変化に頑健なエッジをマッチングの際の特徴量に用いる．
- (7) 街並の変化に対応するため，細かなテクスチャのエッジは除去する．
- (8) 撮影位置の違いによるスケール変化に対応するため，パノラマ画像の縮小・拡大を行う．

このような画像補正と調節により，異なる時間，撮影条件で撮影された画像から相似したパノラマ画像を作成し，画像間で重なりあう領域を容易に探索することを可能とする．

次に，2段階目のマッチングで行う異なるカメラで撮影された画像間の対応を特徴点の対応付けによって行う手法も提案されている．まずは，画像から対応付ける局所特徴量を抽出する必要がある．局所特徴量を検出する手法として Harris Corner Detector[21] などが知られているが，提案手法では物体認識やイメージモザイク等に利用されている，Scale-Invariant Feature Transform(SIFT)[22] を利用する．SIFT は今回の撮影画像に生じている回転，照明条件，スケール変化に頑健な特徴点抽出手法であるが，計算コストが非常に高く多数の画像処理が必要な場合には不向きである．さらに，画像には駐車車両などによる遮蔽など部分的な変化が多く生じるため，局所特徴量の対応付けはこれらの影響を受けやすい．提案手法では，段階的なマッチングによって探索範囲を限定することで計算コスト削減と精度向上を目指す．

3.1.3 提案手法の流れ

提案手法の流れを図9に示す．

まず，マッチングの注目地点となる画像（以下，起点画像とする）を過去画像群から決定する．起点画像は入力画像から同じ位置情報を持つ画像とする．入力画像の起点画像はマッチングの注目地点の位置情報を持つ画像とする．

マッチングにおける特徴量は，撮影位置を特定しやすく一時的な変化の少ない走行車線に面している建物を含む街並領域から抽出する．しかし提案システムではカメラは車両進行方向に向けられているため，図4(a)の撮影画像例からもわかるように街並領域は画像中の左側の領域に歪んで写っている．そこで提案手法では，図4(a)の白い四角に囲まれた領域（撮影画像の左側）から特徴量を抽出する．一方通行の道路の場合は画像の右側からも抽出される．

第1段階のパノラマ画像間のマッチングにおいて利用するフレームは，位置情報の前後15m程度の誤差を考慮し起点画像の前後フレームとする．利用するフレーム数はパノラマ画像幅により決定される．パノラマ画像幅については3.2節

にて述べる．撮影画像の左側の領域から画像の一部を切り出し，それらを時系列順に繋ぎ合わせることでパノラマ画像を作成する．さらにパノラマ画像を作成する際に，さまざまな画像補正を行うことで異なる撮影条件下で撮影された画像から相似した見えのパノラマ画像を作成する．第2段階では，このパノラマ画像間のマッチングで得られた画像間の対応関係を利用し，入力画像と対応付けられた過去画像前後3フレームの計7フレームから特徴点の対応付けによって，もっとも近い位置から撮影された画像を抽出する．

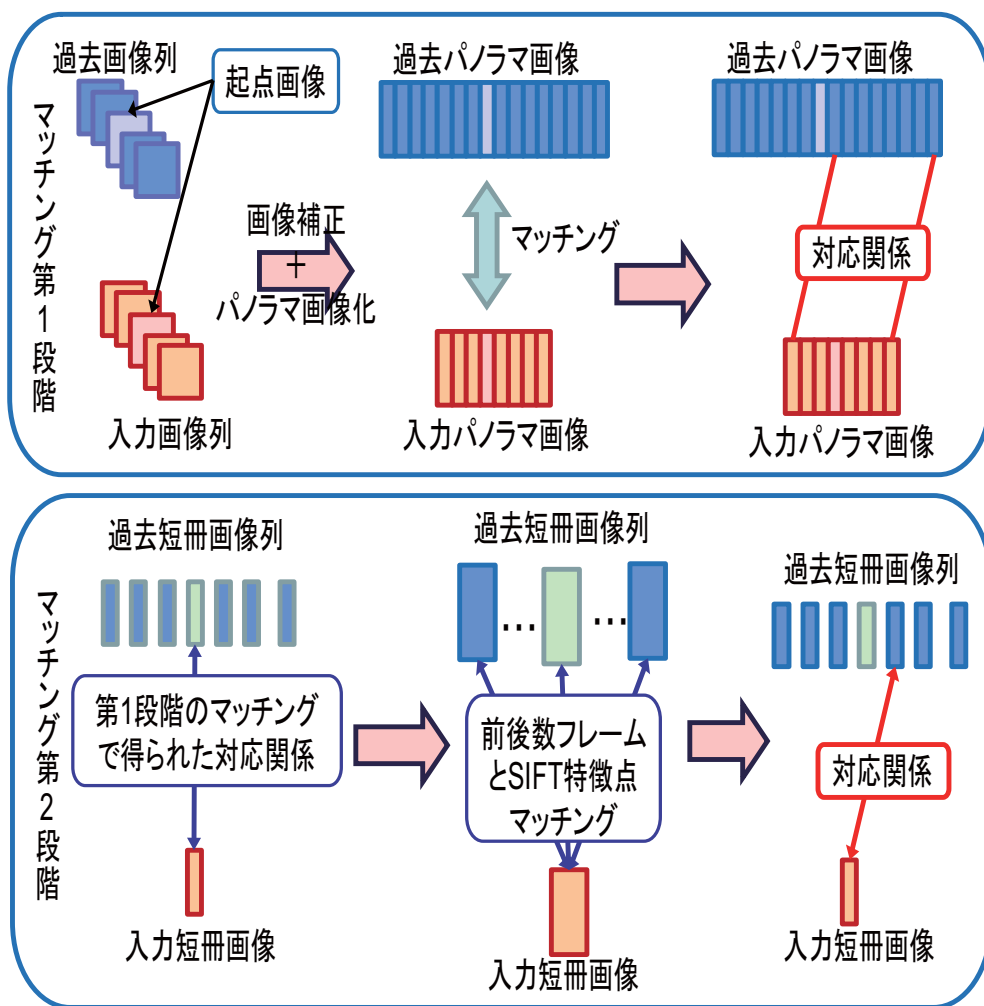


図 9 提案手法の流れ

3.2 第1段階：パノラマ画像による撮影条件の変化に頑健な車載カメラ画像間の対応付け

低フレームレートと車両速度の違いから生じる撮影位置の違いに対応するため、提案手法では長さの異なるパノラマ画像間で、平行移動しながらマッチングを行う。

パノラマ画像の幅は正しい対応付けの結果が得られるように、以下のように定めた。入力パノラマ画像は入力撮影画像群からGPSの位置情報をもとに、起点画像をパノラマ画像のもっとも左側に位置するとし、そこから、建物の平均的な大きさなどを考慮した40-50mの区間をパノラマ画像に投影するため、パノラマ画像幅900pixelをとした。さらにパノラマ画像は時系列順に右側につなぎ合わせることで生成する。過去パノラマ画像の起点画像は入力パノラマ画像中の起点画像とGPSの位置情報が完全に一致する画像とする。GPSの位置情報の誤差(約15m)を考慮し、過去パノラマ画像の幅は入力パノラマ画像の幅よりも長くする必要があり、また、上記の理由から過去パノラマ画像幅は最小で±15mの区間が投影されている必要がある。過去パノラマ画像の幅は安定したマッチングのために、80m程度の区間を投影可能な1800pixelとする。

しかしながら、もし車両の速度が非常に低速な場合、生成される短冊画像の幅は短くなってしまふ。そこでフレーム間の移動量から判断し、停車時の画像は利用しないこととする。

さらに、撮影条件の違いにより画像間は見えの違いが生じる。提案手法ではこれらの問題に対し画像補正を行うことで相似した見えのパノラマ画像を生成し、マッチングの精度向上を目指す。本章では、相似した見えのパノラマ画像生成方法とマッチング方法について述べる。

3.2.1 撮影条件の変化によって生じる問題点

過去、入力パノラマ画像のマッチングの際に問題点となる、多様な要因(車両・カメラ・撮影時刻の違いなど)により生じる撮影条件の変化に対応する必要がある。以下で、それぞれの撮影条件の変化がパノラマ画像にどのような影響を及ぼ



図 10 異なるカメラ向き，車線から撮影された同一地点の画像

すのか，またその対処方法について述べる．

(1) カメラの回転角度の変化

カメラ設置方向が各車両ごとに異なるため，たとえ同一地点で撮影された画像間でも図 10(a) と (c) のように見えが大きく異なり，これらの画像から生成されたパノラマ画像の見えにも違いが生じる．そのため車両進行方向を基準としてカメラ光軸方向の回転角を求め，撮影画像から光軸と車両進行方向が一致した仮想画像を作成する．この仮想画像から生成されるパノラマ画像を過去・入力画像間の対応付けに利用することにより，対応付けの安定性を向上できる．文献 [13] でも，道沿いを観測して得られる画像からパノラマ画像を生成している．しかしながら，このシステムでは固定カメラで得られる高いフレームレートのラインスキャン画像をつなぎあわせているため，容易に歪みのないパノラマ画像が生成できる．一方，提案システムでは低いフレームレートで撮影された普及型のカメラで撮影された画像からは，対応付けに適したパノラマ画像は得られない．

(2) 車載カメラの左右揺れ

車両の左折・右折やカーブ走行により車両に左右揺れが生じた画像から短冊画像の切り出し位置を固定値にして、パノラマ画像を作成すると、パノラマ画像に視差の違いが生じる。そこでパノラマ画像を作成するための短冊画像切り出し位置をフレーム間の移動量をもとに決定する。

(3) 観測車両の車両速度の違い

観測車両の車速は車両ごとに異なり、さらに専用の観測車両とは異なり高速かつ一定ではない。そのため、同じフレーム数でパノラマ画像を作成しても時系列のフレーム間の重なりは異なる。提案手法では水平方向の移動量をもとに撮影画像から切り出す画像の幅を決定することで、車両速度の違いに対応する。

(4) 車両進行方向にカメラが向けられている

カメラが車両進行方向に向けて設置されているため、街並は撮影画像の両端にのみ存在し、マッチングに適した画像は得られない。そこで、パノラマ画像作成時に時系列順に画像つなぎ合わせるのではなく街並に対して平行に存在する仮想パノラマ画像平面に時系列順に投影していくことでパノラマ画像を作成し、マッチングに適したパノラマ画像を作成する。

(5) 車載カメラの上下揺れ

撮影時に道路の段差などにより車両に上下揺れが生じた画像からパノラマ画像を生成すると、上下方向に歪みが生じる。同じ場所を通過する全車両にまったく同様の揺れが生じるなら、対応付けの精度への影響はない。しかし、車両の揺れは道路環境のみが原因とは限らず、走行状態によっても生じる。そこで特徴点追跡の結果をもとに車両の上下の移動量を推定し画像補正を行う。

(6) 照明条件の変化

図 10 の (a) と (b) は同じ日の異なる時刻に撮影された画像であるが、照明条件が大きく異なる。このように、たとえ同じ日であり、撮影時刻が規定範囲内であっても照明条件は著しく変化する。観測車両から送られてくる撮影情報には撮影時間も含まれるため、およそ同じ時間帯に撮影された画像系列間で対応付けを行うことは可能であるが、天候条件の違いもあるため画像には著しい輝度変化が生じ、画像間の比較が難しくなる。文献 [8] では、画像の色情報を特徴量として Dynamic Programming (DP) matching ([23],[24]) による画像の対応付けを行っている。しかし、色情報は照明条件の変化に影響を受けやすく、今回のように照明条件が激しく変化する場合には特徴量として不向きである。そこで、照明条件の変化に頑健なエッジを特徴量としてマッチングを行う。

(7) 複雑なテクスチャや形状の微細な変化

対応付けでは、たとえ街並の変化が生じていてもマッチングを成功させる必要がある。

よって、時系列画像間でも形状が変化するのぼり旗や街路樹や、看板などの複雑なテクスチャを持つ人工物の細かな変化は、「過去画像と入力画像の対応付け」において安定したマッチングのノイズとなる。そのため、これらの領域から抽出される細かなエッジは除去する。このように生成されたエッジ画像は時間経過による変化に頑健な特徴量（建物の輪郭線など）を持ち、正確なマッチングを可能とする。

(8) スケールの変化

車両の走行車線が異なれば、カメラと撮影対象との距離の変化により図 10(a) と (b) のように画像中の構造物の見えが異なる。このような画像中の構造物のサイズの違いに対応するため、入力パノラマ画像を様々な倍率に拡大・縮小し、過去パノラマ画像とマッチングを行なう。

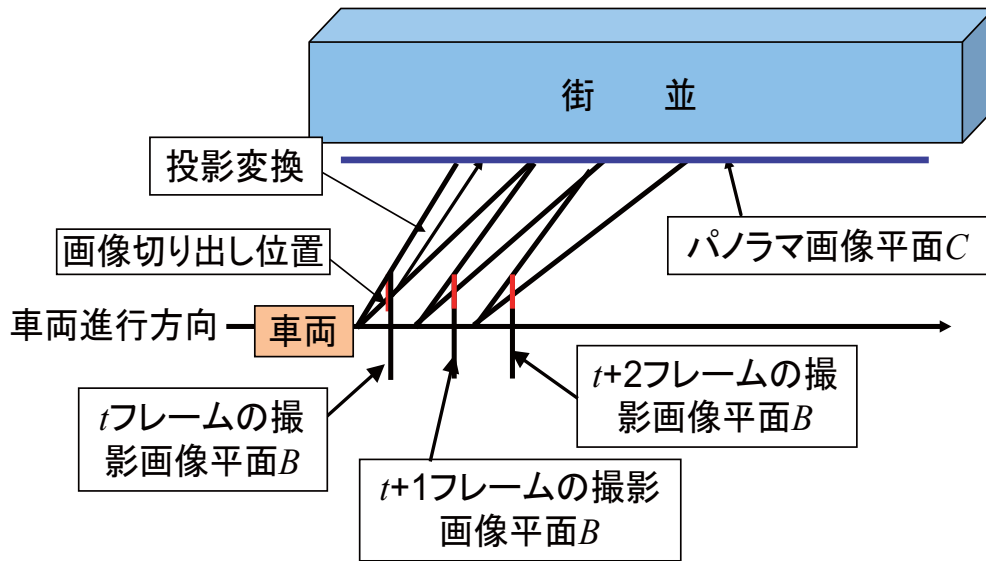


図 11 パノラマ画像生成の考え方

3.2.2 撮影条件の変化に頑健なパノラマ画像生成法

提案手法でパノラマ画像は，道路に面した街並が存在する撮影画像左側の一部を切り出して短冊画像を作成し，それを時系列に繋ぎあわせることで生成する（図 11）．パノラマ画像平面 C は街並の正面と車両進行方向に対して平行だと仮定し，時系列撮影画像はパノラマ画像平面に投影される．

パノラマ画像生成の詳細について以下に述べる．パノラマ画像生成の流れを図 12 に示す．

(1) カメラの回転角度補正

まず，カメラの回転角度を求めるために必要な FOE (Focus of Expansion) の算出方法について述べる．カメラを搭載している車両が直進している状態であれば，車両の移動にともない生じる全てのオプティカルフローは，延長すると画像平面上のある 1 点を通る．この交点を車両の進行方向を示す FOE (f_x, f_y) と呼ぶ．FOE を算出するために，提案手法では 1 フレーム目の撮影画像から Harris operator [21] を用いて抽出した特徴点を，次のフレームで Lucas-Kanade tracker

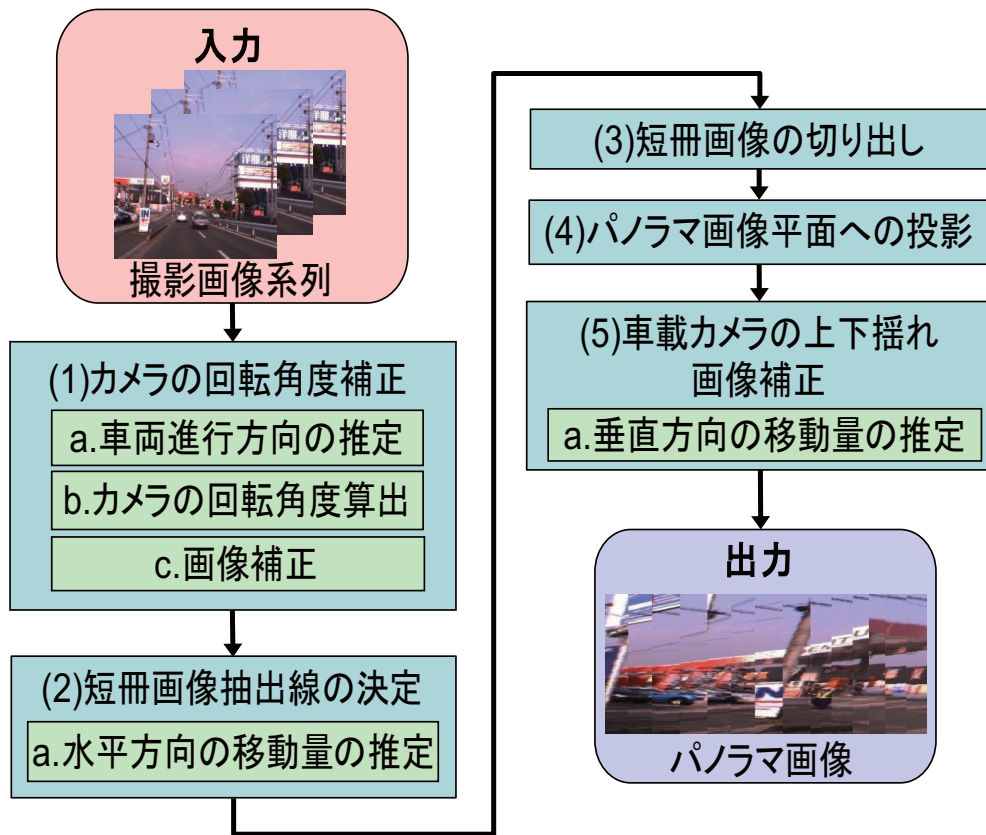


図 12 パノラマ画像生成の流れ

[25] によって追跡する．特徴点のフローは N フレームの間追跡された結果から算出される²．画像全域から求められたすべてのオプティカルフローの交点を $\text{FOE}(f_x, f_y)$ とする．

もし求めた FOE と投影中心を結んだ直線とカメラ光軸（画像中心と投影中心を結んだ直線）が一致していれば，撮影画像平面が車両進行方向に対して垂直な状態であるとみなせる．よって，光軸と直線が一致するパン (θ) とチルト (ϕ) 角度を下式より算出する．

$$\theta = \tan^{-1} \frac{f_x}{f_l}, \quad (1)$$

²本論文の実験では，特徴点の抽出した時間を t としたとき，最低でも次の 2 フレーム ($t+1$ と $t+2$) までは追跡できること確認し， $N=3$ とした．

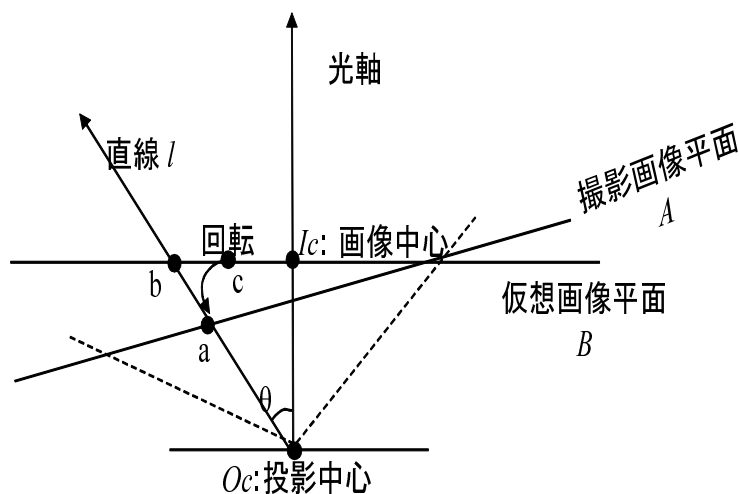


図 13 回転補正

$$\phi = \tan^{-1} \frac{f_y}{\sqrt{(fl)^2 + f_x^2}}, \quad (2)$$

式に用いるカメラの内部パラメータである焦点距離 fl は、第 2.1 節で述べた通り撮影情報として撮影画像と共に収集されるため既知である。求まった θ と ϕ をカメラの回転角度とする。

各車両でカメラ向きが変更されない限り、この節での処理は (FOE の算出とカメラ回転角度の推定) は処理のはじめに一度だけ行われる。

次に、カメラの回転角度に基づき各撮影画像を補正する。図 13 は、画像平面 x 軸と光軸からなる平面を上方から見た図である。仮想画像平面 B が光軸に垂直対して垂直になるように、撮影画像面 A を θ, ϕ により回転させる。補正画像生成のために、仮想画像平面 B 上の点 b に b と光学中心 O_c を結んだ直線 l と撮影画像面 A の交点の点 a の画素の値を代入する。撮影画像面 A 上の点 a の 2 次元座標は仮想画像平面 B の 2 次元座標 (図 13 の点 c) を $-\theta$ と $-\phi$ によって回転させること

で求められる． a の回転は以下の回転行列によって行う．

$$R = \begin{bmatrix} \cos \theta & \sin \theta \sin \phi & -\sin \theta \cos \phi \\ 0 & \cos \phi & \sin \phi \\ \sin \theta & -\cos \theta \sin \theta & \cos \theta \cos \phi \end{bmatrix} \quad (3)$$

この処理の結果，仮想画像平面 B 上のすべての画素値が得られ，回転補正された画像が生成される．図 14 に回転補正なしの撮影画像からパノラマ画像を作成した例 (a),(b),(c) と回転補正後の画像からパノラマ画像 (d),(e),(f) を作成した例を示す．撮影画像は 3 つのカメラ向きの撮影画像列である．図 14 に示したとおり，回転補正なしのパノラマ画像 (a),(b),(c) では短冊画像切り出し位置が異なるため，同じ撮影画像であってもそれぞれパノラマ画像化されている領域が異なる．次にカメラの回転角度の限界について述べる．たとえカメラの回転角度の補正を行っても，街並の存在する領域が小さければ安定したマッチングは困難である．そこで提案手法では変化領域抽出を行うマスクの領域は 200pixel(撮像サイズが VGA の場合) であるため，街並の存在する領域がそれより小さければマッチングを行わない．回転補正の際に得られた FOE(f_x, f_y) の $f_x - 200$ が，画像中心 (ix, iy) の $ix - 200$ 以下であった場合の撮影画像は使用しないこととする．

(2) 短冊画像抽出線の決定

次に撮影画像から短冊画像を抽出する位置である短冊画像抽出線の決定方法について述べる．提案手法では短冊画像の幅は，時系列画像間の水平方向の移動量により得られ，短冊画像抽出線はこの移動量推定のための特徴点抽出領域となる．特徴点は，他の移動物体などによる遮蔽が起こりにくく特徴的な街並領域からの抽出が適している．撮影画像の外側の領域から短冊画像を抽出した場合，透視投影法により画像の見えに歪みが発生する．この歪みはパノラマ画像生成時の時系列短冊画像のスムーズな連結に影響する．さらに画像の外側の領域の特徴点はフレーム間の移動距離が大きいため，安定して長いフローを推定できるが，移動距離が大きすぎると特徴点の追跡に失敗する可能性が増す．一方，画像中心の周辺は歪みが小さいがカメラから遠く離れた背景から抽出されるため，フレーム間の移動距離が小さい．そのためフローは短くノイズが多くなり安定してフローを推定するに不向きである．

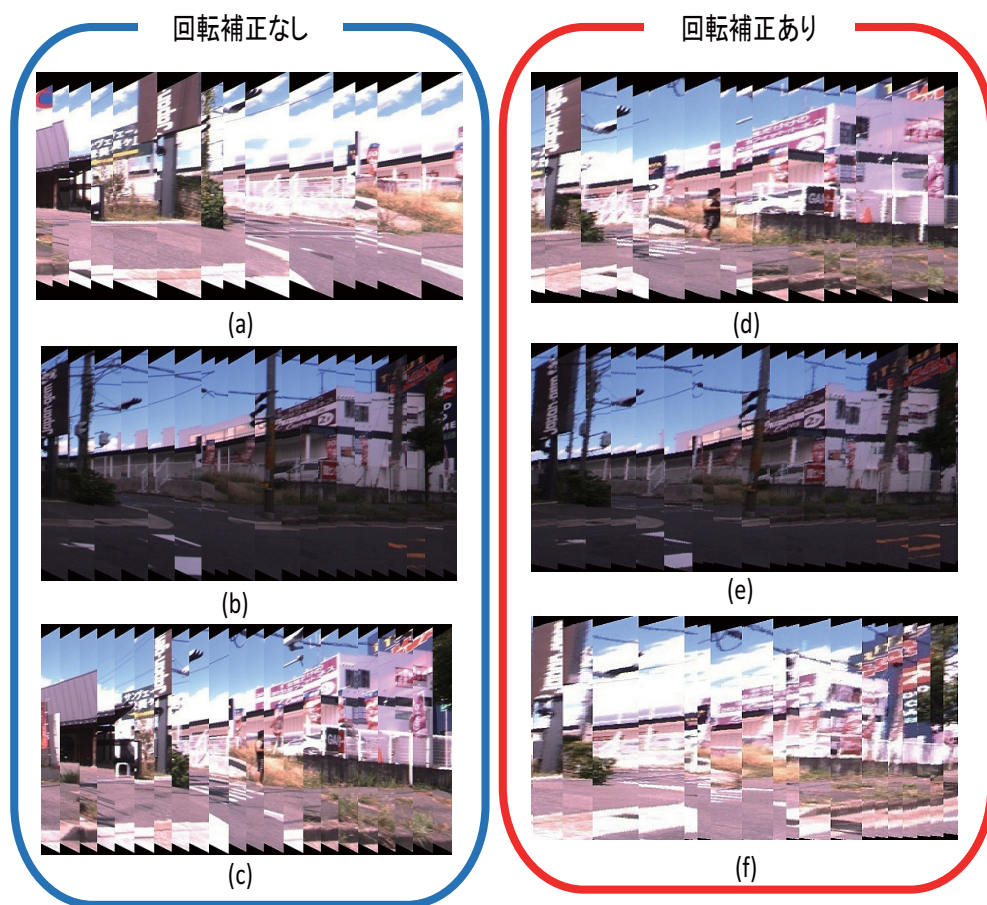


図 14 (回転補正なしのパノラマ画像例:(a),(b),(c), 回転補正ありのパノラマ画像例:(d),(e),(f))

そこで提案手法では、画像中心と画像の左側の中間点の抽出線から短冊画像が切り出される。右側から短冊画像を切り出す場合も同様である。また、想定される全ての画像のサイズ (VGA, XGA, etc.) に無関係に抽出線は同じ決定法とする。これは撮影画像間の視差は画像が撮影されたカメラ向きにより決定するためである。つまり、安定した対応付けのためには全てのカメラの画像から視差のない短冊画像を抽出しなければならない。

短冊画像抽出線の初期値はカメラの光軸からある固定角度 (車両が直進している場合に、十分な特徴点抽出が可能な角度を実験的に求め 11 度とした) に基づい

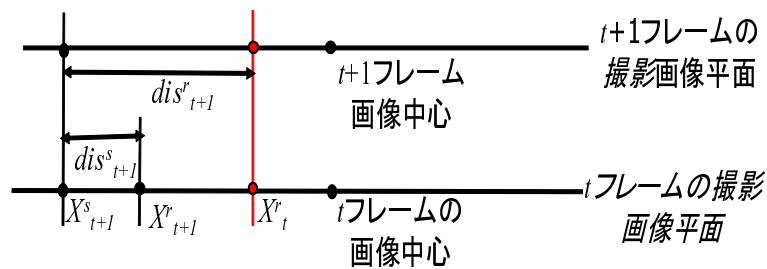


図 15 抽出線の決定方法

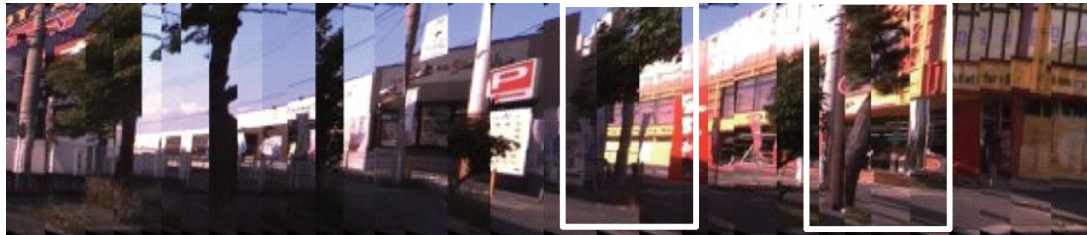
て決定される．つまり特徴点抽出に最適な中間地点は，光学中心と光軸から 11 度の方向を結んだ 3 次元直線と画像平面の交点に位置する³．この y 軸と平行で中間点を通る直線を抽出線 とする．

しかしながら，車両が直進している場合は抽出線の位置は毎回変わらず視差のないパノラマ画像を作成できるが，カーブ走行，左・右折時などの左右の揺れが車両に生じた場合，直線時とはカメラ向きに違いが生じる．そのため直進時と同じ抽出線の位置では，カメラ向きの違いから抽出位置に違いが生じる．ハンドルの回転角度などから画像補正を行うことが可能であるが，提案手法では，左右揺れが生じてても同じ位置から短冊画像抽出を行うために，初期抽出線は固定角度より計算されるが，次フレームからの抽出線は以下の方法で決定する (図 15)．回転補正後の t フレームと $t+1$ フレームにおいて，時刻 t に撮影された画像の抽出線 X_t^r の $t+1$ フレームでの位置 X_{t+1}^s は X_t^r の周辺のオプティカルフローの水平方向の移動量 dis_{t+1}^r から $X_{t+1}^s = X_t^r - dis_{t+1}^r$ となる．この移動量は左右の揺れにより変化するオプティカルフローの水平方向の成分とし，そのオプティカルフローは FOE の推定の際と同様とする．そして $t+1$ フレームでの抽出線 X_{t+1}^r は X_{t+1}^s の周辺から得られた移動量 dis_{t+1}^s により， $X_{t+1}^r = X_{t+1}^s - dis_{t+1}^s$ となる．図 16 にカーブ走行時の撮影画像から，固定値 (初期値の 11 度) の抽出線を用いて作成したパノラマ画像例 (図 16(a)) と提案手法により作成したパノラマ画像例 (図 16(b)) を示す．固定値パノラマ画像上の白い四角に囲まれた領域がカーブによってパノラ

³本論文の実験ではカメラの焦点距離 (810 pixel) を用いた．それゆえに，特徴点は y 軸と平行で画像中心から $(810 \times \tan(11)) \simeq 170$ pixel の線の周辺で抽出された



(a) 抽出線を固定して生成したパノラマ画像



(b) 提案手法により決定した抽出線を利用して生成したパノラマ画像

図 16 左右揺れに対応した抽出線を用いた結果

マ画像に歪みが生じている領域である．提案手法のパノラマ画像の同一の領域では，歪みの発生を抑え滑らかな繋ぎあわせが行われていることがわかる．

(3) 短冊画像の切り出し

観測車両ごとに異なる車速に対応するために，短冊画像の幅は2つの連続したフレームの水平方向の移動量によって決定する．この移動量はオプティカルフローの水平方向の成分とし，そのオプティカルフローはFOEの推定の際と同様の方法で算出する．さらに移動量には，抽出線を中心に前後80pixelの区間から得られたオプティカルフロー(上限は100)の中央値とした．ただし，FOEとは異なり連続するフレーム間の移動量は各フレームごとに算出する．

車速が非常に早い場合は，画像にブレが発生する．提案手法では画像のブレ検出を行う代わりに，上記の移動量の算出時にもし移動量が閾値(実験では70pixel)より長い場合は車速の早い車両によって撮影されたとみなす．そのため，このような画像系列は画像にブレが発生している可能性が高く，パノラマ画像生成には使用しない．また，車両が停止している画像も使用しないため，オプティカルフローを検出できなかった場合も同様である．

(4) パノラマ画像平面への投影

短冊画像は回転補正後の撮影画像から作成されると前述した．時系列画像から作成された全ての短冊画像は仮想パノラマ画像面に透視投影され，つなぎ合わせられる．図 17(a) で仮想パノラマ画像面 C は回転補正画像面 B に対して垂直である． t フレームの回転補正画像面 B から仮想パノラマ画像面 C への投影は， t フレームでの光学中心の位置によって決定される（その位置を O_{c_t} とする）．また O_{c_t} は O_{c_t} から仮想パノラマ画像平面 C までの距離 d と，光軸 Oa に沿う O_{c_t} の位置の 2 つにより決定する．

また各回転補正画像面 B の仮想パノラマ画像面 C への透視投影において，光軸 Oa と仮想パノラマ画像面 C の距離 d はそのスケーリングを決める重要な要素となる．そのため d は任意で決定する．

O_{c_t} から P_t^L/P_t^R を通る直線 L_t^L/L_t^R は時刻 t で抽出された短冊画像の最左線/最右線であり， PP_t^L/PP_t^R は，それぞれ O_{c_t} から P_t^L/P_t^R までの直線と， L_t^L/L_t^R と C の交点である．

図 17 で，まず時刻 t での B_t 上の P_t^L と P_t^R の領域は仮想パノラマ画像平面 C 上では PP_t^L と PP_t^R 間の領域に投影される．同様に， B_{t+1} 上の P_{t+1}^L と P_{t+1}^R の領域は仮想パノラマ画像平面 C 上では PP_{t+1}^L と PP_{t+1}^R 間の領域に投影される．パノラマ画像面 C に投影された画像では段差 ($PP_t^R \cdot PP_{t+1}^L$) や重なり ($PP_{t+2}^L \cdot PP_{t+1}^R$) が生じているが，次の時系列投影画像は図 17(b) のように連結する．すなわち O_{c_t} と $O_{c_{t+1}}$ の相対的な位置は P_t^R の点対応によって決まり，また連続した投影領域の相対配置 $PP_t^L \cdot PP_t^R$ と $PP_{t+1}^L \cdot PP_{t+1}^R$ は $PP_t^L \cdot PP_t^R$ の最右点 (i.e., PP_t^R) と $PP_{t+1}^L \cdot PP_{t+1}^R$ の最左点 (i.e., PP_{t+1}^L) を連結することで近似を行う．

しかしこの単純な投影モデルではいくつかの矛盾が生じる．矛盾を抑止するために，本実験ではパノラマ画像面 C から光軸 Oa までの距離 d はパノラマ画像面から街並の物理的な距離に対応するのではなく，道路に面した建物の輪郭線などの直線が C 上で水平になり，スムーズに画像をつなぎ合わせることが出来る値 d を実験的にもとめた．本実験では焦点距離が 810pixel の際に， $d = 320\text{pixel}$ を用いた．

抽出された短冊画像はパノラマ画像平面 C 上に投影され，時系列順につなぎ

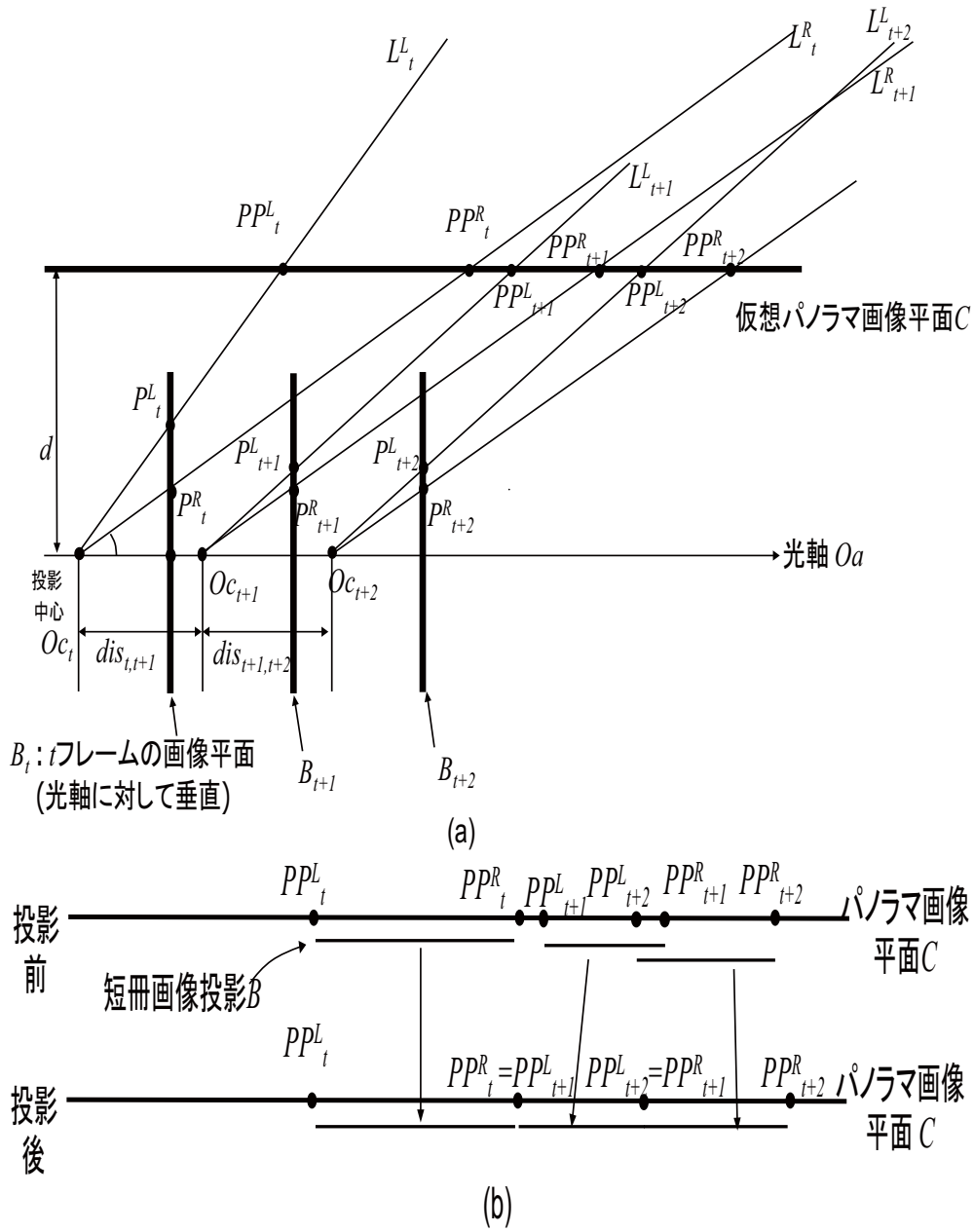


図 17 (a):パノラマ画像平面上の 3 次元投影 (b):投影画像の連結：各矢印は余白や重なりがなく投影画像を連結するための変換を示している

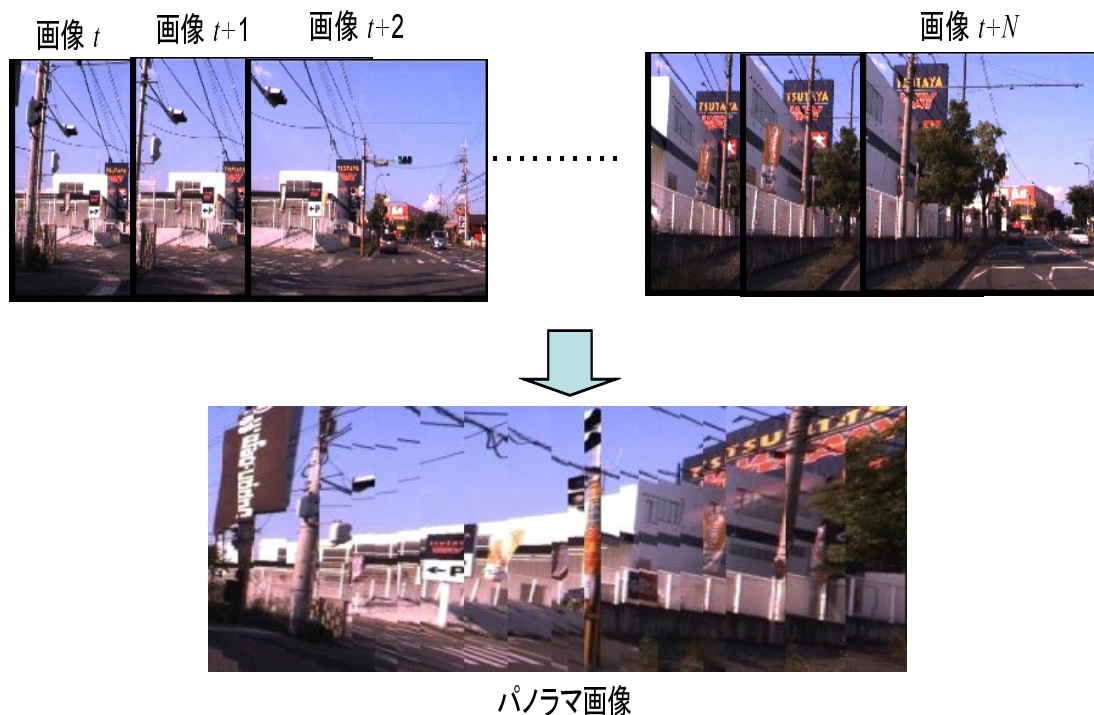


図 18 生成されたパノラマ画像例

合わせられると前述した．実験的に得られた d ではパノラマ画像生成のためのフレーム間移動量 $dis_{t,t+1}$ を得られない．よって，パノラマ画像は以下の手順で生成される．

1. t フレームにおいて，任意の地点に位置する O_{c_t} からの撮影画像から，抽出された B_t 上の短冊画像はパノラマ画像平面 C に投影される．
2. パノラマ画像平面 C 上の全ての投影された画像は時系列順につなぎ合わせられる (図 17(b))．

生成されたパノラマ画像の例が図 18 である．図 18 で示すようにパノラマ画像には隣接する短冊画像間で段差が生じている．これは奥行きが物体に対し，異なる光学中心から撮影された画像を単純に連結しているために生じる．文献 [26] といった精緻な画像を作り出せる画像統合アルゴリズムを用いることで，段差のな

いスムーズなパノラマ画像を作成することも可能である。

また、カメラを街並に対して垂直に向けて撮影すればこのような段差が生じることはない。しかしながら、先述したとおりパノラマ画像の見えに影響を与えるのはカメラ向きの違いだけではなく、観測車両の揺れや走行車線の違いなど撮影条件の変化も含まれる。これはカメラ向きが垂直であっても解決しなければならない問題点である。さらに、道路に対して垂直に在する看板（運転者に対して示される看板）などは垂直方向のカメラでは観測が困難である。また、提案手法の目標はスムーズなパノラマ画像作成することではなく、マッチングに最適なパノラマ画像を作成することである。そして、本システムでは多数の画像を解析することが必要となるので、出来る限り処理は簡素であるべきである。よって、簡単な透視投影を基にしたパノラマ画像生成手法を提案手法では用いる。この有効性は第3.4.2節の様々な提案システムと同様の撮影条件下のもと撮影された画像によるマッチングの実験で証明されている。

(5) 車載カメラの上下揺れ画像補正

パノラマ画像生成時に車両の上下揺れを考慮しなければ、図19(a),(b),(c)の補正前のパノラマ画像のように段差の歪みが生じる。パノラマ画像の上下揺れの補正についてラインスキャンカメラの画像を用いて作成したパノラマ画像中から直線検出を行い、その近似により上下移動量を求める文献[27]が提案されている。しかし、提案手法のパノラマ画像はマッチングに適さない車両進行方向に対して垂直な撮影画像から、投影変換を行いパノラマ画像を作成しているため、隣接短冊画像間の境界がスムーズではなく段差が生じており、直線検出を行うことは困難である。そこで以下の方法で、画像中の上下揺れの補正を行う。

回転補正後の t フレームと $t+1$ フレームにおいて、特徴点の抽出とその対応付けを行う。特徴点の抽出方法とその対応付け方法は、FOEを抽出した際と同様とする。特徴点はパノラマ画像を作成するための領域から抽出される。もし車両が直進している場合、対応点は特徴点とFOEを結んだ直線上に存在するはずである。そのため、3点(特徴点, 対応点, FOE)が直線上に並ぶように、 $t+1$ フレームを上または下へ垂直方向に移動させる。上下方向の移動量は全ての特徴点

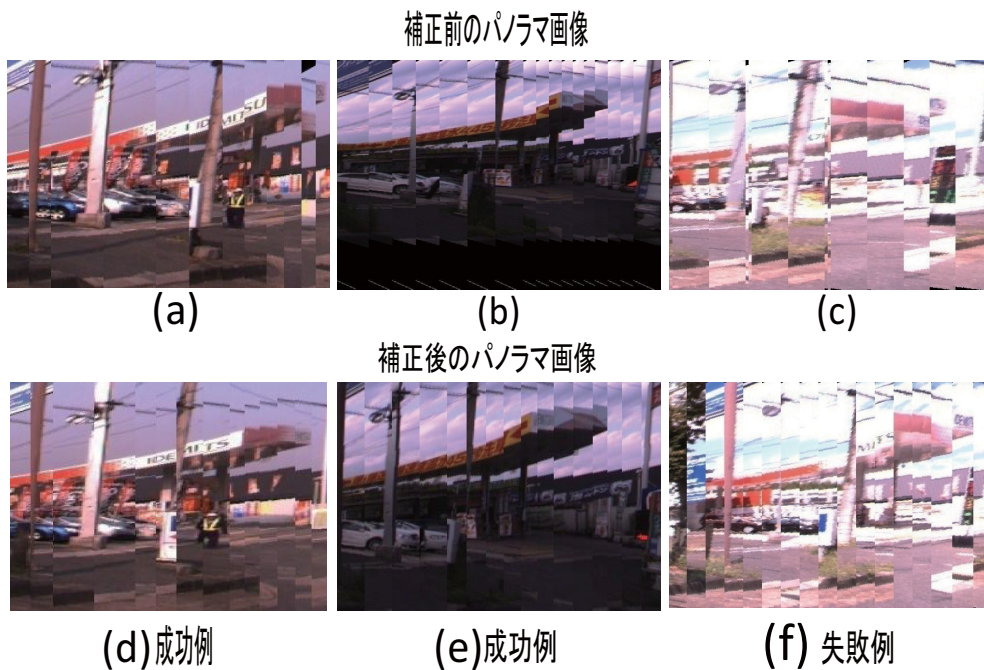


図 19 車載カメラの上下揺れ補正結果

から得られた移動量の中央値とする．この修正の結果，上下揺れによる歪みは図 19 の補正後パノラマ画像の (d),(e) に示すように補正された．補正後のパノラマ画像 (f) は補正に失敗した画像である．補正後のパノラマ画像 (f) は照明条件の変化が著しく，移動量を推定するための特徴点の対応付けに失敗した．

3.2.3 パノラマ画像を用いた画像間の対応付け

本節ではパノラマ画像を用いた画像間の対応付けについて述べる．図 20 に過去画像系列と入力画像系列から各々作成したパノラマ画像間の対応付け処理の流れを示す．対応付けでは，パノラマ画像からエッジ特徴量を抽出し，入力パノラマ画像のサイズを縮小・拡大しながらマッチングを行う．

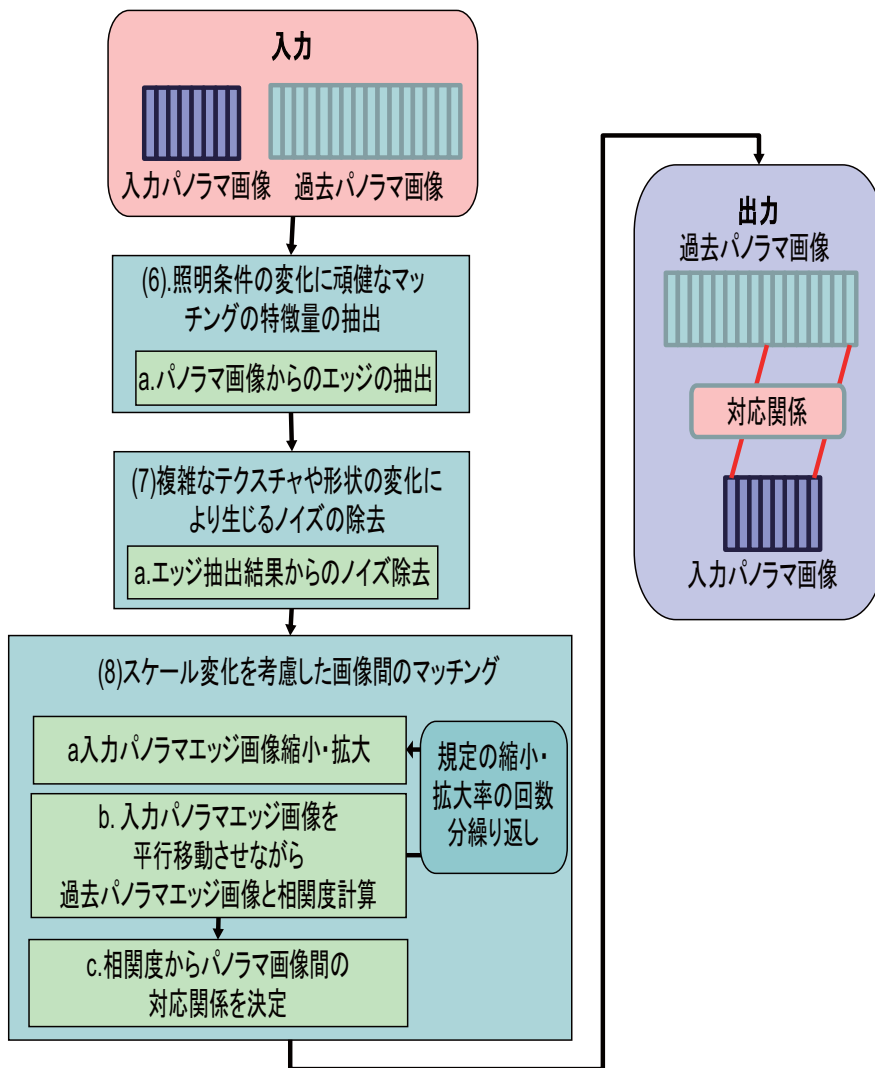


図 20 パノラマ画像を用いたマッチングの流れ

-1	-2	-1
-1	-2	-1
0	0	0
1	2	1
1	2	1

Vertical

-1	-1	0	1	1
-2	-2	0	2	2
-1	-1	0	1	1

Horizontal

図 21 エッジ抽出に用いるフィルタ

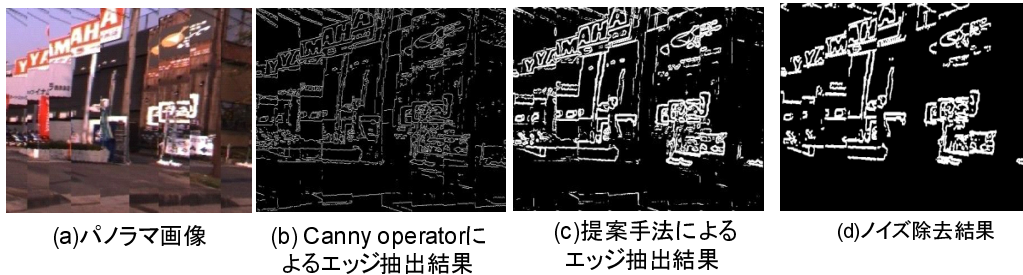


図 22 エッジ抽出とノイズ除去結果

(6) 照明条件の変化に頑健なマッチングの特徴量の抽出

提案手法では照明条件の変化に頑健なエッジを、マッチングの特徴量として用いる。しかし、照明条件の違いに影響を受けやすい複雑なテクスチャを持った看板、時間や風などの影響で形状が変化する街路樹やノボリ旗などが原因で一時的な変化が大量に生じる可能性がある。そこで、安定して検出可能な特徴でありマッチングに適した、建物、看板などの大きな構造物の輪郭線である長いエッジ直線に注目し、それらを抽出することで一時的な変化となりやすい細かなエッジによるマッチングへの悪影響を抑止する。そこで提案手法では、以下のようなエッジオペレーターでエッジ抽出後、その抽出結果から短いエッジ直線または小さなエッジ点のブロックを除去する。はじめに、Sobel オペレータを改良した図 21 のフィルタを用いて、Sobel オペレータと同じ局所積和演算でグラジエントの微分値を計算する。このオペレータは各 RGB チャンネルにそれぞれ適用し、その 3 つの結果 (RGB) の微分値合計を閾値によって二値で表現する (図 22(c))。実験では閾値を 500 とした。この閾値は異なる時間、撮影条件下 (様々な天候条件、ビルの多

い都市部，樹木の多い郊外など) で撮影された様々な画像を用いて多数の実験を行った結果をもとに決定した．このオペレーターによって一般的なエッジ抽出に用いられる Canny operator に比べて (図 22(b) 細かなエッジ直線の抽出を抑えられる (図 22(c)) ．

(7) 複雑なテクスチャや形状の変化により生じるノイズの除去

提案手法でのパノラマ画像は幾何的，光学的にスムーズではない．そのため，短冊画像の境界には人工的なエッジ直線が生じる．よって，このエッジ検出は短冊画像間の境界には適用しない．

さらに上記のオペレーターを用いるだけでは，必要のない街路樹や看板などの密集した細かなエッジも抽出される．そこで次にエッジ抽出結果からこれらのエッジを，連続して存在しているエッジセグメントのサイズによって以下の通りに除去する．

(1) 全てのエッジを 8 近傍内に連続して存在するエッジ画素を元にセグメント化する．

(2) もしそのサイズ (1つのセグメントに属するエッジ画素数) が閾値 (50pixel) 以下であれば，エッジセグメントを除去する．

閾値は，第 2 車線を走行中の観測車両から得られた撮影画像中の検出対象である看板 (変化検出を行う領域に存在する) の面積より小さい値とした．これらの処理の結果，マッチングに適したエッジ画像を生成する (図 22(d)) ．

さらに，過去パノラマ画像のエッジにはなをかけることで，微妙なずれが生じていても入力画像との相関度が高くなるようにする．

(8) スケール変化を考慮した画像間のマッチング

画像マッチングはこれまでの処理で得られた過去・入力エッジパノラマ画像を用いて行う．具体的には，車線の違いなどによる建造物の撮像サイズの違いに対応するために様々なサイズに入力画像を拡大・縮小し，過去画像マッチングを以下の手順で行う (図 20(8)) ．

- (1) 入力パノラマエッジ画像を様々なサイズに拡大・縮小する。
- (2) 各サイズの入力パノラマエッジ画像を過去パノラマエッジ画像上で上下左右に平行移動させながら各位置ごとの正規化相関を計算する。

入力画像とのマッチングで過去画像上で正規化相関値が最大となる位置を求める。そして、過去パノラマ画像の左端に投影されている元画像が入力画像の元画像と対応付けられたものとみなす。提案手法では、2車線までの位置違いに対応可能となるように入力画像の倍率変更は0.5倍から1.5倍まで0.1倍間隔とした。これは一般道路には2車線が多く、さらに3車線目以降からは画像中の街並領域が小さく、他車両による遮蔽が大きいためマッチングが困難なためである。また、平行移動は1pixel間隔とした。

3.3 第2段階：SIFT特徴点による車載カメラ画像間の対応付け

撮影位置が大きく異なり、さらに撮影条件の違いから同一地点を撮影した画像間であっても画像間の重なりは小さく、また見えも著しく異なる。そのため、先述したとおり1対1の画像間のマッチングでは、マッチングの起点画像の前後15mの区間の画像間を直接比較するには、計算コストが高くかつ街並の変化や駐車車両などによる遮蔽の影響を受けやすい。そこで、提案手法では時系列画像を繋ぎ合わせて作成したパノラマ画像間のマッチングと特徴点による1対1の画像間のマッチングの2段階によるマッチングを行う。

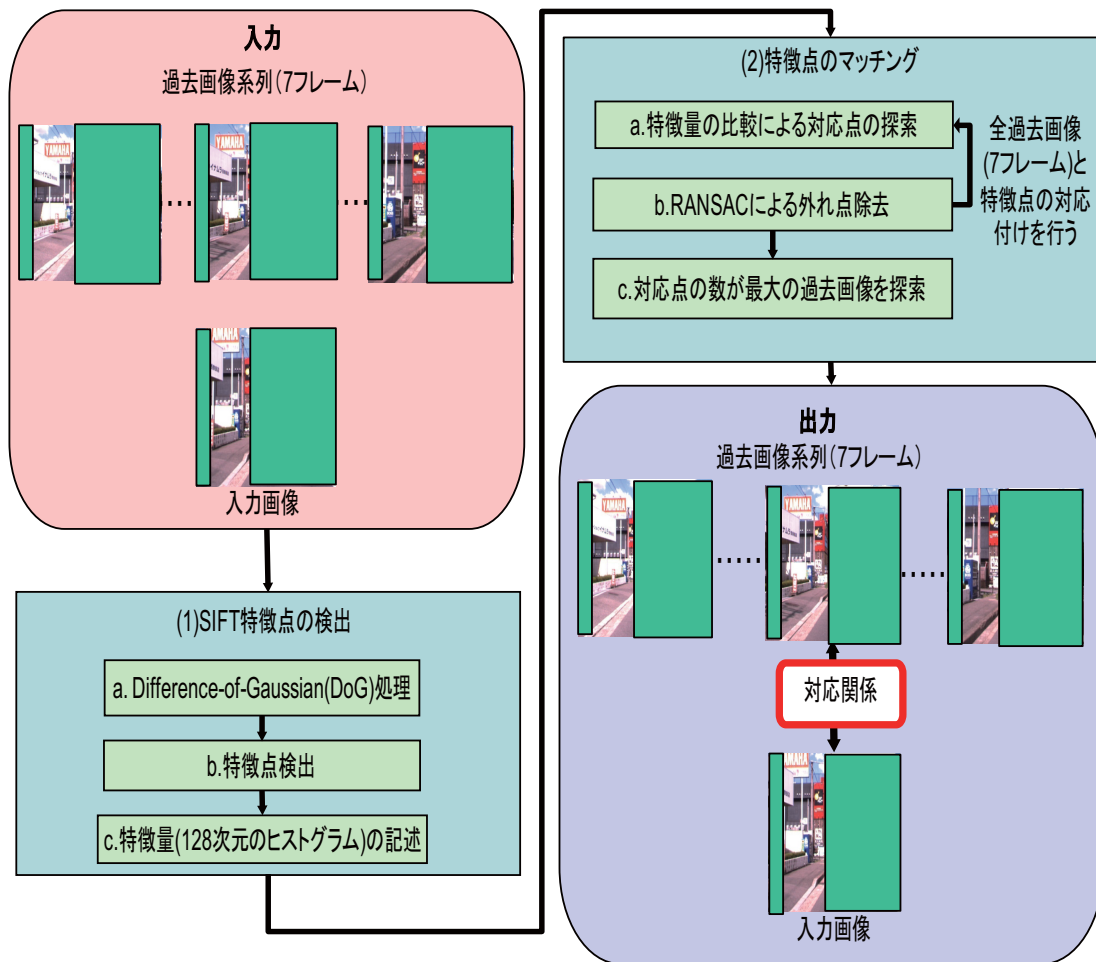


図 23 第2段階:SIFT特徴点による車載カメラ画像間の対応付けの流れ

本節では、第2段階である特徴点による画像間の対応付けについて述べる。

図23に第2段階のSIFT特徴点による車載カメラ画像間の対応付けの流れを示す。第2段階では第1段階であるパノラマ画像による画像間の対応付けの結果から得られた、画像の対応関係を利用し、入力画像1フレーム対過去画像数フレームとの対応付けを行う。過去画像は対応付けられた画像を含め前後3フレームずつで、計7フレームとした。想定するフレームレート(3-4fps)の場合、撮影画像の特徴点抽出領域から得られた特徴点の追跡が可能なフレーム数が平均的に3フレーム前後である。そのため第1段階のマッチングで得られた起点となる撮影画像の前後3フレームを特徴点による対応付けの探索範囲とした。この対応付けでは各画像中から局所特徴量を抽出し、特徴点の対応付けを行う。そして、もっとも多くの対応関係がえられた画像を対応付けられた画像とする。

また、計算コスト削減のため、マッチングを行う際の画像はパノラマ画像を作成する際に撮影画像から切り出して作成した短冊画像と同様に街並が存在する確率の高い画像左端より画像を切り出して作成する。その切り出し位置は短冊画像抽出位置と同じであるが、短冊幅はVGAの画像に対し左側の街並を対象とするためVGAサイズの画像の横幅(480pixel)を考慮し、200pixelとした。200pixelであれば、画像の左側に存在する街並領域を十分に含む。

3.3.1 短冊画像からのSIFT特徴点の抽出

提案手法では、この対応点関係を得るために必要な特徴点に Scale-Invariant Feature Transform(SIFT)[22] を用いる。SIFT[22] は、特徴点の検出と特徴量の記述を行うアルゴリズムである。画像の回転・スケール変化・照明変化に頑健な特徴量を記述可能であり画像のマッチングや物体認識でも用いられている特徴量である。SIFTは特徴点(以下、キーポイントとする)の検出と特徴量の記述の2つで構成されている。キーポイントの検出では、照明条件や視点の変化など、異なる撮影条件下でも同一物体から同様に定まる点や尺度を見つける必要がある。つまり、画像のスケールを変化させた場合でも、点が一定に定まればキーポイントの候補とする。SIFTでは、Difference-of-Gaussian(DoG)処理を用いて、分散 σ の値を変えて得られる複数の平滑化画像を順に並べたスケール空間におい

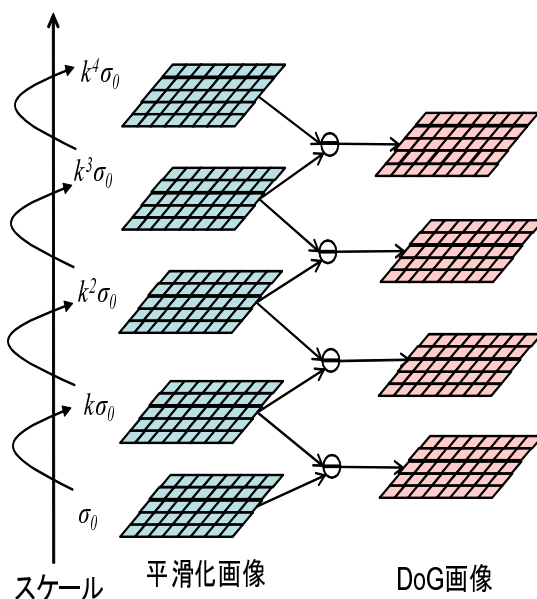


図 24 Difference-of-Gaussian 処理

て、極値探索をすることでキーポイントの位置とスケールを決定する。処理の詳細について次に述べる。

まず、グレースケール化した入力画像 $I(u, v)$ をガウシアンフィルタ $G(x, y, \sigma)$ により平滑化することで平滑化画像 $L(u, v, \sigma)$ を作成する。

$$L(u, v, \sigma) = G(x, y, \sigma) * I(u, v) \quad (4)$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (5)$$

そして、スケールの異なる平滑化画像 $L(u, v, \sigma)$ 間で差分をとり、DoG 画像 $D(u, v, \sigma)$ を作成する。DoG 画像 $D(u, v, \sigma)$ は次式で得られる。

$$\begin{aligned} D(u, v, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(u, v) \\ &= L(u, v, k\sigma) - L(u, v, \sigma) \end{aligned} \quad (6)$$

この処理を σ_0 から k 倍ずつ異なるスケール間で行い、複数の DoG 画像 (図 24) を求める。さらに SIFT では、入力画像をダウンサンプリングすることで、 σ の変

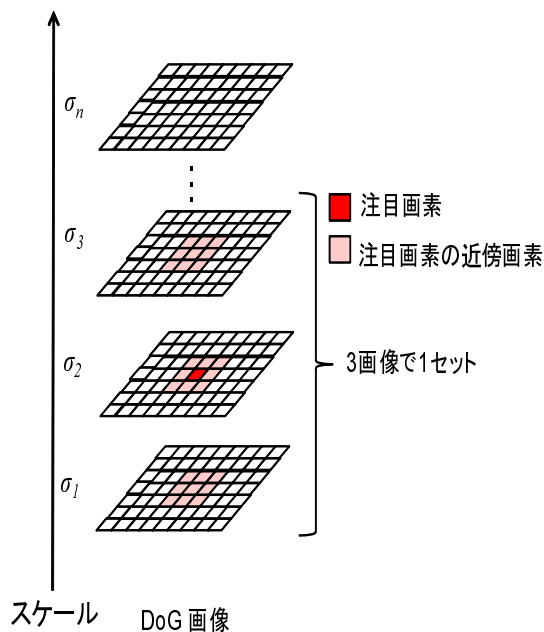


図 25 特徴点検出

化の連続性を保持した平滑化処理を行っている．差分のため DoG の値が大きくなる σ では，スケールの変化領域にエッジ等の情報量を多く含んでいる．次にこの DoG 画像から極値を求めキーポイントとスケールを決定する．

平滑化処理では σ の初期値 σ_0 で平滑化し，平滑化画像 $L_1(\sigma_0)$ を得る．次に k 倍した値 $k\sigma_0$ で平滑化を行う．これを繰り返すことで σ の異なる平滑化画像を得る．これらの一回の処理を 1 オクターブと呼ぶ．この σ の増加率 k は 1 オクターブのスケールスペースの分割数により決定される．スケールスペースの分割数を s としたとき，1 オクターブではスケールスペースは σ_0 から $2\sigma_0$ まで増加するため，増加率 k は $k = 2^{1/s}$ となる．また，文献 [28] から分割数 $s = 3$ ，初期値 $\sigma_0 = 1.6$ としたとき，最適なキーポイントが得られることがわかっているため，提案手法でも同様の値を用いる．

極値の検出は，注目画素の存在する DoG 画像とさらに上下のスケール DoG 画像の 3 枚一組で行う (図 25)．現在のスケールの DoG 画像中の注目画素の 8 近傍，その上下のスケールのそれぞれ近傍 9 点，計 26 近傍に対して値の比較を行う．注

目画素の値が極値であった場合その画素をキーポイント候補点として検出する。この極値検出は σ の最小のDoG画像から行い、一度極値が検出された画素は次のスケールで極値が検出されてもキーポイント候補点としない。この処理をスケールの異なるDoG画像の全画素に対して行う。SIFTは、特徴を最も含むスケール σ を決定できるため、空間的に同範囲の領域から特徴量を記述することで、拡大・縮小に不変な特徴量となる。

次に検出したキーポイントに対して特徴量の記述を行う。検出された各キーポイントに対して、オリエンテーションを求める。オリエンテーションはキーポイントにおける方向を表し、特徴量記述の際にオリエンテーションにより向き正規化を行うことで、回転に不変となる。

キーポイントのオリエンテーションを求めるには、まずキーポイントが検出された平滑化画像 $L(u, v)$ の各画素において、勾配強度 $m(u, v)$ とその勾配方向 $\theta(u, v)$ を以下の式により求める。

$$m(u, v) = \sqrt{f_u(u, v)^2 + f_v(u, v)^2} \quad (7)$$

$$\theta(u, v) = \tan^{-1} \frac{f_v(u, v)}{f_u(u, v)} \quad (8)$$

$$\begin{cases} f_u(u, v) = L(u + 1, v) - L(u - 1, v) \\ f_v(u, v) = L(u, v + 1) - L(u, v - 1) \end{cases} \quad (9)$$

上記から得られた局所領域における勾配強度 $m(x, y)$ と勾配方向 $\theta(x, y)$ から重み付方向を求め、ヒストグラム h を以下の式により作成する。

$$h_{\theta'} = \sum_x \sum_y w(x, y) \cdot \delta[\theta', \theta(x, y)] \quad (10)$$

$$w(x, y) = G(x, y, \sigma) \cdot m(x, y) \quad (11)$$

$h_{\theta'}$ は全方向を36方向に量子化したヒストグラムである。 $w(x, y)$ は局所領域の画素 (x, y) の重みであり、この36方向のヒストグラムの最大値の80%以上となるピークをキーポイントのオリエンテーションとして割り当てる。

次に、このオリエンテーションからSIFT特徴量により128次元の特徴量を記述する。キーポイント周辺領域の持つ勾配情報を用いるため、キーポイントの周辺領域を、割り当てられた代表オリエンテーションを基準とした軸に回転させる(図

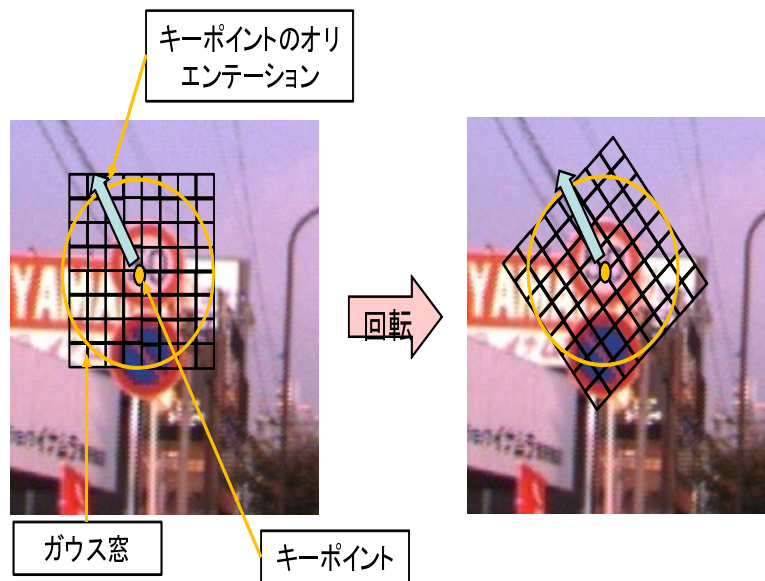


図 26 特徴量の記述の範囲

26) . キーポイントが持つ代表オリエンテーションを基準として特徴記述するため、回転に対して不変な特徴量となる。また、128次元の各特徴ベクトルの長さはベクトルの総和で正規化する。これにより、キーポイントは照明変化に対して影響の少ない特徴量となる。そして、この特徴点周辺領域を16ブロックに分割し、ブロックごとに8方向(45度ずつ)の勾配ヒストグラムを作成する。これにより4ブロック×4ブロック×8方向の128次元の特徴量が得られる(図27)。

本システムの撮影画像に多く存在する道路領域のように輝度が低く、テクスチャが単調な領域からSIFT特徴量を計算すると、撮影条件の変化によって周辺画素と著しく輝度値が異なる画素の影響を受け、対応づくべき画素間でもSIFT特徴量の相違度が大きくなる可能性がある。そこで、SIFT特徴量のヒストグラムに対してベクトルの総和による正規化後、輝度勾配の著しく大きい値に対しては切り捨て(0.2以上を切り捨て)を行い、再びベクトルの総和による正規化を行う。この外れ値の除去処理により、道路領域に属する画素間の相違度を抑えマッチングの誤りを低減させる。なお本論文に使用する、すべてのSIFT特徴量に対して同様の処理を行う。

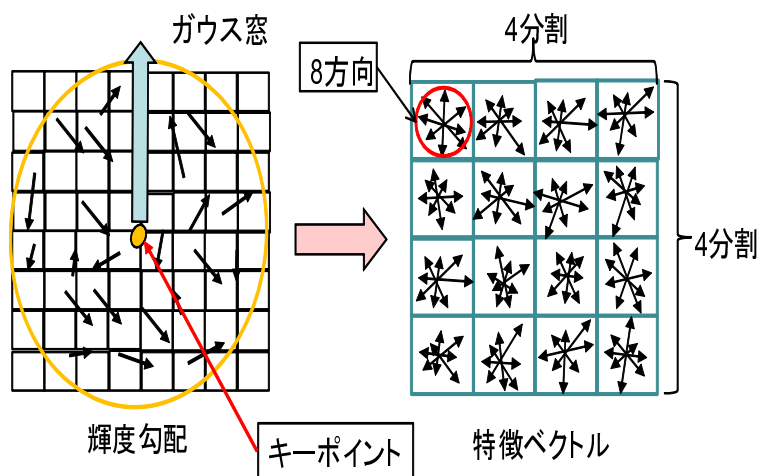


図 27 特徴量の記述

3.3.2 SIFT 特徴点によるマッチング

得られた特徴点のマッチングは SIFT 特徴点のもつ 128 次元の SIFT 特徴量のヒストグラムの比較により行う。対応付けを行った結果を図 28(a) に示す。またさらに RANSAC[29] を用いて外れ点の除去を行う。その結果が図 28(b) である。7 フレームの過去画像間と入力画像間において対応点を抽出し、対応点数が多い画像間を対応付けられた画像とする。もし、対応点数が同数である場合は、パノラマ画像の抽出結果である起点画像に近い画像を優先する。また、遮蔽などにより対応点が得られない場合を考慮し、得られた対応点数が閾値より少ない点数であった場合はパノラマ画像の抽出結果を優先する。なお、特徴点は最大 100 点まで求め、閾値は最大値の 20% にあたる実験的に定めた 20 点以下とした。

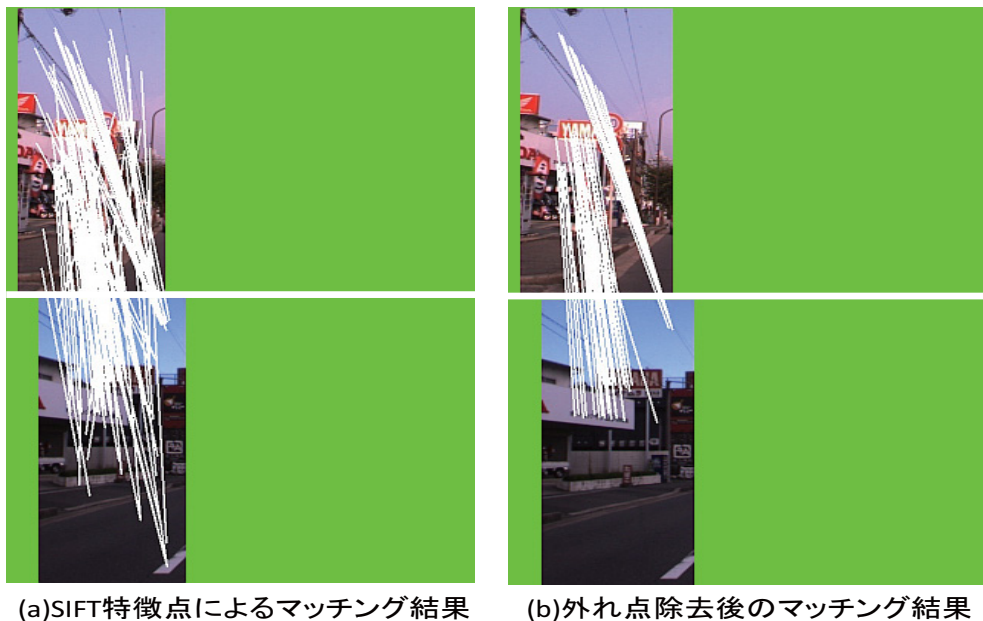


図 28 SIFT 特徴点の対応付けによるマッチング結果

3.4 実験と評価

3.4.1 実験概要

評価実験には、奈良県奈良市・生駒市内の7.3kmの区間(図29)を走行して収集したデータを使用した。図29の赤い線が走行経路である。情報収集には本システムで想定するシステム構成に従い、一般車両に単眼カメラと普及型GPSを設置して行った。利用した車両は3種類で、車両速度は一定ではなくおよそ30~60km/hであった。カメラは、バックミラー下に車両進行方向に向けて設置し、そのカメラの回転角度は走行車両ごとに異なっている。撮影画像はJPEG圧縮のかかった640×480pixelで、フレームレートは3fps~4fpsである。カメラの内部パラメータ(画像中心や焦点距離など)は文献[30]を利用して求めた。GPS測位更新時間は一秒間に一回であり、この間に撮影された画像には同一のGPS測位座標を併せて記録した。

様々な撮影条件の変化(カメラ向きの違い、照明条件、道路状況、街並変化、撮影時の走行車線の違いなど)への頑健性を確認するため、異なる日時、異なるカ



図 29 実験データを収集した走行区間

メラ向きで撮影された 20ヶ所の地点の画像列を GPS の座標をもとに抽出し、対応付けを行った。地図情報更新に必要な街並変化は以下の地点に生じている。

地点 G, N, Q: 建物の構造とテクスチャに変化が生じている

地点 D, L, Q: 建物と独立して存在する看板の構造とテクスチャに変化が生じている。

地点 D, E: 建物のテクスチャに変化が生じている。

地点 N, O: 建物と独立して存在する看板のテクスチャに変化が生じている。

次に各地点の道路環境条件を以下に述べる。

地点 C, F, K: 上りまたは下り坂を走行時に撮影された。

地点 D, N, M: 道路の段差により車両に上下揺れが生じている。

地点 F, I, N, O, Q: カーブ走行時に撮影された。

地点 K, N, P, O: 多数の街路樹により、街並みに遮蔽が生じている。

また、以下の一時的変化や撮影条件の違いも画像間には生じている。

- 照明条件が著しく変化し、画像の一部では色とびが生じている。
- 天候状態は雨天(少量)・曇天・晴天時の3パターンが含まれる。
- 走行車線が異なる。
- 車線変更が行われる。
- 道路沿いにバスや軽トラックなど大きな遮蔽物が存在している。
- 街路樹やのぼり旗の変化が生じている。

さらに、1地点につき異なる日時に撮影された過去画像3シーケンス、入力画像3シーケンスをそれぞれ用意した。先述したとおり、過去画像の3シーケンス間には街並変化は生じていない。

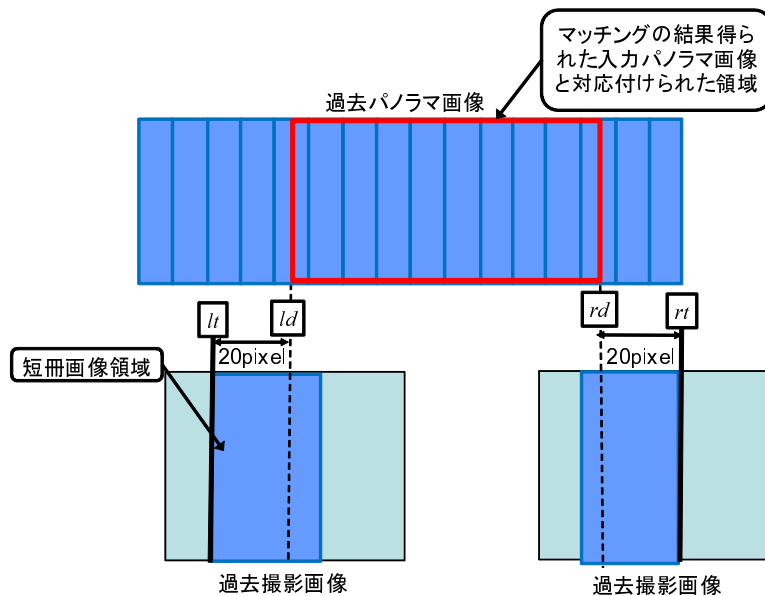


図 30 第 1 段階の対応付け結果の確認方法

3.4.2 実験結果

第 1 段階による画像間対応付けの実験結果

対応付けの結果の成否は以下の方法で行う．まずマッチングしている領域を実撮影画像を目視で探す．そして実撮影画像の最左線 lt と最右線 rt とマッチングの結果得られた画像の最左線と最右線の間を距離を ld/rd とし，この ld と rd が 20 pixel 以内であれば，対応付けの結果を成功とする (図 30) ．

20 地点全 180 シーケンスでの対応付けを行った結果，180 シーケンス中 145 シーケンス (精度 80%) の対応付けに成功した．次に本論文で問題点として解決を図った車両の上下および左右揺れ，街並変化，駐車車両による遮蔽，照明条件の変化，撮影時の走行車線の変化への頑健性について対応付けを行った結果を図 31 に示す．各過去パノラマ画像中の赤い四角形で囲まれた領域がマッチングの結果対応付けられた領域である．対応付けの結果，題が生じている画像に対して安定してマッチングが成功し，提案手法の有効性が確認できた．各地点の成功数を表 3 に示す．次に図 32 に対応付けに失敗した画像例を示す．各過去パノラマ画像中の赤



図 31 撮影条件・街並・移動物体による変化が生じている画像に対しての第1段階の対応付け結果例(成功例)

地点	結果 (成功数)	地点	結果 (成功数)
地点 A	9/9	地点 K	8/9
地点 B	9/9	地点 L	5/9
地点 C	8/9	地点 M	5/9
地点 D	8/9	地点 N	9/9
地点 E	9/9	地点 O	8/9
地点 F	3/9	地点 P	4/9
地点 G	8/9	地点 Q	9/9
地点 H	6/9	地点 R	9/9
地点 I	3/9	地点 S	9/9
地点 J	8/9	地点 T	8/9

表 3 第 1 段階による画像間対応付け結果

い四角形で囲まれた領域がマッチングの結果対応付けられた領域で，白い四角形で囲まれた領域が正しい対応づくべき領域である．以下に対応付けの失敗の原因をまとめる．

図 32(a): 撮影途中に車線変更が行われた．フレームレートが低いため，このような急激な左右揺れが生じるとオプティカルフローの抽出に失敗する．

図 32(b): 過去画像では大きなカーブに加え，停車から走行という急激な車速の変化と街路樹の遮蔽によりオプティカルフローの抽出に失敗し，短冊画像の抽出に影響が生じた．

図 32(c),(d),(f): 街路樹の遮蔽により激な左右揺れが生じるとオプティカルフローの抽出に失敗した，さらにパノラマ画像間においても相関度が低下し，かつ街路樹は似たテクスチャであるため誤対応の原因となる．

図 32(e): 走行車線の違いにより，街並が見える範囲が異なり，さらに広範囲にわたるのぼり旗による変化がマッチングの相関度を下げた原因であると考えられる．

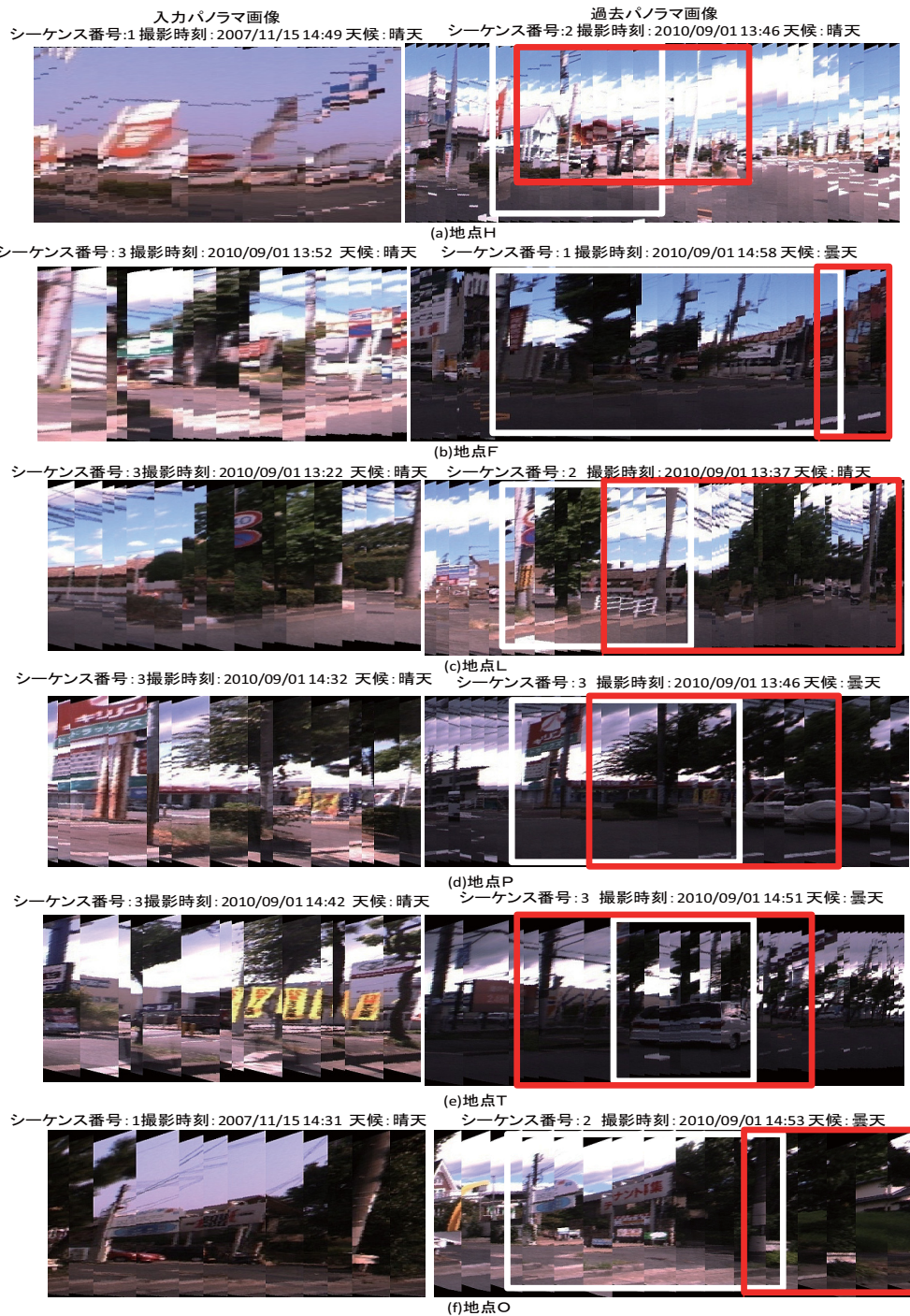


図 32 失敗した第1段階の対応付け結果例

一時的な変化による対応付け失敗は、大量画像取得が前提となる提案手法では影響が少ないが、(b)でおきたカーブなどの道路環境によるマッチングの失敗はシステムに影響を与えるため、改善が必要である。さらに、同じ20ヶ所の地点を高いフレームレート(4-5fps, 6-7fps)で撮影した画像を用いた場合でも、同じ対応付け結果となることを確認した。しかしながら、さらに低いフレームレート(1-2fps)の撮影画像を用いた対応付けでは20ヶ所中11ヶ所の地点で失敗した(この際の対応付けに用いたのは1シーケンスのみ)。

第2段階による画像間対応付けの実験結果

次に、第1段階による対応付けの結果をもとに、第2段階のSIFT特徴点による対応付けを行った。20地点の対応付け結果から1地点につき過去画像3シーケンス、入力画像3シーケンスを用意した。ただし、今回の入力画像の1シーケンスは10フレームで構成されているとする。対応付けの結果、1800フレーム中1557フレームの対応付けに成功し、その精度は86%となった。図33に成功例を示す。図33(a)(b)(c)は第1段階で対応付けに成功した画像、(d)(e)(f)は第1段階で対応付けに失敗した画像である。第1段階において対応付けに失敗している画像においても、第2段階の対応付けにより対応付けが成功していることが確認できる。各地点における対応付けの結果を表4に示す。次に失敗例を図34を示す。図34(a)(b)(c)は第1段階で対応付けに成功した画像、(d)(e)(f)は第1段階で対応付けに失敗した画像である。対応付けに失敗した原因として以下が考えられる。

図34(a) カーブ走行中であるため、入力画像中の街並領域が小さく対応付けに失敗した。

図34(b) 第1段階の対応付け失敗による位置ずれが大きく、重なりあう領域が小さいため対応付けに失敗した。

図34(c) 電線上のSIFT特徴点の誤対応が大量に発生した。

図34(d) 第1段階の対応付け失敗による位置ずれが大きく、重なりあう領域が存在しなかった。

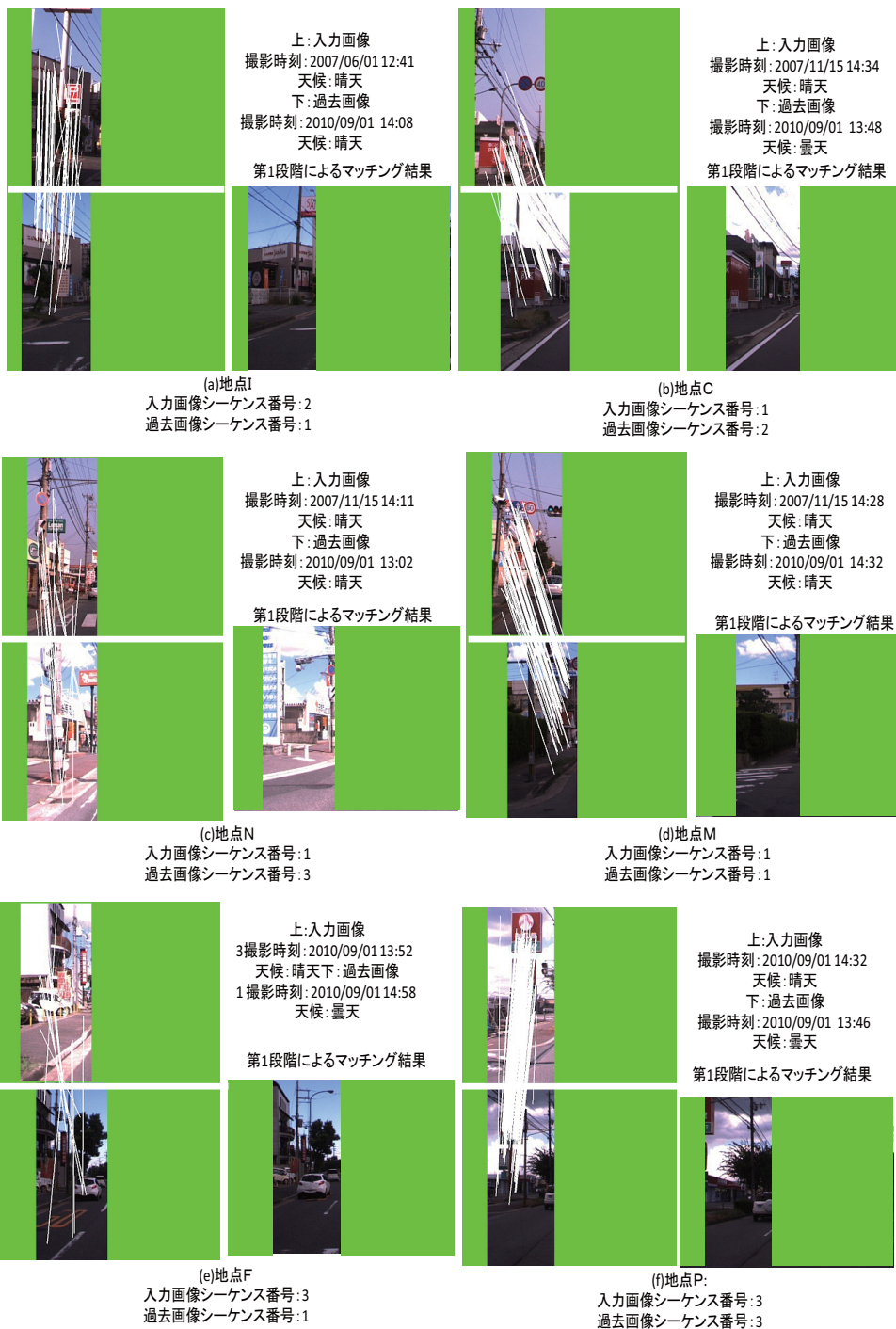


図 33 第 2 段階の対応付け結果 (成功例)



図 34 第 2 段階の対応付け結果 (失敗例)

地点	結果 (成功)	地点	結果 (成功)
地点 A	90/90	地点 K	83/90
地点 B	90/90	地点 L	62/90
地点 C	88/90	地点 M	72/90
地点 D	89/90	地点 N	86/90
地点 E	90/90	地点 O	83/90
地点 F	45/90	地点 P	43/90
地点 G	87/90	地点 Q	87/90
地点 H	71/90	地点 R	90/90
地点 I	46/90	地点 S	90/90
地点 J	84/90	地点 T	81/90

表 4 第 2 段階の画像間対応付け結果

図 34(e) テクスチャが相似している街路樹により街並が遮蔽されているため，誤対応が生じた．

図 34(f) 第 1 段階の対応付け失敗による位置ずれが大きく，重なりあう領域が存在しなかった．

第 1 段階の対応付け失敗が原因による著しい位置ずれ(前後 3 フレームを超える)に関しては，第 2 段階では対応できない．図 34(c) でみられるような対応付けの失敗は第 2 段階固有の問題である．そのため，抽出された特徴点の評価や街路樹による遮蔽が大きい画像は除くことで改善を図る必要がある．

3.4.3 従来手法との比較

1 対 1 の画像間のマッチングによる画像間対応付けとの比較

提案手法では非同期で低いフレームレートの撮影画像を用いるため，以下の理由から 1 対 1 の画像間のマッチングを用いることはできない．

- 撮影間隔が長く同じ位置からの撮影は不可能であるため，単純なマッチング方法では対応できない．
- 高性能で高価なカメラと特定の撮影条件下で撮影されることを前提としているパノラマ画像生成手法（文献 [12, 13, 14, 15, 16]）は本システムでは用いることができない．

これらの手法の代わりに，パノラマ画像を用いた画像対応付けの有効性を確かめるため比較実験として1フレームの入力画像と複数の過去画像を比較する対応付け実験を行った．ただし，回転補正後の撮影画像を用いる．前述したとおり本システムでは撮影された画像をそのまま比較して対応付けるのは困難である．そのため画像から一部を切り出して短冊画像を生成し，それを複数の撮影画像と比較し，最も相関度の高い画像を対応付けられた画像とする．短冊画像と撮影画像の高さは同じで，短冊画像の幅を 20～80pixel まで変化させた．マッチングと評価方法は提案手法と同じものとする．対応付けの結果を以下に示す．

Width=20 pixel: 成功した地点:なし.

Width=30 pixel: 成功した地点:C, N.

Width=40 pixel: 成功した地点:C, N.

Width=50 pixel: 成功した地点:C, I, L, N, R.

Width=60 pixel: 成功した地点:B, C, I, L, N, R.

Width=70 pixel: 成功した地点:C, I, N, R.

Width=80 pixel: 成功した地点:C, N, R.

もっとも成功数の多かった結果で6/20という結果しか得られなかった．比較手法では対応付けが成功したほとんどの地点は，大きな変化のない比較的な単純なシーンであり，提案手法の有効性が確認できた．

SIFT 特徴点のマッチングのみによる画像間対応付けとの比較

提案手法で撮影位置が大きく異なり、街並変化や遮蔽生じる画像間において安定して対応付けを行うために、段階的なマッチングを行っている。そこで、比較のため SIFT 特徴点のマッチングのみで画像間の対応付けの結果を行った。実験では第 2 段階と同じ方法で対応付けを行う。ただし 1 つの入力画像に対してマッチングを行う過去画像は位置情報の誤差を考慮し前後 15m の距離が撮影されているであろう 25 フレームとする。使用する画像は提案手法の有効性を確認した同じ 20 地点の画像から 1 地点につき入力画像を 10 フレーム選択し、計 200 フレームとした。マッチングの結果 200 フレーム中 73 フレームの画像において対応付けに成功し、その精度は 36% であった。提案手法の対応付け精度 86% と比較し、提案手法においてマッチング精度の向上が確認できた。結果例を図 34 に示す。図 34(a)(b)(c) は対応付け成功例、(d)(e)(f) は対応付け失敗例である。失敗の原因は 1 対 1 の画像間では重なりが少ないため、大きな遮蔽や街並の変化が生じると対応関係がとれなくなるためと考えられる。

次に計算時間であるが、本実験では CPU : Intel Core2 Duo 3.00GHz , メモリ : 2GB の環境で行った。25 フレームとのマッチングに平均 80sec がかった。これに対し、提案手法では 1 つのシーケンスとのマッチングを行うために平均 33sec 必要となる。フレーム数は車両速度に応じて増減する可能性もあるため、提案手法の 2 段階マッチングの計算時間は大量画像処理を行う際に現実的な計算時間であると言える。

動的計画法による対応付けとの比較

文献 [8] で用いられてる動的計画法 (Dynamic Programming) による画像間の対応付けを行った。動的計画法は系列になってるデータ間の系列の部分的な伸縮を考慮しながら類似度を計算しマッチングを行う手法である。本実験では、撮影画像から Sobel operator によりエッジ抽出を行いエッジ画像を作成した。さらに、過去、入力エッジ画像には提案手法の第一段階のパノラマ画像マッチングと同様にガウシアンフィルタをかけ、正規化相互相関により画像間の相関度を得る。また、撮影画像は第 2 段階のマッチングと同じマスク領域を用いた。使用する画像



図 35 SIFT 特徴点のマッチングのみの対応付け結果

地点	結果 (成功)	地点	結果 (成功)
地点 A	24/25	地点 K	20/25
地点 B	22/25	地点 L	13/25
地点 C	14/25	地点 M	21/25
地点 D	12/25	地点 N	14/25
地点 E	24/25	地点 O	15/25
地点 F	8/25	地点 P	9/25
地点 G	16/25	地点 Q	16/25
地点 H	18/25	地点 R	23/25
地点 I	14/25	地点 S	24/25
地点 J	19/25	地点 T	21/25

表 5 動的計画法による画像間対応付け結果

は 1 シーケンスを 25 フレームとし、同一の 20 地点から 1 地点につき入力画像 1 シーケンス、過去画像 1 シーケンスの計 500 組の対応付けを行った。対応付けの結果、500 組中 346 組の対応付けに成功し、その対応付け精度は 69% となった (表 5)。提案手法の対応付け精度 86% と比較し、提案手法において精度向上がみられた。しかしながら、動的計画法では非線形な伸縮を可能とするため、特徴量の工夫を行えばさらにマッチングの精度を向上させることは可能である。しかしながら、地点 F, P で生じている他車両による遮蔽や、地点 G, N, Q で生じている建物の構造・テクスチャの大きな街並変化では画像間の相関度が低く、DP 距離が大きくなり対応付け精度が低下する。DP 距離が大きい場合は、動的計画法ではマッチング精度が落ちることは文献 [31] でも確認されている。

疎に観測された撮影画像間では重なりあう領域が小さいため、動的計画法で行う 1 対 1 の画像間のマッチングではこれらの変化がマッチングに及ぼす影響は大きい。そのため、重なりあう領域自体を増やすパノラマ画像が、このような変化が生じてる疎な撮影画像間の対応付けにおいて、より安定したマッチングを実現可能である。また、より安定したパノラマ画像マッチングを可能とする画像補正

を行うことでマッチングの精度向上を図っているが、補正の失敗によるマッチングの失敗も生じる。また、単純な投影による歪み(隣り合う短冊画像間)も生じるが、このようなパノラマ画像生成により生じるエッジは削除するため影響は少ない。また計算時間であるが、提案手法と同様の環境下(CPU: Intel Core2 Duo 3.00GHz, メモリ: 2GB)で行った場合、1シーケンスで約95secの時間を要した。提案手法の処理時間が33secであったことから、提案手法の計算コストは動的計画法と比べて低いことがわかる。

3.5 まとめ

本章では、システム実現のために異なる時期に多数の車両・カメラから撮影された画像間の対応付けを行う手法を提案した。提案手法では、特性の異なる2段階のマッチングにより街並の変化や遮蔽が生じている画像間の対応付けを実現した。第1段階では、撮影位置の違いによる重なり小さい画像間の安定したマッチングを可能とするため、時系列画像をつなぎ合わせて作成したパノラマ画像によるマッチングを行った。さらに、カメラ向きの違いや車両揺れに対応する画像補正を行うことで相似したパノラマ画像を作成した。第2段階では、パノラマ画像の対応付けの結果をもとに特徴点による対応付けを行い、さらに、マッチング精度の向上を図った。以下に今後の課題をあげる。

パノラマ画像の改善

パノラマ画像を生成する文献[13]ではラインスキャンにより得られる短区間の短冊画像をつなぎ合わせているため、得られるパノラマ画像は滑らかである。一方、本研究では低いフレームレートと幅の広い短冊画像を用いてパノラマ画像を生成している。また、単純な画像の平行化で得られるパノラマ画像上では段差が生じてしまう。この段差は、カメラから観測物体までの距離が既知であれば修正できることが知られているが(文献[32])、広大なシーンすべてにおいて3次元情報を得ることは困難である。この段差を無くすためには、一般のパノラマ画像生

成で行なわれているように，画像間の特徴点对応付けに基づいた画像変換（文献 [26] など）が必要である．

しかし，過去・入力画像のパノラマ画像作成において同様の変換が行なわれる保証がない場合，両画像のマッチングにおいて重要な情報となる強いエッジを示す直線の位置がずれてしまう可能性もある．よって，こうした画像変換の単純な適用には慎重な検討が必要である．また，提案手法では撮影を行なった車線の違いによる画像の変化を拡大・縮小のみで近似的に表現したが，実際には透視投影による歪みが生じ対応付けの失敗を招く恐れがある．この歪みに対処する画像モザイクング手法も提案されているので（[33] など），今後はこうした手法の導入について検討が必要である．

また本システムは大量のパノラマ画像を生成することが必要となる．各フレームを高速に安定してつなぐ手法として画像間の重複してる枚数を基にキーフレームを識別する文献 [34] が提案されている．地図更新システムを実現する際は，このような高速化の手法の導入も必要となる．

正確な車両揺れへの対応

フレームレートが低く，さらに街路樹などによる遮蔽で特徴点の追跡が困難であるため，特徴点追跡を利用して得られる上下揺れ補正や抽出線の決定に失敗してしまう場合がある．そのため，一般的なカーナビゲーションでも利用されている，ハンドルの回転角度やジャイロセンサーから得られる車両の傾きから画像補正を行う必要がある．

照明条件の著しい変化への対応

天候条件の大きな変化へ対応するためにエッジ抽出の閾値を調節する必要がある．ただし，多数データの収集が可能な目標システムでは，雨天時のように照明条件の変化に限らず見えが大きく変化する際には変化検出を行わないことで問題回避できる．

照明条件の著しい変化が生じても対応付けを成功させるために、パノラマ画像からエッジ特徴量を抽出する際の閾値は自動的に調節されるべきである。しかしながら、本システムでは異なる車両から同じ場所の多数の画像から得られるため、全ての撮影画像を必ずしも用いる必要はない。そのため、提案手法では撮影された入力画像が過去画像とのマッチングに適しているかどうか、またはマッチングの結果が信頼できるものであるかを分析する必要がある。

SIFT 特徴点の対応付けによるマッチングの高速化と精度向上

SIFT は回転・照明条件・スケール変化に頑健な特徴量といわれているが、その計算コストは膨大であり、実用化を目指す際には高速化手法の導入の検討が必要となる。文献 [35] では、SIFT のキーポイント検出と特徴量記述の処理を Box フィルタと積分画像を用いた近似によって高速化を図っている。また、文献 [36] では Graphics Processing Unit(GPU) を用いたハードウェア処理により高速化を行っている。また、精度向上と高速化を目指し学習を取り入れることで高速かつ高精度化を目指す手法が提案されている。文献 [37] では、Randomized Trees[38] を用い、事前にキーポイント分類のための決定木を作成し、対応点探索を決定木のトラバーサルによる分類とすることで高速かつ正確なキーポイント分類が可能となる。また、学習データにアフィン変換した画像を用いることで、射影変換への頑健性を高めている。

4. 複数車載カメラ画像からの変化領域抽出による街並の変化検出

4.1 はじめに

地図情報更新のためには、街並の変化を検出する必要がある。また、ここで述べる街並の変化とは第 2.3 節で定義した現在のデジタル地図に収録されている固定された人工物（建物・看板）の構造とテクスチャの変化である。本章では、現在は人手により行われている同一地点を異なる日時に撮影した 2 画像間を比較し、街並の変化を検出する手法を提案する。

まず、第 3 章の対応付けにより得られた異なる日時・撮影条件下で同一地点を撮影した画像例を示す。図 36 は、街並の変化はないが、撮影位置と撮影条件の変化による見えの変化と他車両などの移動物体、さらに街路樹の形状の変化といった街並以外の変化が存在する。図 37 は、図 36 と同様の見えの変化と他車両に加えて白い線で囲まれた領域に検出すべき街並の変化（建物の構造、テクスチャの変化）が生じている画像である。

このように撮影画像中には街並の変化以外にも、撮影条件の違いにより生じる見えの変化、移動物体による変化や遮蔽が生じる。これらの撮影画像から街並の変化をその領域のサイズに関わらず、漏れなく検出することが求められる。しかしながら、画像中にどのような建物や看板があるかといった事前知識はなく、また対象となる街並の構造やテクスチャは多種多様である。下記に、街並の変化検出を行うために解決が必要な問題点をまとめる。

問題 (1) 撮影条件の違いによる見えの変化が生じる。(撮影位置・カメラ向き・照明条件など)

問題 (2) 街並以外の変化が生じる(他車両, 通行人, 形状の変化しやすい街路樹・のぼり旗)

問題 (3) 検出対象(街並)に関する事前知識がない。(建物・看板の色・形状は多種多様)



図 36 街並に変化はなく撮影条件・移動物体などの対象外の変化を含む撮影画像例

問題 (1) は第 3 章でも問題となった撮影条件による変化である．街並の変化検出においてもこの問題は同様に解決する必要がある．

また，画像中に存在する全ての物体の変化検出を目的とする一般的な画像の変化検出問題とは異なり，提案手法では変化を検出する対象となる物体が限定されている．しかしながら問題 (2) のとおり，検出対象ではない変化が存在するためこれらの識別が必要となる．

問題 (3) では，撮影画像中にどのような形状とテクスチャを持った検出対象が存在するのかが不明であり，さらに多種多様であるため全てを事前に学習しておくことは不可能である．もし検出対象に関する事前知識があれば，あらかじめ存在する検出対象のみを限定して照合を行うことも可能である．さらに検出した変化がどのような物体の変化であるかといった識別の際に，物体認識などで行われている物体の特徴を学習しておく手法や，事前知識をもとにした制限を設けることが可能であるが，問題 (3) のため提案手法では事前知識なく変化の検出を行わなければならない．

提案手法と同様に画像間の比較による街並の変化検出を目的とした手法が提案されている．文献 [3] では，衛星写真から物体のエッジを線分として抽出しそれらの比較を行っているが，エッジのみでは街並の構造やテクスチャの大きな変化

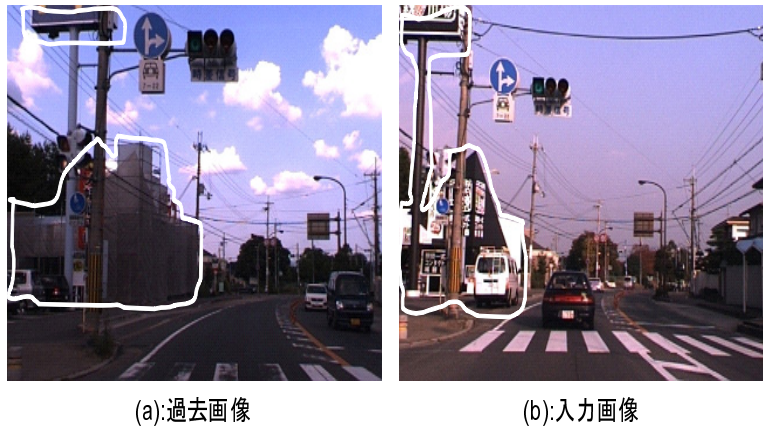


図 37 街並の変化とその他の様々な変化を含む撮影画像例

を捉えることは可能であるが、テクスチャのみの変化は捉えることが困難である。

また文献 [8] では、車載全方位カメラ画像から画像全体の RGB 値を特徴ベクトルとし、主成分分析によって低次元化したものを特徴量としてそれらの相違度を比較している。この文献 [3], [8] で行われている、画像全体から抽出される特徴量を大域特徴量と呼ぶ。大域特徴量は、特徴量が画像 1 フレームにつき 1 つであるため、大量画像を処理する際に計算コストが低減できるという利点もあるが、画像の一部に遮蔽が生じている場合は同一の特徴が得られない。そのため、他車両などの移動物体が多く存在し、さらに撮影条件の違いから見えの変化が著しい提案システムの撮影画像には不向きである。また画像全体を特徴として、画像間の相違度の閾値処理による街並の変化検出を行うと、建物や看板の構造やテクスチャなどの変化が生じている領域が大きな街並の変化は検出可能であるが、変化が生じている領域の小さい街並の変化は検出が困難である。

4.1.1 研究目的

本論文では、事前知識なく街並の変化が生じている領域の面積や種類を問わない街並の変化検出実現を目的とする。

そこで地図情報更新に必要な街並の変化を検出するため、本論文では過去画像と入力画像の特徴点マッチングと 2 画像間の画素単位の密なマッチングを行う。

まず初めに、局所特徴量である SIFT 特徴点による対応付けにより、画像間の対応関係を推定するとともに画像全体に渡る大きな街並の変化を検出する。次に、推定された対応関係を利用した画素単位の密なマッチングによる、変化領域抽出を行う。この変化領域は、入力画像中の注目画素と対応する過去画像中の画素をマッチングにより探索し、対応する画素が存在しなければ、注目画素を「変化」として判断することで抽出される。

特徴点間の対応付けによる変化検出と変化を領域として抽出する画素単位の密なマッチングにより、生じた街並変化のサイズに関わらず検出することを可能とする。さらに、部分的な変化に対してはオペレーターによる確認時に画像のみを提示する場合に比べ、画像中の変化領域を示すことでオペレーターによる確認の手間を削減する。

特徴点間の対応付けによる変化検出では、画像全体に渡る大規模な街並の構造やテクスチャの変化を検出することを目的とする。そして、画素間のマッチングによる変化領域抽出では部分的な物体の構造・テクスチャの変化を抽出することを目的とする。そのため、撮影条件の違いから生じる見えの変化が生じている同一物体のマッチングを可能としないといけない。そこで、特徴点对応付けではこれらの撮影条件の変化に頑健な特徴点の抽出を行う必要がある。また、画素間のマッチングでは入力画像中のある1つの画素に対応する過去画像中の画素を見つけるために、過去画像中の全画素とのマッチングを行うと、撮影画像サイズに比例して計算コストが高くなる。また、画像中の似たテクスチャをもつ他の物体との誤対応が生じる可能性がある。そこで、提案手法ではステレオマッチングで利用されている画像間の対応関係を推定し、拘束条件として画素間のマッチングに用いることで、計算コスト削減とマッチングの精度向上を目指す。

このようにして得られた変化には、対象物体(街並)以外の変化も含まれる。そこで本論文では、変化が生じていると判断された画像や、画像中から抽出された変化領域から街並の変化を識別するため、3フレームの過去画像と1つの入力画像から得られた変化抽出結果をもとに多数決的に街並の変化を検出する。図36,37でも示したとおり、抽出された変化には対象となる街並以外の変化(街路樹、のぼり旗、他車両、通行人など)が多数含まれる。しかしながら、前節の問題(2),(3)

で述べたとおり街並の形状やテクスチャは多種多様でありこれらの事前知識なく、街並の変化のみを識別することは困難である。そこで、第 2.3 節でも述べたとおり、変化を「変化が生じる時間」により下記の 3 つのクラスに識別する。

常時変化：歩行者や車両といった移動物体，形状が変化しやすい街路樹・のぼり旗

短期変化：一時的に存在する駐車車両や工事の看板など

長期変化：固定された人工物の構造とテクスチャの変化

この街並変化を含む長期変化を識別するため、提案手法では複数 (3 フレーム) の過去画像と 1 つの入力画像のマッチングにより得られた変化抽出の結果で多数決を行う。3 フレームの過去画像において、変化と判断された画像または領域を街並の変化として検出する。また、過去画像間は異なる日時・撮影条件に撮影された画像であるが、街並の変化は生じていないものとする。

4.1.2 関連研究と位置付け

提案手法では街並の変化検出のために画像間のマッチングを行う。画像間のマッチングは、動作認識、3 次元復元、画像探索など様々な用途に用いられている。そこで、マッチングの手法ごとに、(a) テンプレートマッチング、(b) 特徴点間のマッチング、(c) 画素間のマッチングに分類し以下にそれぞれの詳細を述べる。

(a) テンプレートマッチング

画像の一部からテンプレートとなる画像を作成し、もう一方の画像からテンプレートと相関が高いまたは相違度が低い領域を探索することによりマッチングを行う手法である。文献 [39] では、カメラ位置やシーン変化が生じている 2 つの異なる時系列画像間をテンプレートマッチングにより対応付けることで動作認識を実現している。しかしながら、この時系列画像間の視差と変化は非常に小さく歪みも小さいため、対象物体 (人) の見えの変化は小さい。しかしながら、視差や歪みが大きい場合は同一の対象物体であっても見えの変化が大きいため、入力画像から作成したテンプレートによるマッチングは困難である。

(b) 特徴点間のマッチング

画像中から特徴点を抽出し，それらの対応付けによりマッチングを行う手法である．文献 [40] では，3次元復元のための同時刻に同じシーンを撮影した多視点画像間でのマッチングが行われている．同時刻の撮影画像であるため，画像間に移動物体・照明条件の変化はないが，撮影位置の違いによる見えの変化が生じた画像間のマッチングを可能としている．また，Photo Tourism[41] では，画像に様々な変化（歩行者や照明変化）が生じていても，これらの異なる領域を除き，多視点画像間の特徴点マッチングを可能としている．しかしながら，提案手法では対象物体の変化とその他の変化を識別する必要がある．特徴点間のマッチングは遮蔽やシーン変化に頑健ではあるが，特徴点間のマッチングのみで街並の変化を全て検出することは困難である．

(c) 画素間のマッチング

文献 [42] では，ステレオカメラで撮影された2画像間（左右の画像）の画素間でマッチングを行うことで，左画像の画素と対応する右画像との間の視差を推定している．同時刻に撮影された画像であるため，撮影位置の違いによる見えの変化は生じているが，画像間に照明条件の変化やシーンの変化はなく，画素の特徴量として輝度値を用いマッチングを行っている．さらに文献 [43] では，画素の特徴量に照明条件・回転・スケール変化に頑健な SIFT 特徴量を用い，マッチングが困難な異なるシーン中から構造が相似した物体（建物や車両などの）の探索を可能としている．これらの画素間のマッチングでは，1つの注目画素に対して画像中の全画素とのマッチングを行っているため，撮影画像のサイズに比例して計算コストが高くなり，大量画像の処理には向かない．また，文献 [43] では似た構造の他の物体（形状の異なる車両など）間の識別は目指していないため，テクスチャのみの変化は変化として識別が困難である．

提案手法では (b) 特徴点間のマッチングと (c) 画素間のマッチングによる変化領域抽出を行うことで，生じた街並変化のサイズや種類（構造やテクスチャなど）に関わらず検出することを目的とする．得られた変化領域から，画像解析により検出対象である長期変化を識別する処理を行うことも考えられるが，そのために

は解決困難な以下の問題点がある。まず、画像中にはどのような人工物が存在するかといった事前知識がなく、さらにその人工物の形状や色は多種多様である。そのため、道路標識など形状や色が既知である物体の特徴をあらかじめ学習しておき、撮影画像中から探索するような手法(文献[44],[45])は用いることができない。また、常時変化である歩行者や車両といった移動物体を画像中から抽出する文献[46]が提案されている。しかしながら、提案手法で得られる撮影画像は車載カメラから得られた動背景の画像であるため、背景差分など静止した背景から移動物体を切り出す文献[47]や、背景モデルを作成するような文献[48]は使用できない。

そこで、車載カメラのような動背景からの移動物体の抽出を行う手法が提案されている。文献[49]では特徴点を追跡することで移動物体の抽出を行っているが、特徴点の抽出が困難な物体や背景と似た特徴をもつ移動物体などは抽出できない。また、手法([50],[51],[52])では物体の特徴をあらかじめ学習しておき、それらの特徴を画像中から探索することで歩行者や車両といった特定物体の抽出を行っている。しかしながら、検出対象のスケールの違い、遮蔽、また撮影向きの違いに対応するには大量の学習や検出ウィンドウをスケール変化させながら複数回ラスタスキャンするといった高い計算コストが必要であり、多様な撮影条件のもと観測された大量画像の処理が必要な本システムでは、これらの手法の適用は現実的ではない。システムとして、大量画像の処理を効率的に行いかつ、対象外となる変化が検出されたとしても、対象となる街並の変化を検出漏れすることなく変化領域として抽出することがデジタル地図情報の正確性を保つには重要である。

4.1.3 提案手法の流れ

提案手法の流れを図38に示す。まず特徴点の対応付けによる画像全体に渡る街並の変化の検出と画像間の対応関係の推定方法について述べる。1つの画素に対し全画素とのマッチングを行うのでは、画像サイズに比例して計算コストが高くなり、大量画像を処理するには不向きである。また、似たテクスチャをもつ異なる物体に対して誤対応が生じる可能性がある。そこで、既存文献[43]の問題点を解決するため、提案手法では、画像間の対応関係を拘束条件として用いること

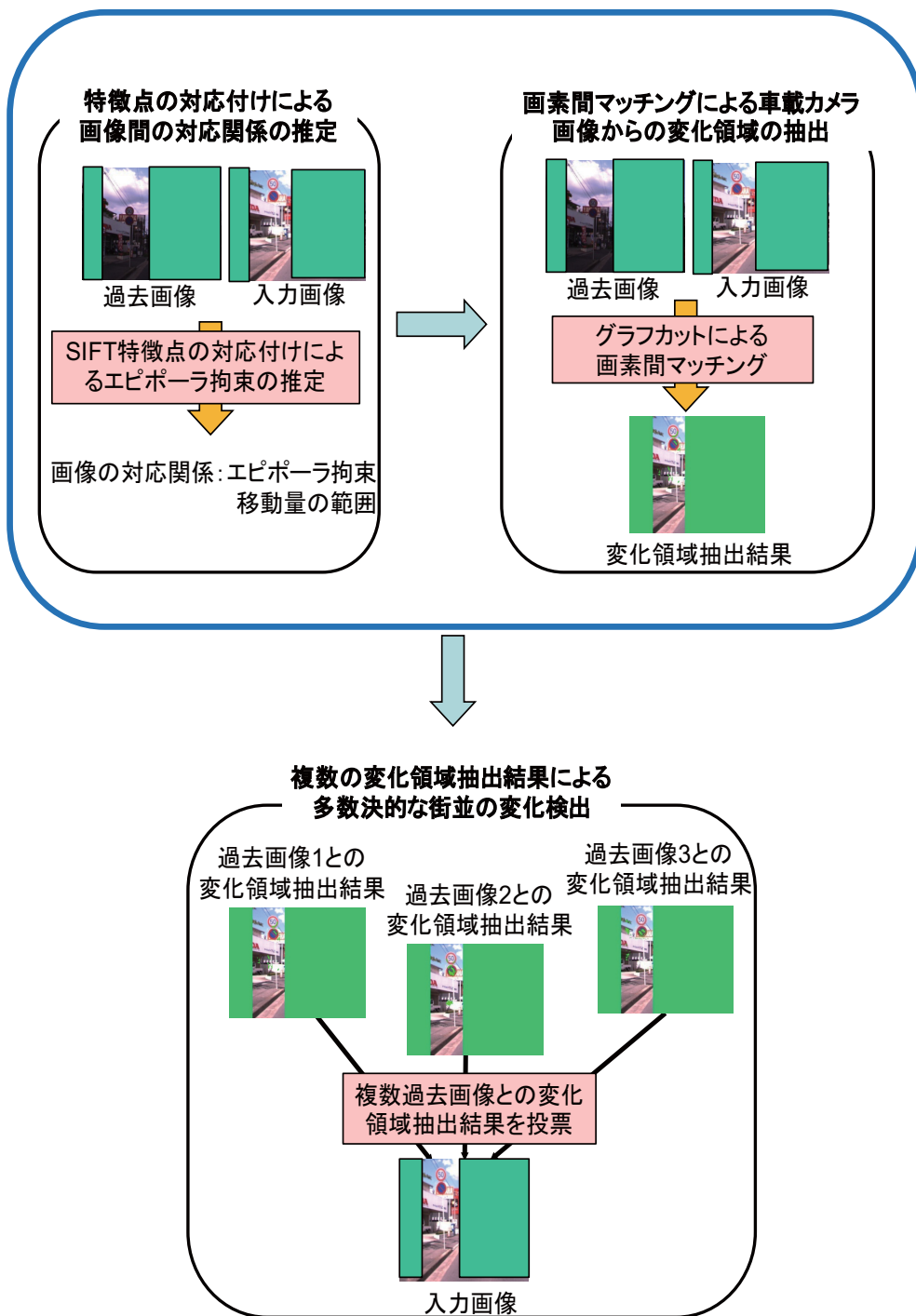


図 38 提案手法の流れ

で計算コストの削減と誤対応の低減を目指す。

この拘束条件には、多視点画像間のマッチングを行うステレオマッチングで用いられている仮定と同様の拘束条件を適用する。ステレオマッチングでは、3次元中のある点を複数の画像平面上に投影したときに生じるエピポラ拘束と呼ばれる関係性を利用し、異なる視点間の画像上の対応点を効率的に探索する。提案システムで得られる画像間においても同一地点をもっとも近い位置から撮影した画像であるため、画像中には同一物体が存在すると仮定できる。よって、撮影時刻は異なるが画像中にはステレオマッチングと同じ関係性が成立すると仮定する。この仮定を利用し、画素間マッチングの際に探索範囲を限定することでマッチングの精度を向上させ、計算コストの削減を図る。この拘束条件は、複数の特徴点の対応関係により得られる。また、この仮定が成立しない、つまり特徴点の対応関係の数が十分に得られない場合は画素間のマッチングが困難な大きな変化が生じていると判断し、変化領域抽出は行わない。

次に、この画像間の対応関係を利用した画素間マッチングによる変化領域抽出手法の流れについて述べる。提案手法では文献 [43] と同様に画素の特徴量として見えの変化の原因となる回転、照明条件、スケール変化に頑健な SIFT 特徴量に加えて、RGB のベクトル角度を用い、画素単位のマッチングを最適化問題として扱う。文献 [43] では、確率伝搬法 (belief propagation)[53] によって相似したシーン探索のためのマッチングを行っている。提案手法では3次元復元のための対応点探索 ([54], [55],[42]) する際にも利用されているグラフカットを用いて、画素単位の密なマッチングを行う。提案手法ではエピポラ拘束を用いて、グラフカットによる画素単位の密なマッチングを行う。この際に比較されるのは、マッチングの第2段階と同様の画像のマスク領域内の部分である。なぜならば、画像はカメラを進行方向に向けて撮影されているため、対象となる街並は左端に多く存在し、さらに中心部分は車両などの移動物体が存在する確率が多いためである。マスク領域の切り出し位置は、対応付けの第1段階においてパノラマ画像作成のための短冊画像を切り出し位置と同一とする。また、その切り出し幅は実験的に固定値とし、画像サイズ (VGA サイズで 200pixel) によって変動する。

最後に、異なる日に撮影された3フレーム過去画像と入力画像において変化抽

出を行う。そして得られた変化抽出結果をもとに入力画像に対し投票を行う。投票では2つ以上の重畳が起きた領域を街並変化とする多数決原理に基づき街並の変化を検出する。また、2つ以上の過去画像との画像間の対応関係が推定できなかった場合も、画像中に街並の変化が生じていると判断する。

4.2 画像の対応関係を用いた画素間マッチングによる車載カメラ画像からの変化領域抽出

撮影条件の変化が生じている車載カメラ画像間から画像の対応関係を推定し、効率的な画素単位の密なマッチングを行うことで変化領域を抽出する。変化領域抽出手法の流れについて図 39 に示す。まず、計算コストの削減とマッチングの

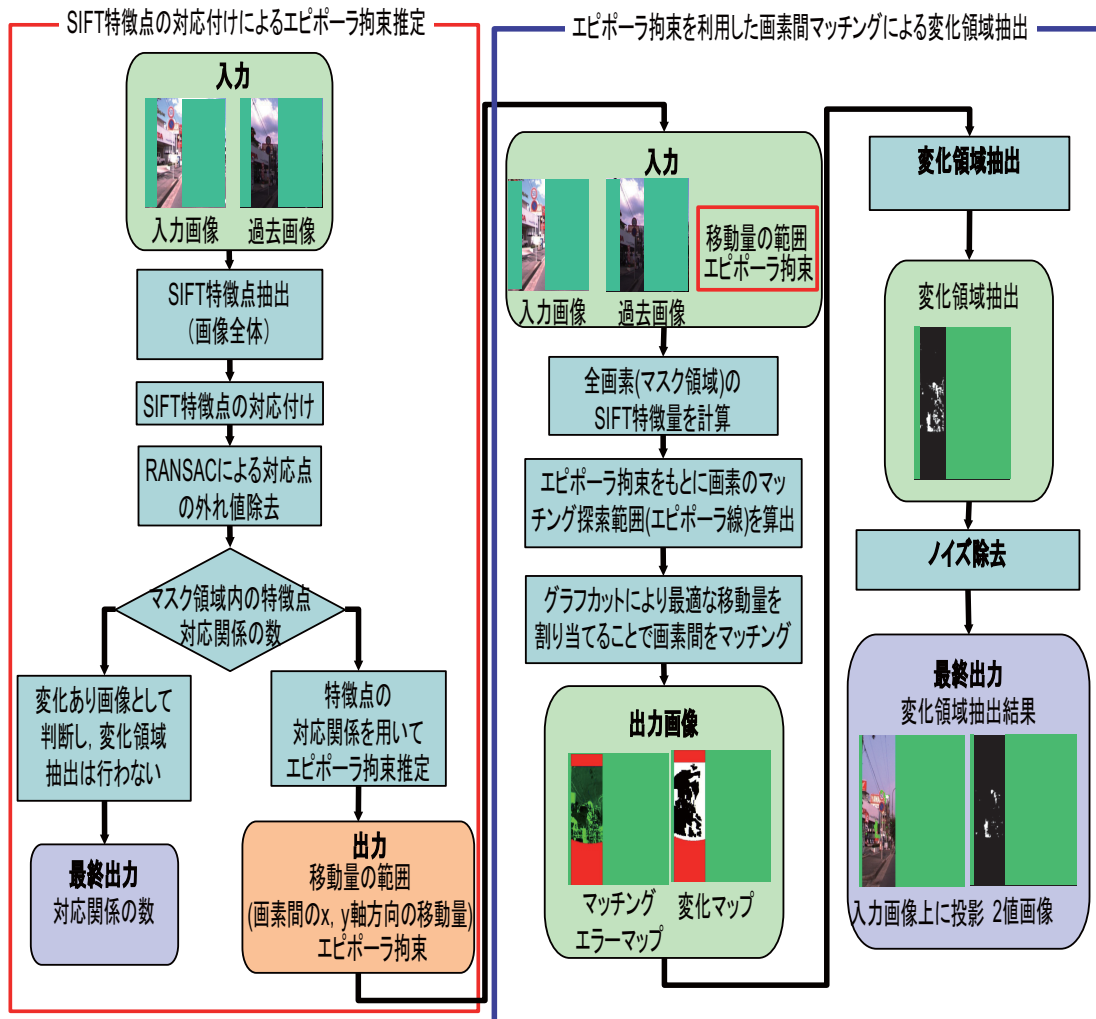


図 39 画像の対応関係を用いた画素間マッチングによる車載カメラ画像からの変化領域抽出手法

精度向上のために、画像間から拘束条件を推定し、それを利用することで画素間マッチングの範囲を限定する。この拘束条件には多視点画像間のマッチングを行うステレオ視で利用されているエピポーラ拘束を利用する。エピポーラ拘束は画像間の特徴点の対応関係により得られるため、提案手法では SIFT 特徴点の対応付けによりエピポーラ拘束を推定する。また、この特徴点の対応関係は画像全体から得られるが、マスク領域内の対応関係の数が閾値以下であった場合は安定したエピポーラ拘束推定が困難となるため、画像中に大きな変化が生じていると判定し、変化領域抽出は行わない。さらに3つの過去画像中、2つの過去画像との特徴点の対応付けにおいて対応関係の数が閾値を下回った場合は、街並の変化が生じているとして検出する。

次に、得られたエピポーラ拘束を用いてグラフカットによる画素単位の密なマッチングを行う。グラフカットによる画素間のマッチングでは、画素間の移動量 (x 軸, y 軸方向の移動量) を変化させながら評価関数のエラー値が最も小さくなる画素を探索することで、2画像間の画素が対応付けられる。このマッチングにより下記が得られる。

マッチングエラーマップ：マッチングにより得られた画素間の特徴量 (SIFT 特徴量と RGB のベクトル角度) を比較した相違度を表している、緑色 (0-255) の値が大きいほど相違が大きい。

変化マップ：各画素にラベリングされた移動量の整合性の確認において、不整合と判断された (移動量のノルムが閾値以上) 画素を「1」、その他を [0] とした2値画像。

この変化マップで1かつマッチングエラーの値が閾値以上もしくはマッチングエラーの値が閾値以上の場合を変化画素として抽出する。また、全てのマップの赤色の画素は以下の原因などにより画素間のマッチングが行われなかった画素である。

- (1) エピポーラ拘束の推定失敗により、適切な探索範囲が得られなかった。
- (2) SIFT 特徴点の対応付け失敗および誤差により、適切な移動量の範囲が与えられなかった。

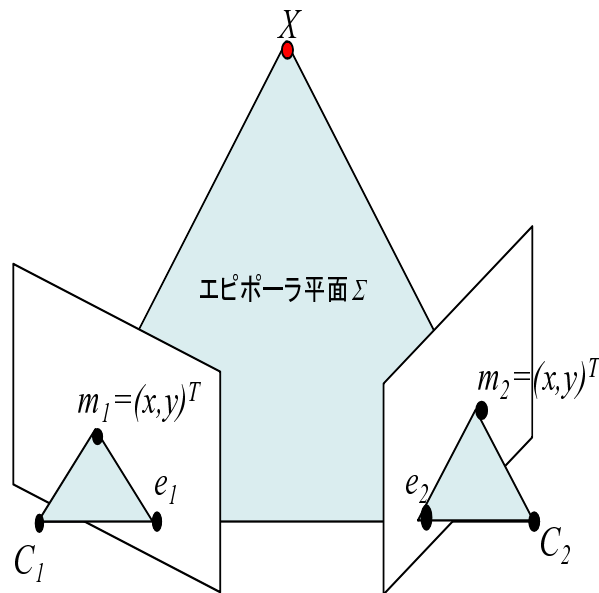


図 40 エピポーラ幾何

まず，提案手法ではマッチングの探索範囲はエピポーラ線上としている．そのため，エピポーラ拘束推定に失敗すると適切な探索範囲が得られずマッチングが行われない．次に，グラフカットによる画素間マッチングでは注目画素に対し，与えられた移動量に対応するエピポーラ線上の画素から対応づく画素を探索する．移動量の範囲は，エピポーラ拘束推定時に行われる SIFT 特徴点間の対応付けにより得られた対応関係の x, y 軸方向の移動量の平均値をもとに最小値と最大値を与える．この時に適切な移動量の範囲が与えられないと，マッチングが行われない画素が発生する．本節ではこのエピポーラ拘束の推定方法と，この拘束条件を利用した画素間の密なマッチングによる変化領域抽出手法について述べる．

4.2.1 SIFT 特徴点の対応付けによるエピポーラ拘束の推定

まずは，ステレオマッチングにも利用されているエピポーラ拘束 ([56],[57],[58]) について述べる．3次元上のある一つの点を複数のカメラによって2次元に投影したとき，複数の画像間に特有の幾何が現れる．このような幾何のことをエピポーラ

ラ幾何と呼ぶ(図 40)．3次元空間上のある点 X を視点 C_1 と C_2 のそれぞれの画像平面に投影したとき， X, C_1, C_2 は平面 Σ を定義し，この平面 Σ をエピポーラ平面と呼ぶ．画像上の点 m_1 と m_2 はこのエピポーラ平面上にあり，エピポーラ面と画像面の交線を l_1 と l_2 をエピポーラ線と呼ばれ，同様に視点 C_1 と C_2 を結ぶ直線と画像平面の交点 e_1 と e_2 をエピポールと呼ぶ．このように空間上の一つの点に対し，一つのエピポーラ平面が規定され，それぞれの画像平面とのエピポーラ線が定まる．エピポーラ線はエピポールを中心とした扇形の構造をなし，すべてのエピポーラ線はエピポールを通る．

このエピポーラ幾何により，視点 C_1 の画像上の点 m_1 と視点 C_2 において対応する点は必ずエピポーラ線 l_2 上に存在する．これをエピポーラ拘束と呼ぶ．ステレオマッチングではこの関係性を用いて異なる視点間の画像上の対応点を探索する．提案手法でも同様に，ある注目画素との対応画素の探索をこのエピポーラ線に限定することで，対応点探索の範囲を限定する．さらに探索範囲を限定することで，画像中に存在する注目画素と相似性の高い異なる物体との誤対応を防ぐことが出来る．

このエピポーラ線を求めるエピポーラ方程式は，一台目のカメラの画像点の点を $m = (x, y)^T$ とし，二台目の画像の点を $m = (x', y')^T$ とした時，文献 [56],[57],[58] では下記のように表す．

$$\tilde{m}^T F \tilde{m} = 0 \quad (12)$$

$$F = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \quad (13)$$

この行列 F は基礎行列とよばれる．カメラの内部パラメータが未知であり，画像座標系のみが得られる場合，この基礎行列は画像間の 8 点以上の対応関係から，最も小さい固有値に対する固有ベクトルとして求められる．

この対応点探索には特徴点ベースと領域ベースの 2 つの手法 ([59],[60]) が提案されている．特徴点ベースのマッチングでは，2 つの画像中からそれぞれエッジや極値などの特徴点を抽出し，特徴点ともっとも相関の高い特徴点を他方の画像中から探索する．また，領域ベースのマッチングでは与えられた基準点の近傍か

ら切り出して作成した領域をテンプレートとし，相関度の高い位置を探索する．これはブロックマッチングとも呼ばれる．領域ベースのマッチングは画像全体にわたって密な対応付けを行うことが可能であるが，画像間に変形や被写体の見えが大きく異なる場合はマッチングが成功しない．そこで提案手法では，遮蔽や見えの違いに頑健な特徴点ベースのマッチングにより対応関係を得る．

また提案手法では，この特徴点には Scale-Invariant Feature Transform (SIFT)[22] を用いる．SIFT[22] は，第 3.3 節でも述べたとおり特徴点の検出と特徴量の記述を行うアルゴリズムである．画像の回転・スケール変化・照明変化に頑健な特徴量を記述可能であり画像のマッチングや物体認識でも用いられている特徴量である．拘束条件の推定ではこの SIFT 特徴点を用いる．また，この SIFT 特徴点でも 3.3.1 節で行った SIFT 特徴量の外れ値の除去を行う．

安定したエピポーラ拘束の推定のために，SIFT 特徴点は画像全体から求める．SIFT 特徴点の対応付けを行った結果を図 41(a) に示す．また，正確な基礎行列の推定のためにさらに RANSAC[29] を用いて外れ点の除去を行う．その結果が図 41(b) である．

このエピポーラ拘束を用いて，画素間のマッチングにより変化領域の抽出を行うが，図 41(c) に示したとおり，マスク領域の広範囲にわたって街並の変化が生じている場合は特徴点の対応関係も得られず，さらに密なマッチングとなる画素間のマッチングは不可能である．そのため，このマスク領域内の特徴点の対応関係数が閾値以下であった場合は安定したエピポーラ拘束の推定と画素間のマッチングが困難であるため，変化あり画像として判断し変化領域抽出は行わない．また対応関係の数に対する閾値は次のように定める．まず，先述したとおりエピポーラ拘束を推定するためには最低 8 組の対応関係が必要である．そこで，安定したエピポーラ拘束が推定できるよう閾値はその 2 倍となる 16 組とする．

4.2.2 エピポーラ拘束を用いたグラフカットによる変化領域抽出

画素間のマッチングにより変化領域を抽出するが，画素の特徴量の単純な比較では撮影位置の違いや撮影条件の変化が生じているため，3 次元上の同一の点を投影した画素間であってもマッチングは困難である．そこで提案手法では画素間

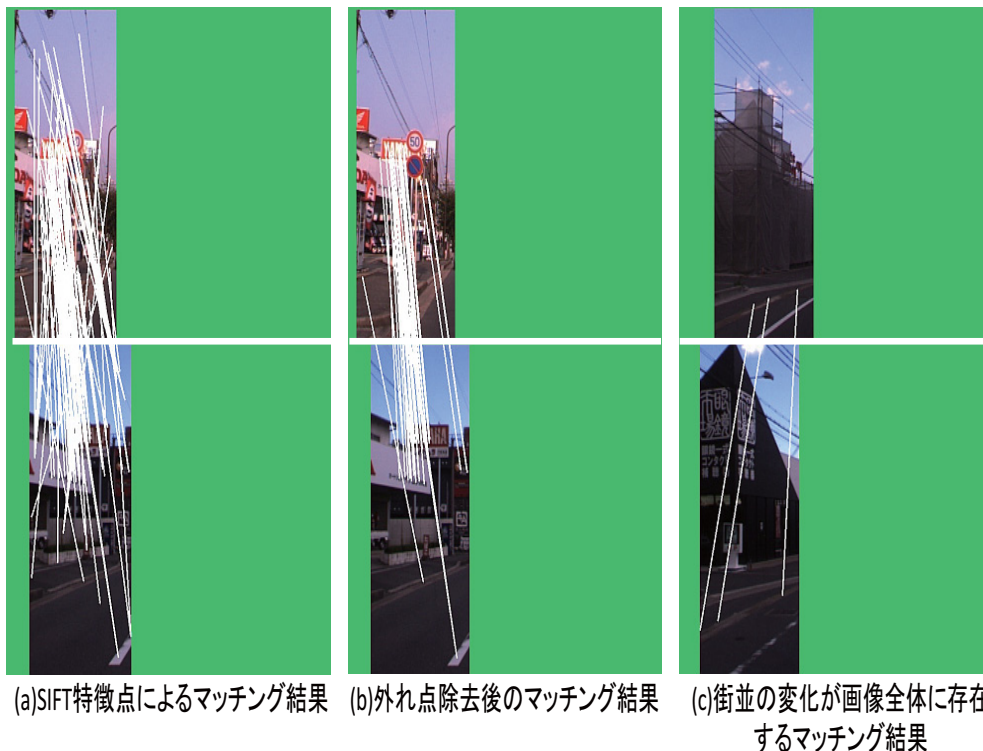


図 41 SIFT 特徴点の対応付け結果

のマッチングを、最適な画素間の移動量 (x 軸, y 軸方向の移動量) をラベリングするためのエネルギー最小化問題として捉える。

まず注目画素に対し、画素間の移動量を与える。注目画素と移動量から注目画素と同一の3次元上の点を投影したとされる画素を評価関数により評価することで、エラー値がもっとも小さくなる画素を探索する。このようにこの与えられた移動量と画像との整合性を評価することで、画素間のマッチングを実現する。このようにして得られた画素のマッチング結果から、エネルギー関数のエラー値と画素の特徴量の比較の値または、画素の特徴量の比較の値が閾値以上であった場合に注目画素が変化画素であると決定する。

さらに、マッチングを行うマスク領域は、第3.3節と同じマスク領域とする。これは想定するシステムの構成上、対象となる街並は左端に多く存在し、さらに画像中心部分の道路領域は車両などの移動物体が存在する確率が高いためである。

(1) グラフカットアルゴリズムについて

コンピュータビジョンの問題をエネルギー最小化問題として捉える手法が提案されている．提案手法と同様に画像を画素単位で輝度，セグメンテーションなど各画素のとりうる値をラベルとしてラベリングしていくこととなるが，実画像にラベリングを行う場合には，ノイズなどの影響により最適なラベリングを行うことは困難である．そこで，文献 [61] では最大流最小切断定理を利用してエネルギーの最小化方法を提案している．文献 [61] では，問題設定に合わせた有向グラフを作成し，そのグラフの最小切断を求め，エネルギー関数の最小化を行う．この最小切断は最大流問題の線形計画法に相対するものであり，最小切断は最大流と一致し，多項式時間で効率的に解くことが可能となる．この最大流最小切断定理に基づいた手法はステレオマッチング ([62], [63], [64])，画像復元 ([54], [64])，セグメンテーション ([65], [66]) などに応用されている．

以下に，グラフカットについて述べる．ラベリングとは画像 I の各画素 p に対して，どのようなラベル L_p をつけていくかを解く問題であり，このラベリングを何に対して行うかは求める問題により異なる．ステレオマッチングでは視差，セグメンテーションでは各物体，提案手法では各画素の移動量に対してラベリングする．このラベリング問題を解くアルゴリズムとしてエネルギー関数 E を定義し，これを最小化する方法がある．グラフカットアルゴリズムでは，定義したエネルギー関数 E に基づき最大流最小切断定理によりエネルギーが最小となる最適な解を求める．

まず，重み付き有向グラフを $G = (V, E)$ と定義すると，頂点の集合 V はノード，頂点間に存在する有向辺の集合 E はエッジとなる．コンピュータビジョンの問題では，各画素やボクセルなどがノードとなり提案手法では画素とする．次に各画素のとりうる値をラベル L と定義する．辺 $(u, v) \in E$ には，それぞれ実数値の重みが付与されており，それを $p(u, v)$ と表す．次に，頂点の集合 V の中から，source $s \in V$ と sink $t \in V$ と呼ばれる特別なターミナル(ラベル)を選択する．エッジはノード間の関係を表現しており，隣接する画素との関係を表したエッジを n-link，各画素と source s ，sink t との関係を表し，2つをつなぐエッジを t-link と呼ぶ．n-link のコストは，周辺画素との連続性を表現したペナルティ関数によ

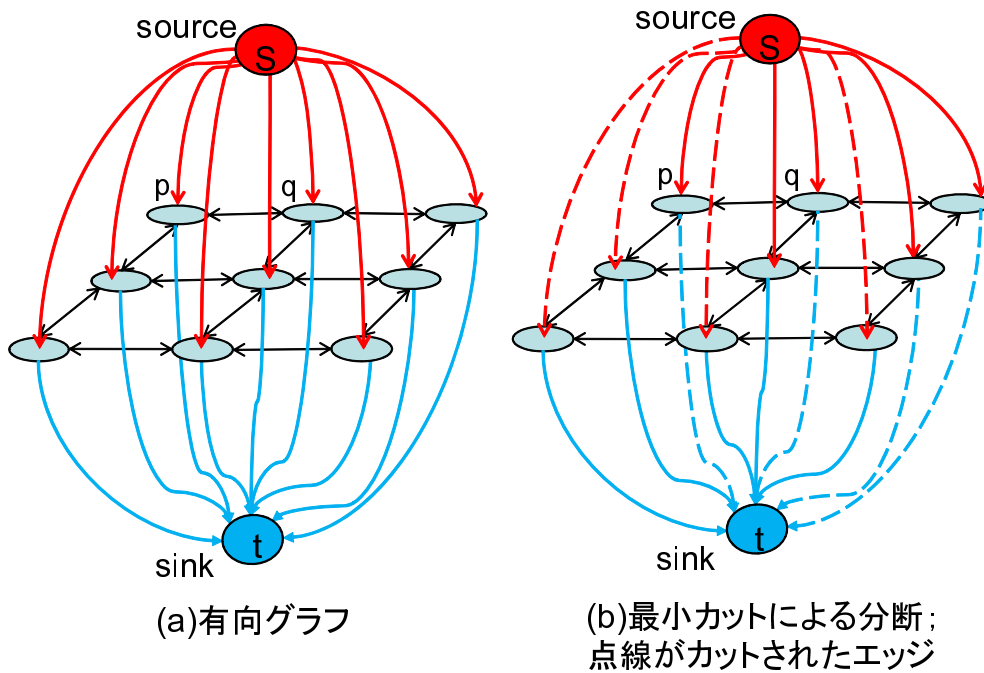


図 42 グラフカット例.

り決定され、 t -link のコストは各画素がそのラベルである確率を表したペナルティ関数により決定される。このように作成されるグラフをネットワークグラフと呼ぶ(図 42(a))。次に重み付き有向グラフ G の s, t についての切断とは、頂点集合 V を $s \in S$ と $t \in T$ を満たすように部分集合 $V \subset S$ と $T = V/S$ の 2 つのグラフへと分割することを指す。この切断のコストとは source s から sink t へ向かう辺の重みの総和

$$P(S, T) = \sum_{u \in S, v \in T} p(u, v) \quad (14)$$

である。この時、sink t から source s への辺の重みは総和に加算されない。この分割の際に切断されるエッジを $s-t$ カットと呼ぶ。その切断されるカットのうち S から T へ接続されているエッジのコストの総和を $s-t$ カットの容量という。このカットの容量である分割コストが最小になる分割の仕方を求めることが最小切断問題である。この最小切断は、最大流最小切断の定理を用いることによって、ネットワークグラフの最大流と一致する。辺の重みが全て非負であるときに最大

流問題を効率的に解く手法が多数提案されており，augmenting path[67] や push-relabel[68] がある．

2 値画像のノイズ除去問題を最適化問題としてグラフカットを適用した際の，グラフと画像の関係を示す．ノイズ除去問題は画像の各画素をノイズ領域とそれ以外の物体領域に分離する問題である．この場合，ラベルは $L = 0, 1$ となる．まず，各画素が頂点に対応する．さらに画素に対応する頂点とは別に，特定の画素に対して人手によって source s と sink t が用意され，source は物体に，sink はノイズに対応する．この与えられた値は seed とよぶ．最小切断の結果，頂点 i が $i \in S$ となれば，それは画素 i が物体と推定されたことを意味し， $i \in T$ となればノイズと推定されたことを意味する．最小切断は，各辺に付与される重みに対して最適な分割結果を与える．従って，本来達成したいノイズ問題 (に対して定義されたエネルギー関数最小化問題) の解が得られるようにグラフの重みを定めることができれば，大域的最適解が求まることになる．しかしながら，常にエネルギー関数の定義と重みの定義の対応付けが可能となるわけではない．ノイズ除去問題のような 2 値最適化問題の場合は，平滑化項が L-1 ノルムや L-2 ノルムなどであれば (実際には，厳密な数理的な表現がある) グラフカットの適用が可能となることが証明されている．

グラフカットは 2 値最適化問題に対しては，上記のとおり適用可能である．一方，多値大域最小化問題にグラフカットを適用する場合，ラベルが一次元で順序に定義でき，また平滑化項が凸性を持っているという条件を満たす必要がある．提案手法のような多値多値大域最小化問題にグラフカットを適用する場合，大域的最適解が求まる保証はないが，2 値最適化問題を繰り返すことで近似最適化求める， $\alpha - \beta$ 交換と α 拡張という以下の手法が提案されている． $\alpha - \beta$ 交換と α 拡張ともにある配置から別の配置への移動を定義するものである．

$\alpha - \beta$ 交換：まず，各サイトにラベルを適当に初期値として与え， α または β のいずれかをもつサイトにのみ注目する．そして，各サイトが現在のラベルをもう一方のものに交換するか，しないかという 2 値に対してエネルギー関数を定義する．交換するかしないかの 2 値最適化問題とすることでグラフカットの適用を可能とする．上記の操作を α, β として取り上げるラベル

を変えて繰り返すことで結果を得る．

α 拡張：各サイトにラベルを適当に初期値として与える．任意のラベル α を選び，各サイトが現在のラベルラベルを α に変更するか変更しないかという2値についてエネルギーを定義し，2値最適化問題とすることで，グラフカットの適用を可能とする．上記の操作を α として取り上げるラベルを変えて繰り返すことで，結果を得る．

α 拡張は α にラベルをかえるかどうかの交換を繰り返し，これを同様に $\alpha \in L$ だけくりかえす．このため， α 拡張が $\alpha - \beta$ 交換にくらべて効率的であるため(文献 [69]) 提案手法では α 拡張方式を利用する．

(2) グラフカットによる画素間のマッチング

提案手法の画素間のマッチングに適用するグラフカットの詳細について述べる．文献 [43] では1つの画素(以下，注目画素とする)に対し，全画素との比較を行っていたが提案手法では計算コストの削減と誤対応の低減のために，第 4.2.1 節で推定されたエピソード拘束を利用する．そのため，グラフカットに与えられる探索範囲は全画素ではなく，入力画像の注目画素に対応する過去画像のエピソード線上の画素とする．

過去画像 IM_o の画素集合を P_o ，入力画像 IM_n の画素集合を P_n としたとき，画素 $p(x, y)$ は $p \in P$ に対応する．まず，全ての画素間の移動量をマッピングする f について述べる．ラベリング f は P から L へのマッピングを行う．このとき L はラベルに対応する異なる移動量 (x, y 方向の移動量) の集合であり， $L = L_1, L_2, \dots, L_p, \dots, L_{|P|}$ と定義する．この移動量は前節の SIFT 特徴点の対応関係から得る(詳細は seed の与え方の際に述べる)．

ここで， $p \in P, l \in L$ のとき，ある点を $\langle p, l \rangle$ のペアとする．例えば $\langle l_o = l_n \rangle, \langle l_o = l_n \rangle \in I$ かつ $l_o = l_n$ であったとき，点 $p_o + l_o$ と点 $p_n + l_n$ は同一の移動量を持つ．この同一の移動量をもつ画素間を対応付けることが提案手法のラベリング問題である．入力画像 IM_n を重み付き有向グラフ $G = (V, E)$ と定義する．入力画像 IM_n の各画素 p はノードである．グラフカットで用いるエネルギー

関数は次のように定義する .

$$E(f) = E_{data}(f) + E_{smooth}(f) \quad (15)$$

このエネルギー $E(f)$ を最小にする , ラベル L_p を与えることがラベル付けの問題の解決方法である . 提案手法でのデータ項 $E_{data}(f)$ では , 入力画像の画素 $p(x, y)$ と移動量から与えられた過去画像 IM_o の対応画素候補 $q(x, y)$ を比較し評価する .

$$E_{data}(f) = \sum_{\langle p, f(p) \rangle, \langle q, f(q) \rangle \in I} D(p, q) \quad (16)$$

つまり同じ 3 次元の点を射影している確率が高い , つまり画素間が同じ特徴を持っていれば値が小さくなるような関数として定義する . この画素間が相似しているかの , 評価関数にはマッチングの第 2 段階でも利用した回転・スケール変化・照明変化に頑健な正規化の行われていない SIFT 特徴量 (第 3.3.1 節で行った SIFT 特徴量の外れ値の除去を行う .) の相違度 $SIFTDis$ と , RGB のベクトルの角度の特徴量 A を用い , 下記の式 (17) を用いる .

$$D(p, q) = \min\{0, (M * SIFTDis + (1 - M)A) - K\} \quad (17)$$

$$M \begin{cases} M \rightarrow 0 (SIFTDis \rightarrow \infty) \\ M \rightarrow 1 (SIFTDis \rightarrow 0) \end{cases} \quad (18)$$

式 (18) の M は重み係数であり $0 \leq M \leq 1$ となる . $SIFTDis$ が大きいほど A に注目し , $SIFTDis$ が小さいほど $SIFTDis$ に注目するように重みづけを行う .

SIFT 特徴量の相違度 $SIFTDis$ は , 注目画素 p, q の正規化を行っていない SIFT 特徴量をそれぞれ $Sift(p), Sift(q)$ としたとき下記の式 (19) により得られた特徴量間のユークリッド距離を用いる . 通常の SIFT 特徴量ではヒストグラムの最大値を用いて正規化を行うが , 道路領域のような黒くテクスチャの少ない領域ではヒストグラム全体の値が非常に小さい . そのため画素の SIFT 特徴量ごとに正規化を行うと , SIFT 特徴量の相違度が大きくなる . そこで , 提案手法では画素の SIFT 特徴量 $Sift(p), Sift(q)$ に対して正規化を行わない . ユークリッド距離は整数値の最大を超えないよう $SD = 500$ とする . この際 , ユークリッド距離をエラー値の閾値 $C = 1000$ と比較し最小値をとる .

$$SIFTDis = \min \left\{ C, \frac{\sum_{i=1}^{128} (Sift(p_i) - Sift(q_i))^2}{SD} \right\} \quad (19)$$

(a)4近傍:斜めの境界線と (b)8近傍:斜めの境界線と

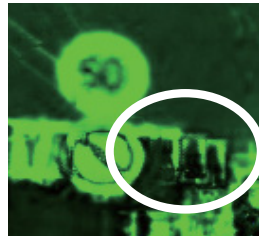
判断されない			判断される		
2	1	2	2	1	2
1	2	1	1	2	1
2	1	2	2	1	2

数値はラベル(移動量の値)
赤色:注目画素 水色:近傍画素



(c)入力画像 (d)過去画像

テクスチャ(看板の文字)
を捉えていない



(e)4近傍の隣接関係を利用した
グラフカットによるマッチングエラーマップ

テクスチャ(看板の文字)を捉えている



(f)8近傍の隣接関係を利用した
グラフカットによるマッチングエラーマップ

図 43 隣接関係数の異なるグラフカットによるマッチング結果

次に RGB のベクトルの角度の特徴量 A は注目画素 p, q の RGB 値を $p_{rgb}(r, g, b)$ と $q_{rgb}(r, g, b)$ としたとき, 下記の式 (20) により得られる. この際, エラー値の閾値 $C = 1000$ と比較し最小値をとる.

$$A = \min \left\{ C, \left(\arccos \frac{\langle p_{rgb}, q_{rgb} \rangle}{\|p_{rgb}\| \cdot \|q_{rgb}\|} \times 100 \right)^2 \right\} \quad (20)$$

定数 K は $K > 0$ であり, 文献 [40] と同様の方法で入力データからデータ項を用いて得られる λ から得られ, $K = 5\lambda$ とする.

また, このデータ項は最大流問題を多項式時間で解くために, 非負であること求められる. 次に平滑化項 $E_{smooth}(f)$ について述べる. 平滑化項は近傍画素との整合性を評価する項である. 提案手法では近傍画素を下記のとおり定義する.

$$N \subset \{\{p, q\} \mid p, q \in P\} \quad (21)$$

近傍画素は注目画素 p, q の上下左右斜めを含む 8 近傍の画素 $f(p), f(q)$ を用いる。提案手法の平滑化項では、ラベルである移動量は同一の物体であれば注目画素と周囲画素間において急激に変化しないという拘束を用いている。今回、街並変化検出の対象となる看板は文字など複雑なテクスチャをもち、斜め方向にもテクスチャが多数存在する。また、上下移動に加えて過去・入力画像の画素間は回転の位置変化も加わっている可能性が高い。そこで、これらのテクスチャを正確にとらえるため、提案手法では 4 近傍ではなく 8 近傍の隣接関係を用いる。4 近傍では図 43 に示したとおり、斜めの境界を正しく判断することができない。実際のグラフカットによるマッチングの結果でも、図 43 に示したとおり、斜めのテクスチャが存在する看板の変化を 4 近傍と比べ 8 近傍では正確に捉えている。同一画像上で $|p_x - q_x| + |p_y - q_y| = 1$ であるとき $p = (p_x, p_y)$ 、 $q = (q_x, q_y)$ は近傍となる。提案手法の平滑化項は下記のとおり定義する。

$$E_{smooth}(f) = \sum_{\{p,q\} \in N} V_{\{p,q\}}(f(p), f(q)) \quad (22)$$

平滑化項は p, q と隣接した画素間は移動量も大きく変化しないと仮定し、ラベルを滑らかにするような関数として定義される。そこで $V_{\{p,q\}}(l_p, l_q)$ は画像の相似した明度をもつ隣接画素間では相似した移動量を持つように設定され、文献 [40] と同様に実験的に定められる。

$$V_{\{p,q\}}(l_p, l_q) = \begin{cases} \lambda(l_p - l_q) & \text{if } \max\{(p_r - q_r), (p_g - q_g), (p_b - q_b)\} < 5 \\ \lambda(l_p - l_q) & \text{otherwise} \end{cases} \quad (23)$$

注目画素と近傍画素の移動量の差分に対し、注目画素と近傍画素の RGB の値から得られた明度の差の最大値をもとに重みづけを行う。また、seed となる移動量の範囲は本来はユーザによって与えられる。提案手法ではエピソード拘束を推定する際に得られた SIFT 特徴点の対応点間の x 軸、 y 軸方向のそれぞれの移動量の平均に対して、 $x \pm 25$ と $y \pm 15$ とすることで、それぞれの最小・最大値を設定する。さらに文献 [61] で提案されている最小最大流定理を用いてグラフの分割を実現した。このグラフカットによるマッチングにより得られた結果をもとに、変化画素かどうかの判断方法について述べる。変化領域抽出は変化マップとマッチングエラーマップの 2 つにより得られる。変化マップ (図 44(c),(f)) は、

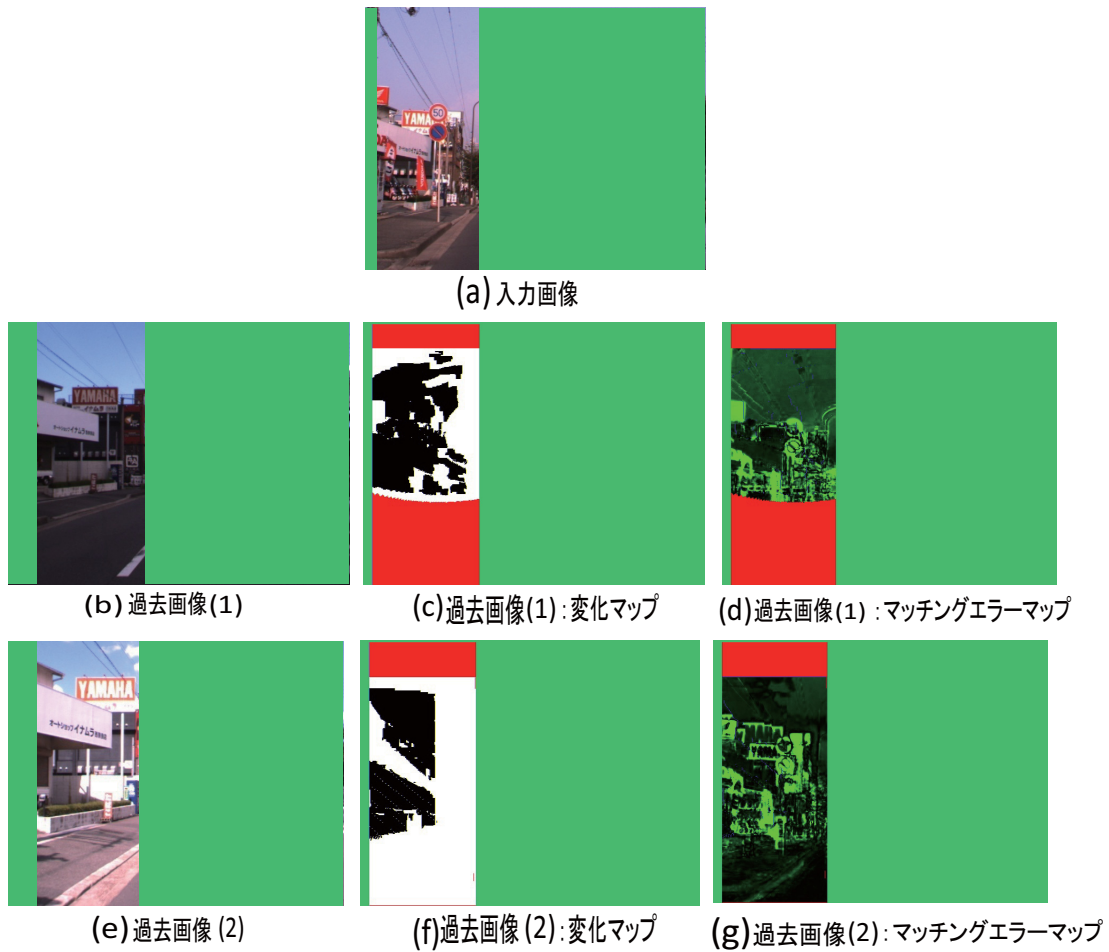


図 44 変化マップ, マッチングエラーマップの出力画像

各画素にラベリングされた移動量の整合性の確認において, 不整合と判断された (移動量のノルムが閾値以上) 画素を「1」, その他を [0] とした 2 値のマップである. 変化マップの作成方法について述べる. 変化が生じることで誤ったラベリングが行われる可能性が生じる. そこで, グラフカットによりラベリングされた画素間の移動量の整合性をみることでラベリングの値を評価する. まず, 過去画像画素 $a(x_a, y_a)$ に対し, ラベリングされた値が $d_a(dx_a, dy_a)$ とすると, 過去画像画素 a に対応する入力画像画素 b は $b(x_a + dx, y_a + dy)$ となる. そこで, 入力画像画素 b にラベリングされた値 $d_b(dx_b, dy_b)$ と過去画像画素 a の $d_a(dx_a, dy_a)$ をのノルムを用いて整合性を評価する. $|dx_a + dx_b|$ と $|dy_a + dy_b|$ の値が閾値以上であっ

た場合を不整合と判断する．閾値は実験的に定めた10とする．これを全画素，さらに過去画像・入力画像を入れ替え交差チェックを行う．

マッチングエラーマップ(図44(d),(g))はマッチングにより移動量がラベリングされた際のデータ項のエラー値，つまり画素間の特徴量の相違度を正規化(0-255)したマップである．よって緑色(0-255)の値が大きいほどマッチングされた画素間の相違が大きい．また，全てのマップの赤色の画素は以下の原因などにより画素間のマッチングが行われなかった画素である．

- (1) エピポーラ拘束の推定失敗により，適切な探索範囲が得られなかった．
- (2) SIFT 特徴点の対応付け失敗および誤差により，適切な移動量の範囲が与えられなかった．

提案手法ではエピポーラ拘束上のみを探索しているため，エピポーラ拘束を推定する SIFT 特徴点の対応付けがえられなかった領域は探索範囲から外れてしまう．また，グラフカットではラベリングされた移動量に対応する画素間を評価するため，与えられた移動量の範囲を超える画素はマッチングが行われぬ．そのため，移動量の範囲が適切でない場合はマッチングが行われぬ画素が生じる．この変化マップとマッチングエラーマップをもとに下記のどちらかの条件を満たした場合，注目画素を変化画素として判定する．

判定条件(1): 変化マップで値1かつマッチングエラーの値が閾値(80)以上．

判定条件(2): マッチングエラーの値が閾値(250)以上．

判定条件(1)が構造の変化を検出するための判定条件である．閾値は街並の構造の変化が生じている画像を用いて実験的に定めた．判定条件(2)はテクスチャの変化を検出するための判定条件であるため，閾値は街並のテクスチャの変化が生じている画像を用いて実験的に定めた．判定結果例を図45(d),(g)に示す．

(3) ノイズ除去

今回，検出したいシーンの変化とは建物や看板の構造やテクスチャの変化である．そこで，遮蔽や細かな誤対応などを変化領域結果から除外する．



図 45 グラフカットによる変化領域抽出結果

まず、マスク境界近くは撮影位置の違いから誤対応が生じやすく、境界から 20pixel 以内 (図 45(d),(g) の黄色の四角で示した) の変化画素は除去する。次に抽出された変化領域結果に対し収縮処理を行う。変化領域抽出結果は二値化された画像である (白が変化画素、黒が変化ではない画素) 収縮処理は周辺に 1 画素でも黒い画素があれば黒に置き換える。これにより、孤立画素や少ない連結を持つ画素を除去する。また、この収縮処理は 1 度だけ適用する。次にこの結果から、連続した変化画素 (以下、変化画素領域とする) の面積が実験的に定めたある一定の面積以下 (50pixel 以下) の場合は変化画素を削除する。第二車線を走行中の車両から撮影した一般的な看板の面積より小さい面積を閾値とした。変化画素領域の面積は、注目画素の 8 近傍の画素において変化画素領域の境界画素を探索し、

処理の高速化のため境界画素数によって領域面積の閾値判定を行う簡易型の処理である。そのため領域内に変化画素ではない画素が多数存在すると、変化画素領域の面積が閾値以下となる場合が生じる。.. ノイズ除去結果を図 45(e),(h) に示す。図 45(f),(i) はノイズ除去結果を入力画像に重畳させた結果である。

4.3 複数画像からの多数決原理に基づく街並変化の検出

抽出された変化領域の中には、移動物体などによる常時変化，駐車車両による短期変化，街並変化を含む固定された物体の変化である長期変化が含まれる．それらを「変化が生じる時間」による識別を行うため，1つの入力画像と3つの撮影日時が異なる過去画像との変化領域抽出結果をもとに，多数決的に変化領域を抽出する．この際，過去画像には街並変化は生じていないが，その他の変化（短

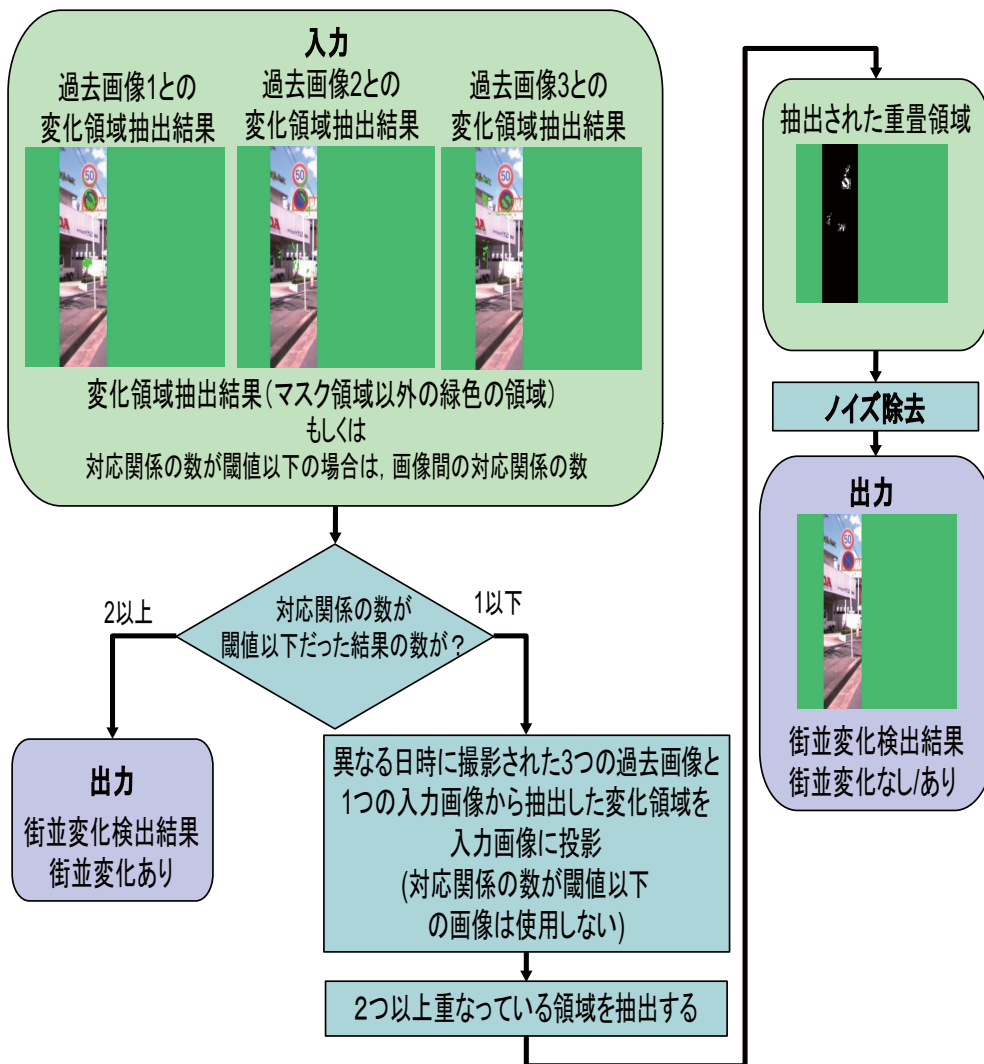


図 46 多数決による街並変化検出処理の流れ

期・常時変化)は生じているものとする。街並変化検出の流れを図46に示す。まず3フレームの過去画像と入力画像の画素間のマッチングにより得られた変化領域抽出結果を入力画像に対してそれぞれ投票する。この際、2つ以上の過去画像との画像間の対応関係推定時に変化ありと判断された場合は、街並に変化ありと検出する。また、変化ありと判断された画像はこの投票には用いない。最後に、2つ以上の領域の重なりが生じた領域を街並の変化として検出する。さらに、街並変化として抽出された領域の面積が50pixel以下である場合もノイズとして除去する。これは変化領域抽出のノイズ除去と同様の処理である。図47に投票結果を示す。投票があった画素を白(1)、その他を黒(0)の2値画像とする。

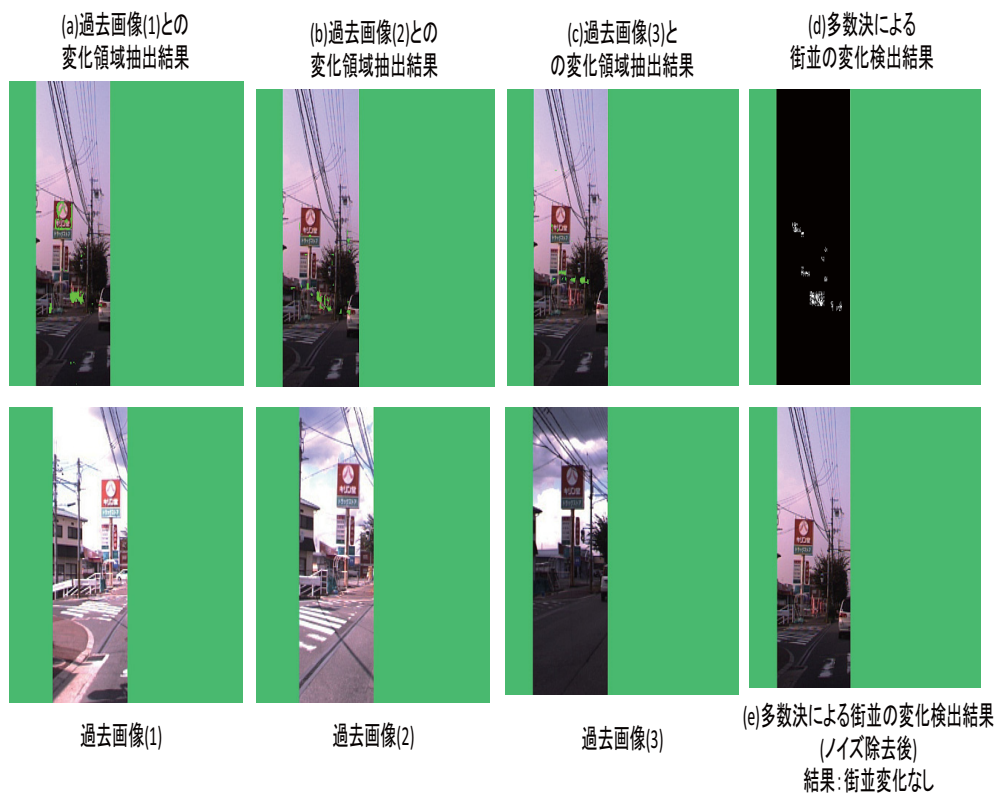


図 47 多数決による街並変化検出結果

4.4 実験と評価

4.4.1 実験概要

第 3.4.1 章の対応付け実験で得られた 20 地点の画像を用いて提案手法の様々な撮影条件の違いへの頑健性と変化領域抽出結果の精度を評価する。実際のシステムの流れと同じにするため、対応付けが失敗している画像も用いる。対応付けの実験と同様に 20 地点、過去画像 3 シーケンス、入力画像 3 シーケンスの画像を用いた。1 シーケンスは 10 フレームとする。撮影画像には第 3.4.1 節で述べたとおり、撮影条件の違いによる見えの変化に加え、様々な常時・短期・長期の変化が生じている。

まず、変化領域抽出結果の正誤判定基準について述べる。街並変化を含む長期変化の生じている撮影画像に対して、下記の条件により変化領域ありと判定するか、または変化領域抽出を行っていれば、正しく変化領域抽出が行われていると判断する。

判定条件 (1): エピポーラ拘束推定時の SIFT 特徴点の対応付けにおいて、マスク領域内の対応関係の数が閾値以下であった。

判定条件 (2): 変化領域抽出結果が実際の長期変化と面積の 50%以上が重畳している

また、長期変化が生じているにも関わらず 2 つのいずれかが行われなければ検出漏れとする。

次に複数画像からの街並変化検出結果に対する正誤判定基準について述べる。定義した街並の変化が生じている画像に対し、下記の 2 つのどちらかの判定を下していれば街並の変化を正しく検出していると判断する。

判定条件 (1): エピポーラ拘束推定時の SIFT 特徴点の対応付けにおいて、過去画像 2 つ以上でマスク領域内の対応関係の数が閾値以下であった。

判定条件 (2): 多数決により得られた街並の変化領域が実際の街並の変化と面積の 50%以上が重畳している

また、長期変化が生じているにも関わらず変化なしと判定した場合は検出漏れとする。そして、それ以外の変化を街並の変化として検出した場合は街並の変化の誤検出とする。

4.4.2 変化領域抽出結果

20 地点計 1800 ペアの画像に対して、変化領域抽出を行った。この中で街並の変化が生じている画像は 327 ペアであり、検出漏れは 28 ペアであった。まずは変化領域抽出結果画像の成功例を示す(図 48)。図 48 は各地点の過去画像、変化マップ、マッチングエラーマップ、入力画像に重畳させた変化領域抽出結果である。結果画像のマスク領域内の緑色の画素が抽出された変化領域である。図 48(b) は建物のテクスチャの変化が生じている画像であるが、正しく変化領域として抽出されている。図 48(a), (d), (e), (f) は街路樹、他車両などの街並の変化以外の短期変化・常時変化が生じている画像であり、それらを変化領域として抽出している。また図 48(c), (e) では自然物体である雲を変化領域として抽出しているが、提案手法では背景と前景を区別するような奥行き推定は行っていないため、これは正しい抽出結果である。次に変化領域抽出に影響を及ぼす照明条件の変化例を図 49 に示す。図 49(a), (c) では照明条件の著しい変化により色情報の変化が生じ、変化が生じていない領域を変化領域として抽出している、提案手法では RGB のベクトル角度を色情報の特徴量として用いているため、著しい照明条件の変化によるこのような色情報の変化には対応できない。さらに図 49(b) では色とびが生じ、それを変化領域として抽出している。

次にエピポーラ拘束推定について述べる。図 50 に示すようにエピポーラ拘束の推定失敗により、正しい画素間のマッチングを行うことができない場合が生じた。図 50 では電線上の SIFT 特徴点の誤対応により、正しい対応関係が得られずエピポーラ拘束の推定に失敗している。この他にも、図 51 に示すように街路樹など似たテクスチャをもつ物体がマスク領域内に多く存在する場合は誤対応が起こりやすくなり、安定したエピポーラ拘束の推定が困難となる。次にエピポーラ拘束を使用せず、全探索を行った結果を図 52 示す。グラフカットの評価関数などは同じである。変化のない領域を変化領域として抽出している。さらに処理時間

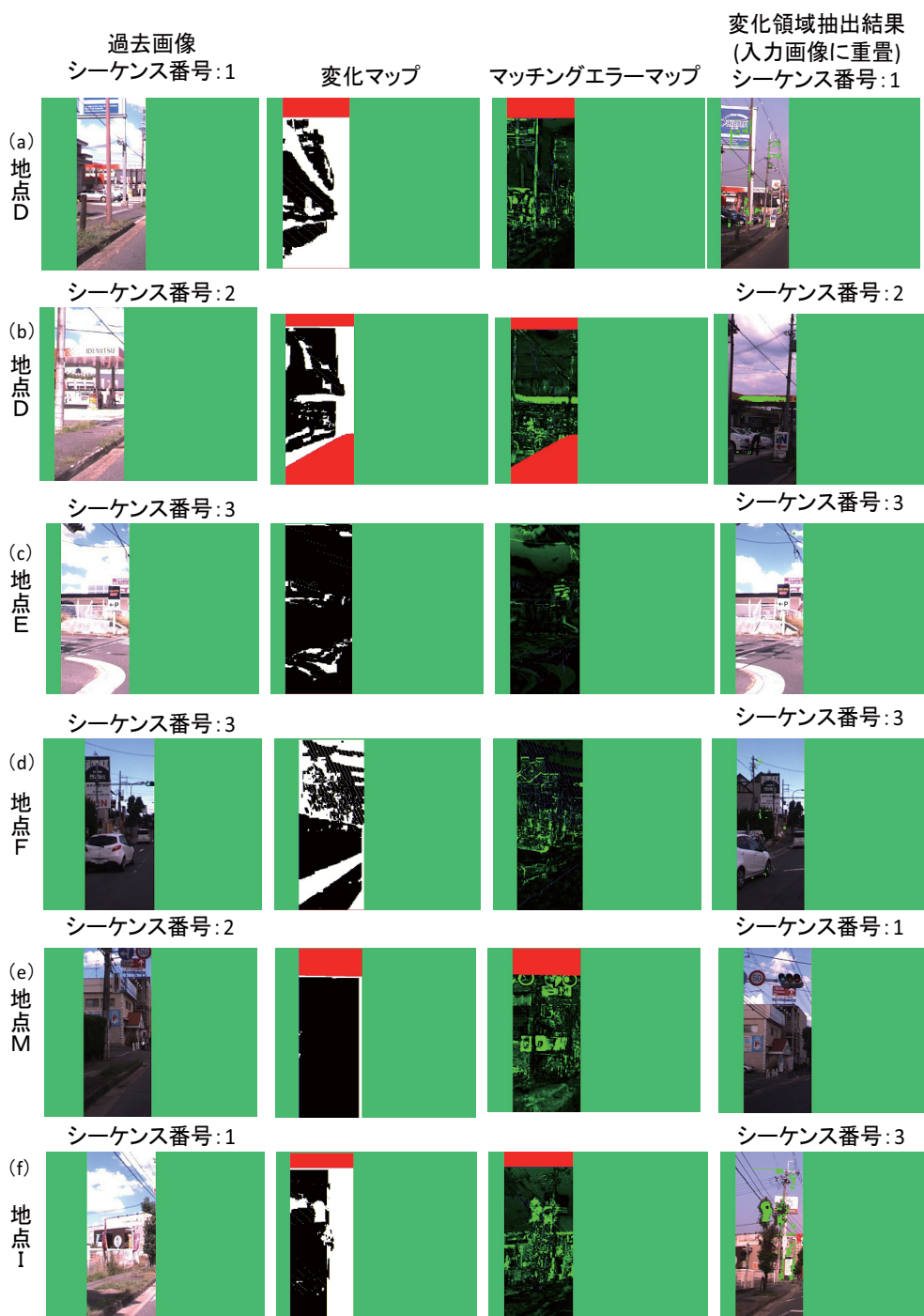


図 48 変化領域抽出結果 (成功)

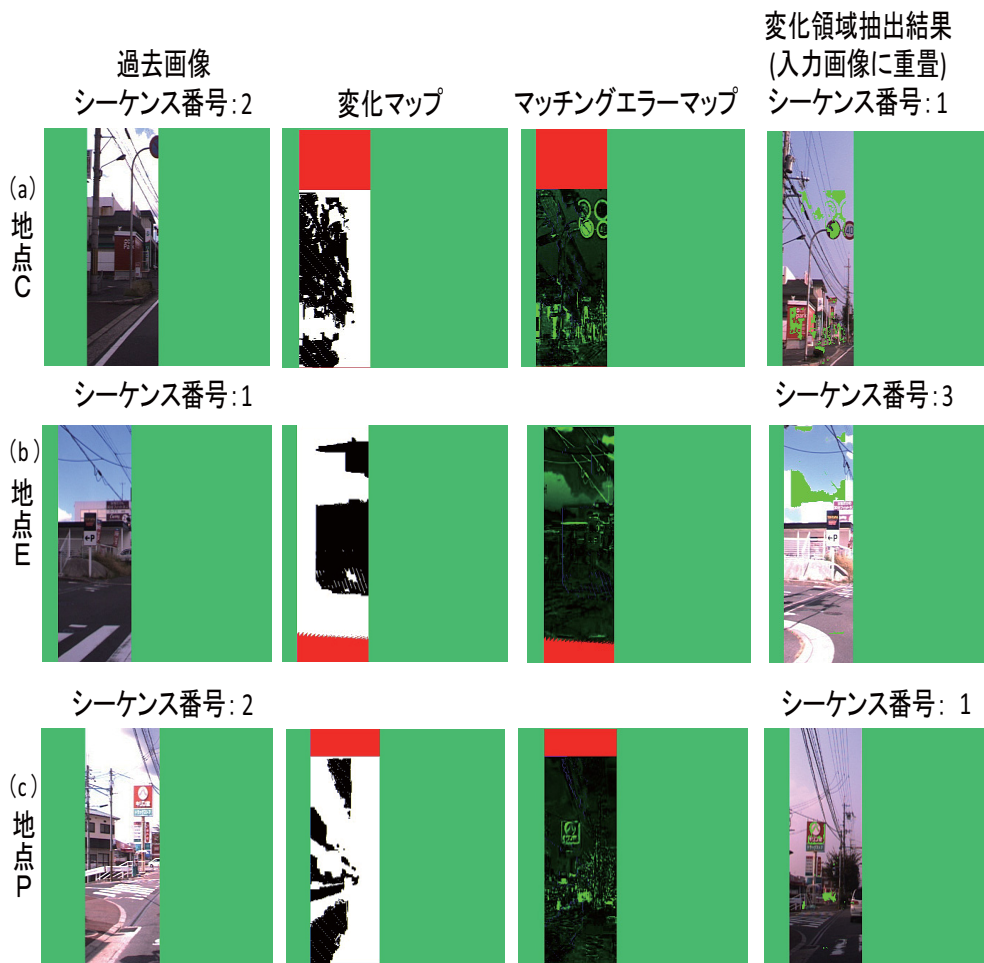


図 49 変化領域抽出に影響を及ぼす照明条件の変化.

も提案手法が 290secs だったのに対し，エピソード拘束を使用しない全探索では 1053secs かった．また，エピソード拘束を使用しない場合は，誤対応が多く発生している．

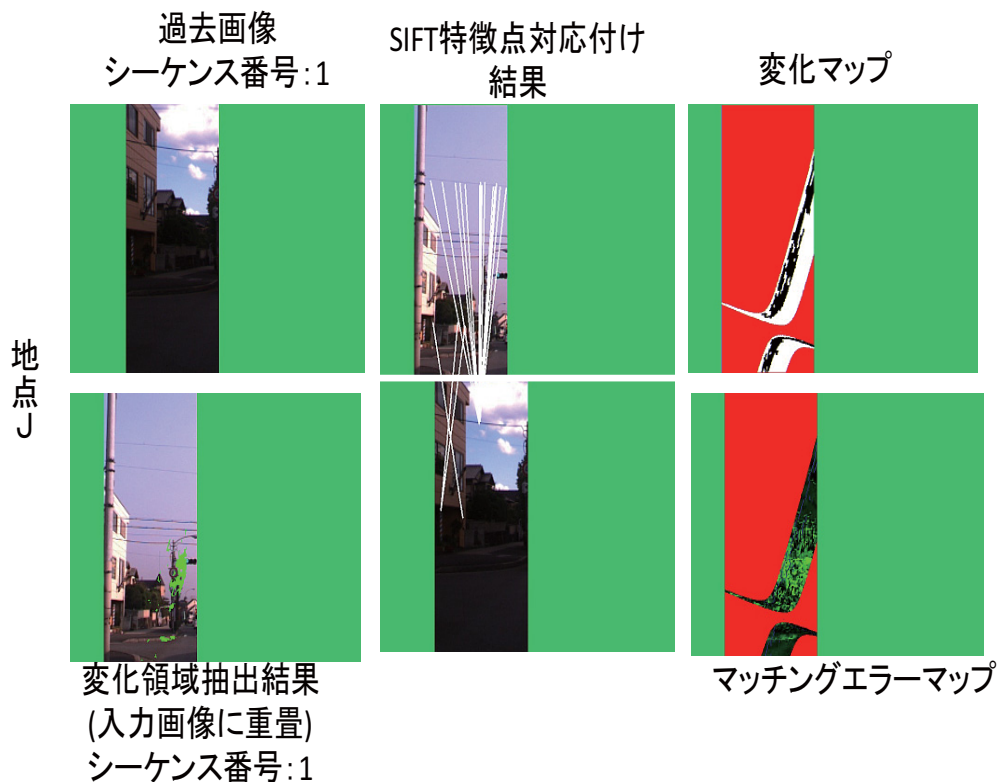


図 50 エピポーラ拘束推定の失敗例.

複数画像からの街並変化検出結果

異なる日時に撮影された複数(3フレーム)の過去画像と入力画像からの変化領域抽出結果をもとに多数決を行い,街並の変化検出を行った.画像は20地点で1地点30組の計600組とし,一組は3フレームの過去画像と1フレームの入力画像で構成される.実験の結果,600組の中に街並変化のあったのは124組で,そのうち正しく変化ありと判断されたものは117組であった.さらに600組中,192組が変化ありと判断された.つまり最終的な街並変化検出漏れ率は6%,街並変化誤検出率が39%となった.これは第2.3節で検討した最低精度目標である街並変化誤検出率を50%,街並変化検出漏れ率を10%以下の範囲内であり,目標精度を達成できたことが確認できた.各地点の結果を表6に示す.街並変化検出漏れは実際の街並変化に対して,街並の変化を検出できなかった組数を表す.また,街並変化誤検出は街並の変化がない画像に対して,街並の変化検出ありと判断した

SIFT特徴点对应付け結果
 入力画像(シーケンス番号:1)

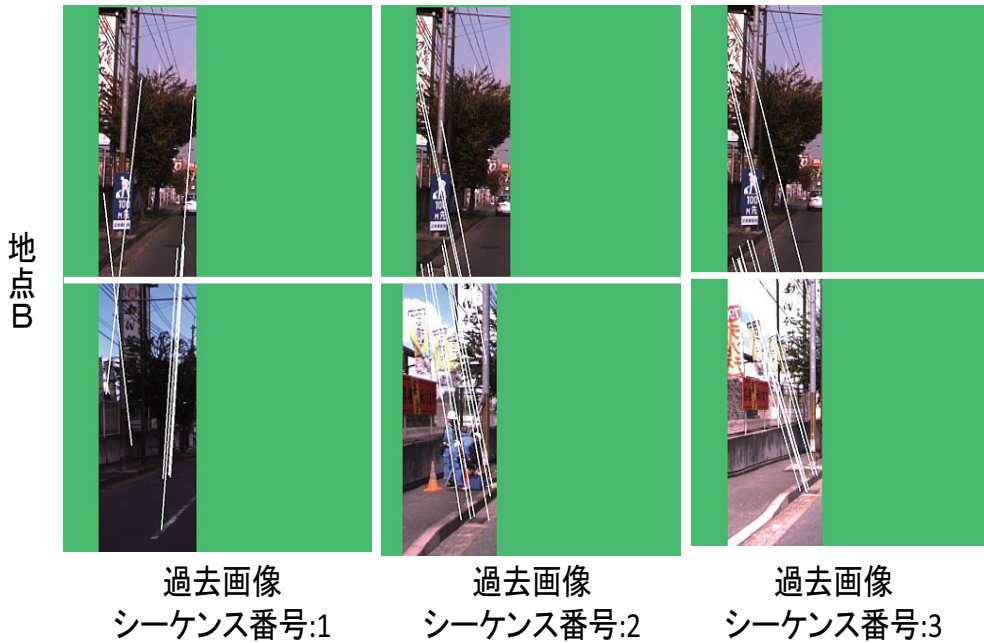


図 51 街路樹による遮蔽が生じている撮影画像からの SIFT 特徴点抽出結果例.

組数を表す．変化検出結果の成功例を図 53，54，55 に示す．

図 54 の上部は，建物の構造およびテクスチャに変化が生じている地点 G の画像である．過去画像 1,2,3 で対応関係の数が閾値以下であったため，街並の変化として検出している．図 54 の下部は，建物の構造およびテクスチャに変化が生じている地点 Q の画像である．過去画像 1,2,3 で対応関係の数が閾値以下であったため，街並の変化として検出している．図 53 の地点 D の画像には，建物のテクスチャに変化が生じている．過去画像 1 と 2 で対応関係の数が閾値以下であったため，街並の変化を検出している．また，過去画像 3 との変化領域抽出結果においても街並の変化を变化領域として正しく抽出している．図 53 の地点 G の画像には，街並の変化は生じていないが他車両による遮蔽と撮影位置の違いによる著しい見えの変化が生じている．判定結果は街並の変化なしと正しい判断を行っている．

図 55 の地点 F,H は共に街並の変化が生じていないの画像である．変化領域抽

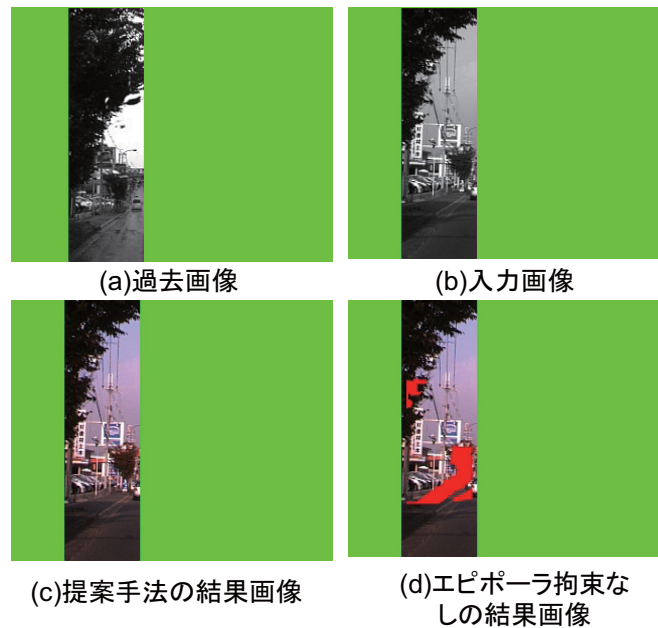


図 52 エピポーラ拘束なしのマッチング結果

出結果の多数決の結果，街並の変化なしと正しい判断を行えている．

次に変化検出結果の失敗例を図 56,57 に示す．図 56 の地点 E は建物 (店舗) の壁面の一部のテクスチャに変化が生じている画像である．街並に変化ありと判断はされているが，抽出された街並の変化は実際の街並変化とは重畳せず，地面のテクスチャの変化 (植物からコンクリートに変化) を街並の変化として検出している．地面のテクスチャは固定された物体ではあるが，今回定義した街並の変化には含まれない．しかしながら，提案手法の「変化が生じる時間」による変化の識別法では，このような固定物の変化は長期変化と判断されるため，正しく識別することはできない．図 57 の地点 N は看板と建物のテクスチャに変化が生じている画像であるが，街並に変化なしと判断し検出漏れを起こしている．テクスチャのみの変化であるため，誤対応を起こし変化領域抽出に失敗していると考えられる．

図 56 の地点 I は，街並に変化がないが同じような位置に形状が異なる雲が生じている画像である．この雲を，街並の変化として検出している．提案手法では抽出した変化が人工物であるかどうかの判断は行っていないため，同じような位置に変化が生じると，街並の変化として識別する．図 57 の地点 T は対応付けの失

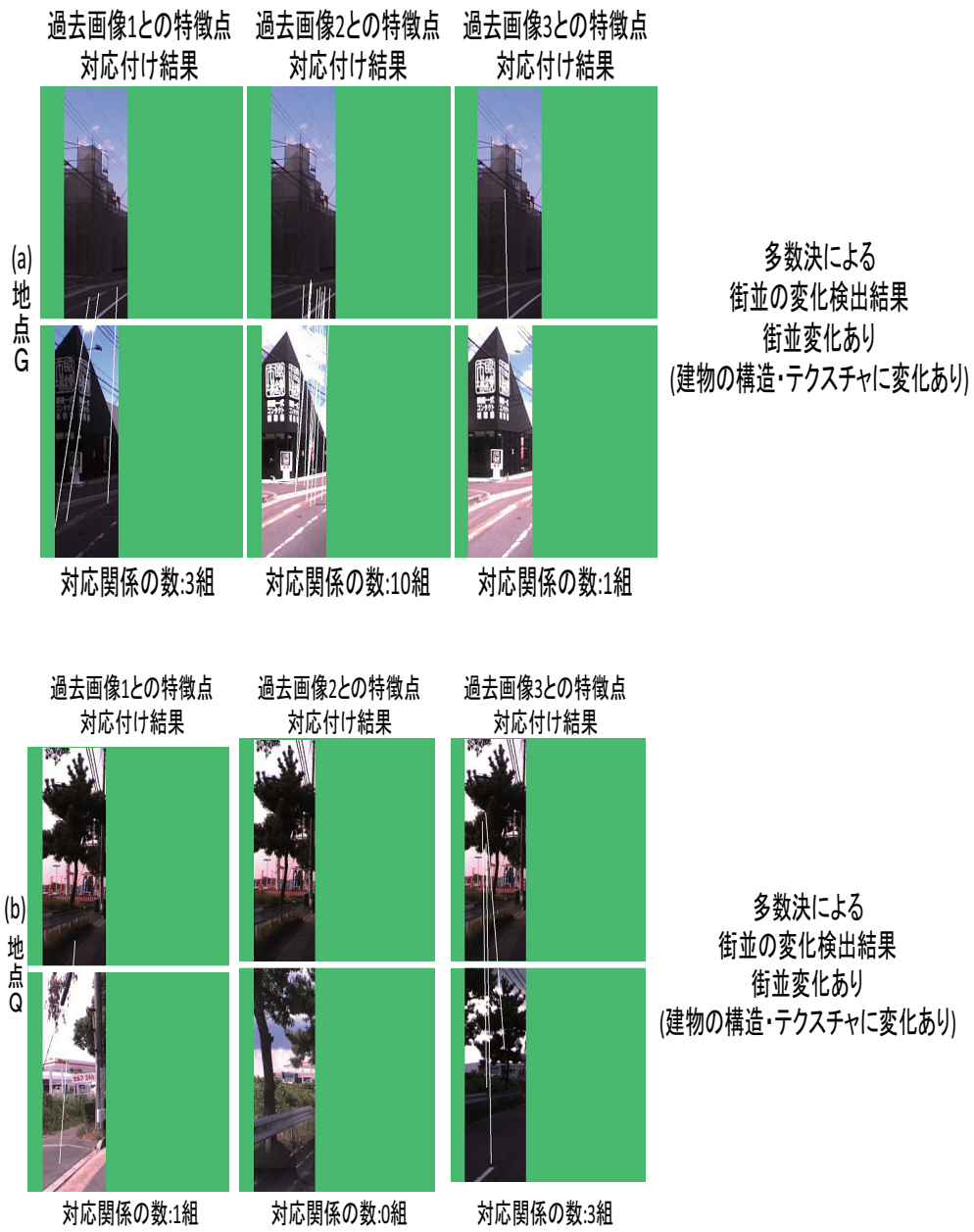


図 53 多数決による変化検出結果 (成功例) : 地点 G , Q

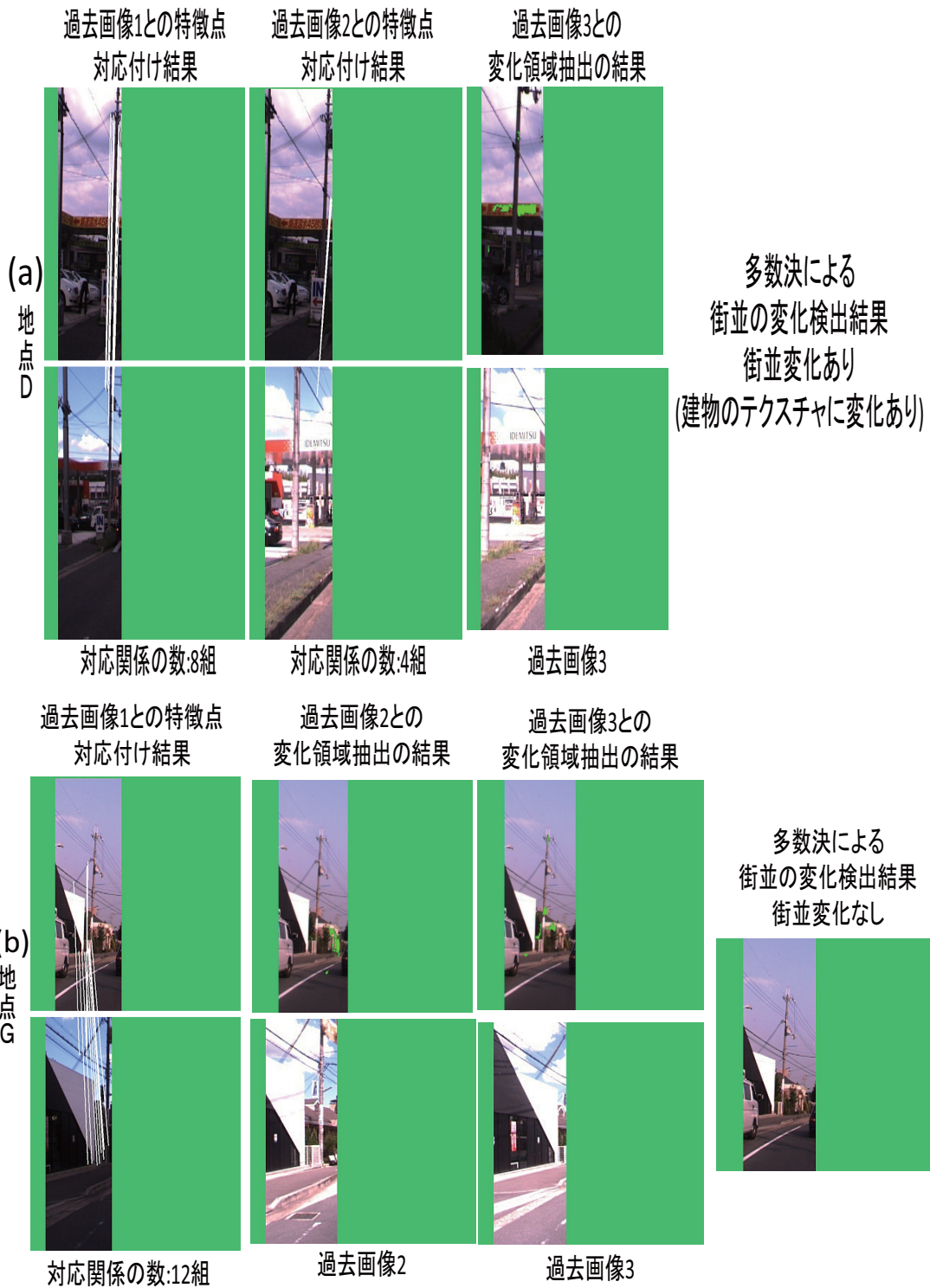


図 54 多数決による変化検出結果 (成功例) : 地点 D,G



図 55 多数決による変化検出結果 (成功例) : 地点 F,H

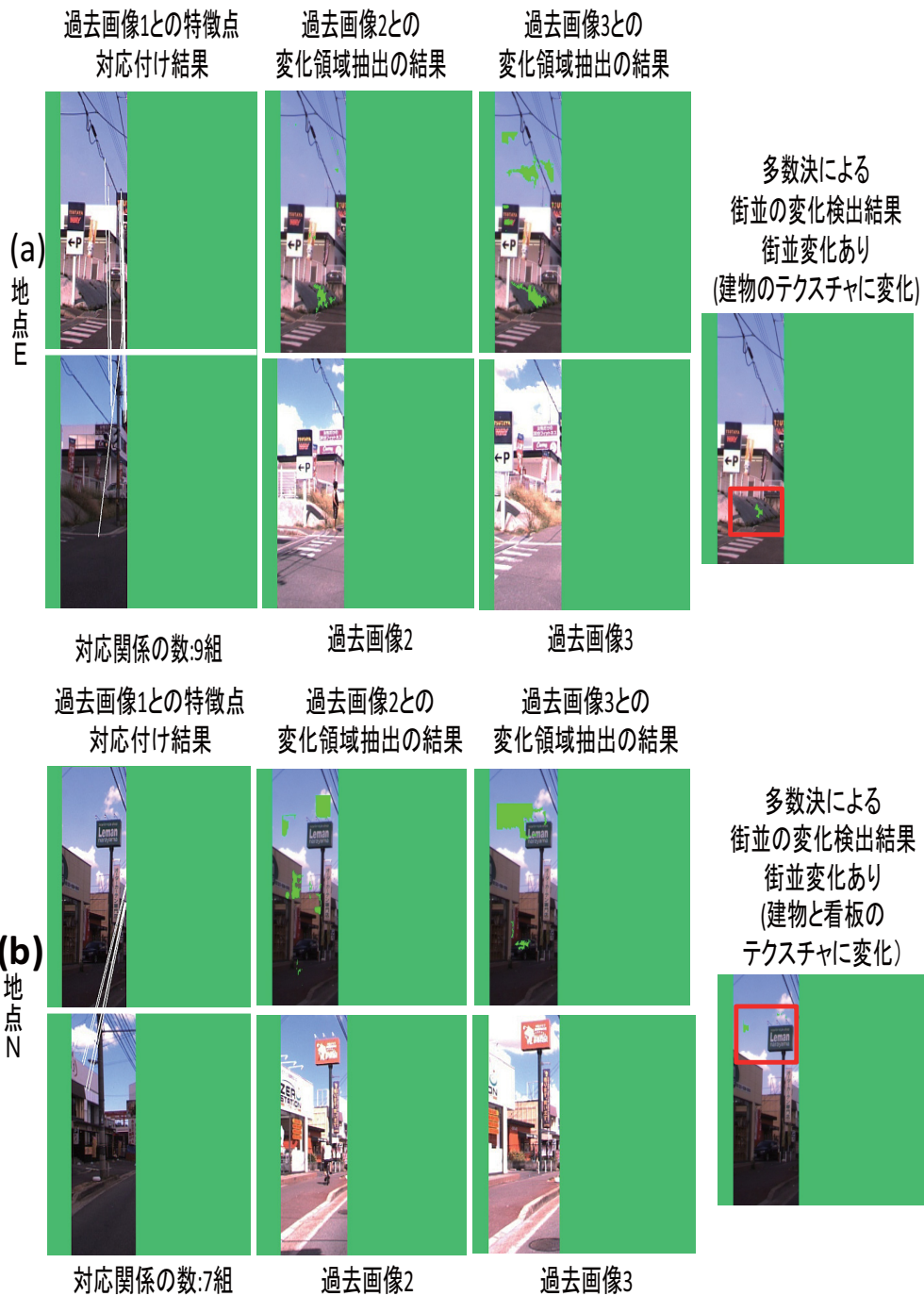


図 56 多数決による変化検出結果 (失敗例) : 地点 E,N

地点	街並変化 検出漏れ	街並変化 誤検出	地点	街並変化 検出漏れ	街並変化 誤検出
地点 A	0/0 組	0/30 組	地点 K	0/0 組	0/30 組
地点 B	0/0 組	0/30 組	地点 L	0/8 組	15/30 組
地点 C	0/0 組	4/30 組	地点 M	0/0 組	0/30 組
地点 D	0/21 組	0/9 組	地点 N	1/24 組	2/6 組
地点 E	2/11 組	0/19 組	地点 O	1/17 組	1/13 組
地点 F	0/0 組	12/30 組	地点 P	0/0 組	16/30 組
地点 G	0/22 組	3/8 組	地点 Q	2/21 組	2/9 組
地点 H	0/0 組	4/30 組	地点 R	0/0 組	0/30 組
地点 I	0/0 組	11/30 組	地点 S	0/0 組	0/30 組
地点 J	0/0 組	0/30 組	地点 T	0/0 組	5/30 組

表 6 複数画像からの街並変化の検出結果

敗と撮影位置とカメラ向きの違いから，過去画像には存在しない（撮影範囲外に存在する）変化のない標識を街並の変化として検出している（図 57 の地点赤い四角の領域）．このように，撮影位置の違いは街並の変化検出の精度に大きな影響を与える．

画像解析の処理時間

1 カ月以内の地図更新が本論文で提案する画像解析要素技術により実現可能か，解析処理時間について考察する．本実験では CPU : Intel Core2 Duo 3.00GHz，メモリ : 2GB の環境で行った．まず，画像間の対応付け処理の時間についてのべる．第 1 段階のパノラマ画像マッチングでは，1 シーケンスの画像間のマッチングの平均処理時間は 21sec であった．第 2 段階の SIFT 特徴点の対応付けは 12sec であり，1 フレームの対応付けには 33sec であった．また，街並の変化検出は 15-874sec であった．よって，1 フレームの画像解析処理時間は計 48-907sec であった．

2.3 節で述べたとおり，精度目標として画像全体の 40%を人手で確認するため

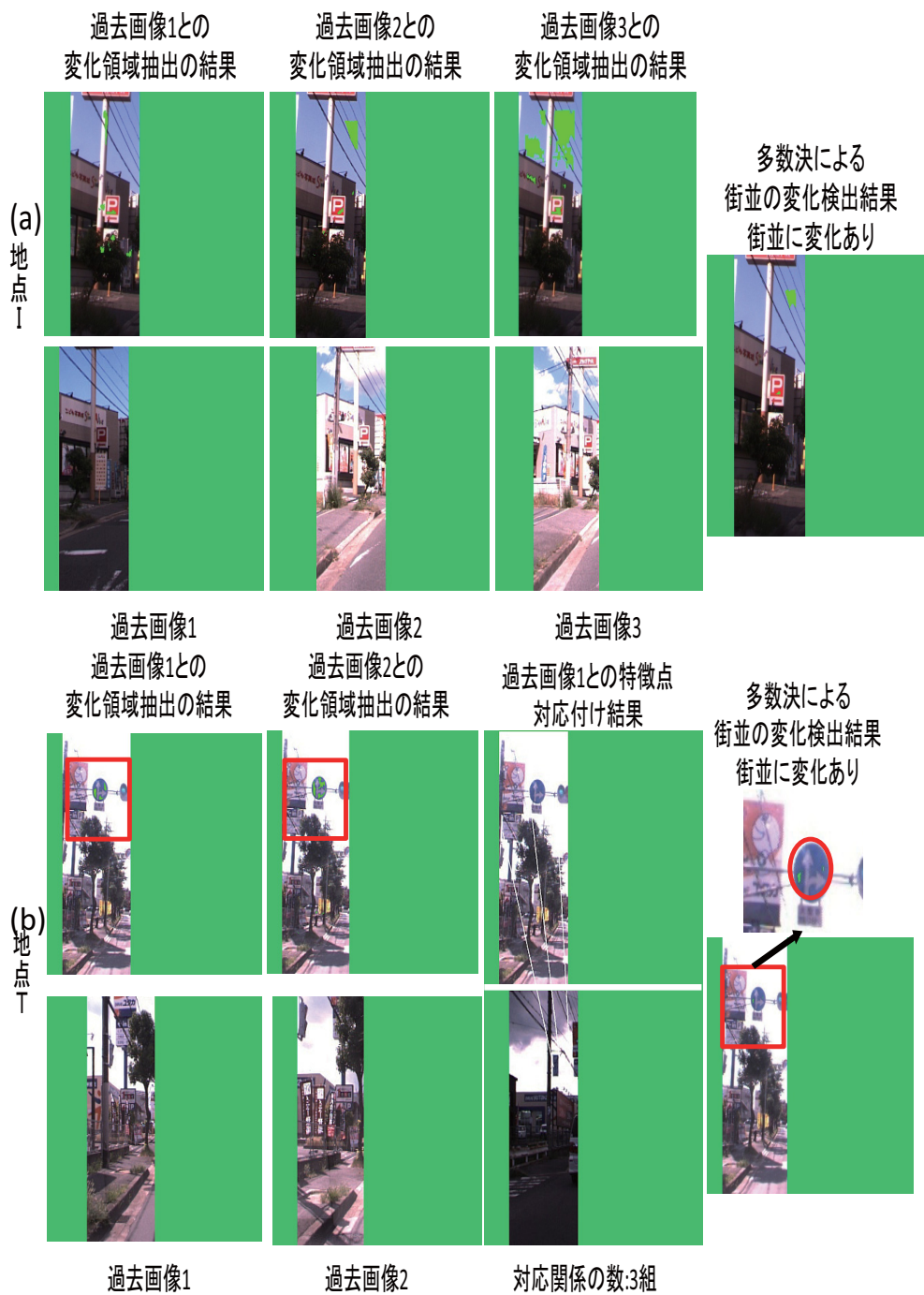


図 57 多数決による変化検出結果 (失敗例) : 地点 I,T

11 日間の処理時間が必要となる．そこで，画像解析処理において1 カ月ごとの更新を実現させるためには全フレーム(約 800,000,000 フレーム)を 19 日以内で処理する必要がある．データベース上の過去画像に必要な処理(画像補正や短冊画像化，SIFT 特徴量の算出など)をあらかじめ行い，プログラム上の効率化と 4 コア以上の CPU による 8 スレッド以上のマルチスレッド処理を仮定すれば，提案手法の処理を 1 フレーム平均 12sec まで抑えることが可能となる．よって，約 330 台(1 都道府県につき 6-7 台)のサーバー(4 コア以上の CPU によるマルチスレッド処理を仮定)により提案手法の画像解析を行えば，18 日間での処理が可能となる．

人手による解析を行っている既存のシステムと比較すれば，サーバーのコスト(1 台 70-100 万円)は安価である．また大量の画像解析が必要であるため，大規模なサーバーセンターの構築は必要不可欠ではある．しかしながら，安定した高頻度更新を実現するにはさらなる処理の高速化が必要となる．提案手法では SIFT 特徴量を用いているが，SIFT 特徴量は撮影条件の変化に頑健ではあるが，計算コストが非常に高い．よってさらなる特徴量計算の高速化が必要である(文献[35])．またグラフカットにおいても，多項式問題を現実的な処理時間で解くことは可能であるが高速とは言い難い．そのため，精度と計算コストを考慮した効率的な最適化手法が求められる．

4.4.3 従来手法との比較

既存手法(文献[43])との比較を行う．既存手法(文献[43])を用い，街並変化が生じている画像 3 ペア，街並変化が生じていない画像 3 ペアに対して変化領域抽出を行った．文献[43]では画素の特徴量に SIFT 特徴量を用い，画素単位のマッチングには信念伝播法を用いている．また提案手法とは異なり全画素とのマッチングを行う．変化画素かの判断条件は，対応づけられた画素間の特徴量の相違度(SIFT 特徴量間のユークリッド距離)が閾値以上であれば変化画素と判断する．閾値は実験的に定め，6.0 とした．さらに，変化領域抽出結果に対して提案手法と同様のノイズ除去処理を行った．街並変化が生じていない画素に対し，変化画素と判断した画素を誤抽出画素とし，街並変化が生じている画素に対し，変化画素と判断した画素を変化抽出画素とした．街並変化が生じているかどうかの判断は，

	既存手法 (誤抽出画素数/変化なし画素数)	提案手法 (誤抽出画素数/変化なし画素数)
(a)	2564/72996	1793/72996
(b)	2821/72996	611/72996
(c)	704/72996	563/72996

表 7 街並変化の生じていない画像に対しての変化領域抽出結果

	既存手法		提案手法	
	(変化抽出画素数 /街並変化画素数)	(誤抽出画素数 /変化なし画素数)	(変化抽出画素数 /街並変化画素数)	(誤抽出画素数 /変化なし画素数)
(a)	180/3681	676/69315	1003/3681	100/69315
(b)	8/2773	2228/70223	10/2773	625/70223
(c)	120/4454	4480/68542	241/4454	1667/68542

表 8 街並変化の生じている画像に対しての変化領域抽出結果

人手により与える．図 58(a)(b)(c) は街並変化が生じていない画像に対して既存手法を適用した結果である．図 58 中の SIFT image は，画素の 128 次元のヒストグラムである SIFT 特徴量を RGB の値によりマッピングしたものである．マッチング結果では，対応付けられた画素間の特徴量の相違度が閾値以上であった画素を 1 とした 2 値画像である．表 7 に提案手法による変化領域抽出結果との比較結果を示す．既存手法と比較し，全ての画像に対して提案手法の誤抽出画素数が少ないことが分かる．図 59(a)(b)(c) は街並変化が生じている画像に対して既存手法を適用した結果である．表 8 に提案手法による変化領域抽出結果との比較結果を示す．既存手法と比較し，提案手法の適用結果の変化抽出画素数がいずれの画像でも多く，また誤抽出画素数が少ないことが分かる．提案手法では特徴量に SIFT 特徴量に加えて RGB 特徴を用いているため，テクスチャのみの変化も捉える事が可能となっている．またエピポーラ拘束を利用し，探索範囲を限定することで似た特徴を持つ異なる物体との誤対応を低減することで，誤抽出画素数が低減していると考えられる．次に処理時間であるが，既存手法では全画素の探索をおこ

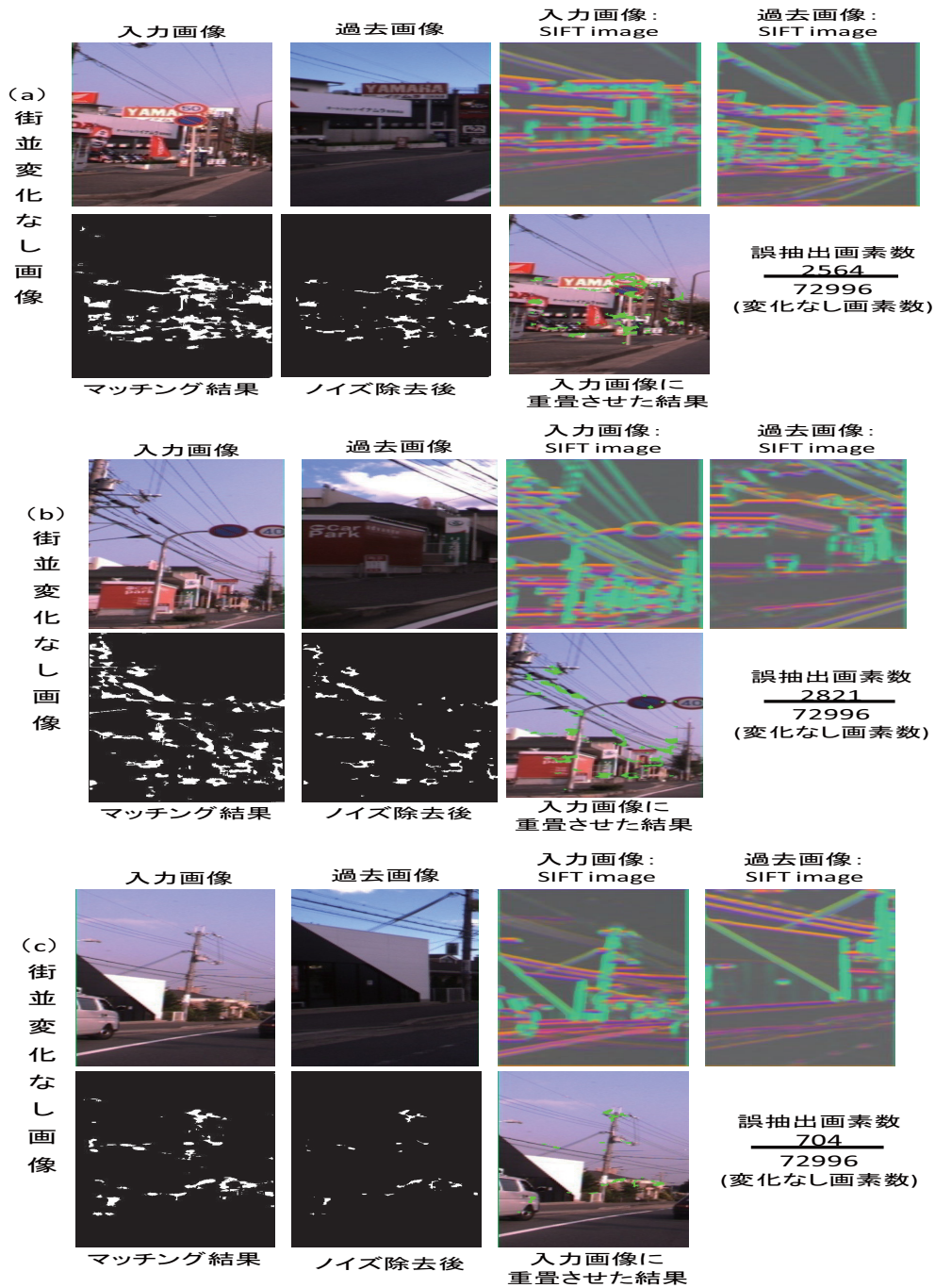


図 58 既存手法による変化領域抽出結果 (街並変化なし)

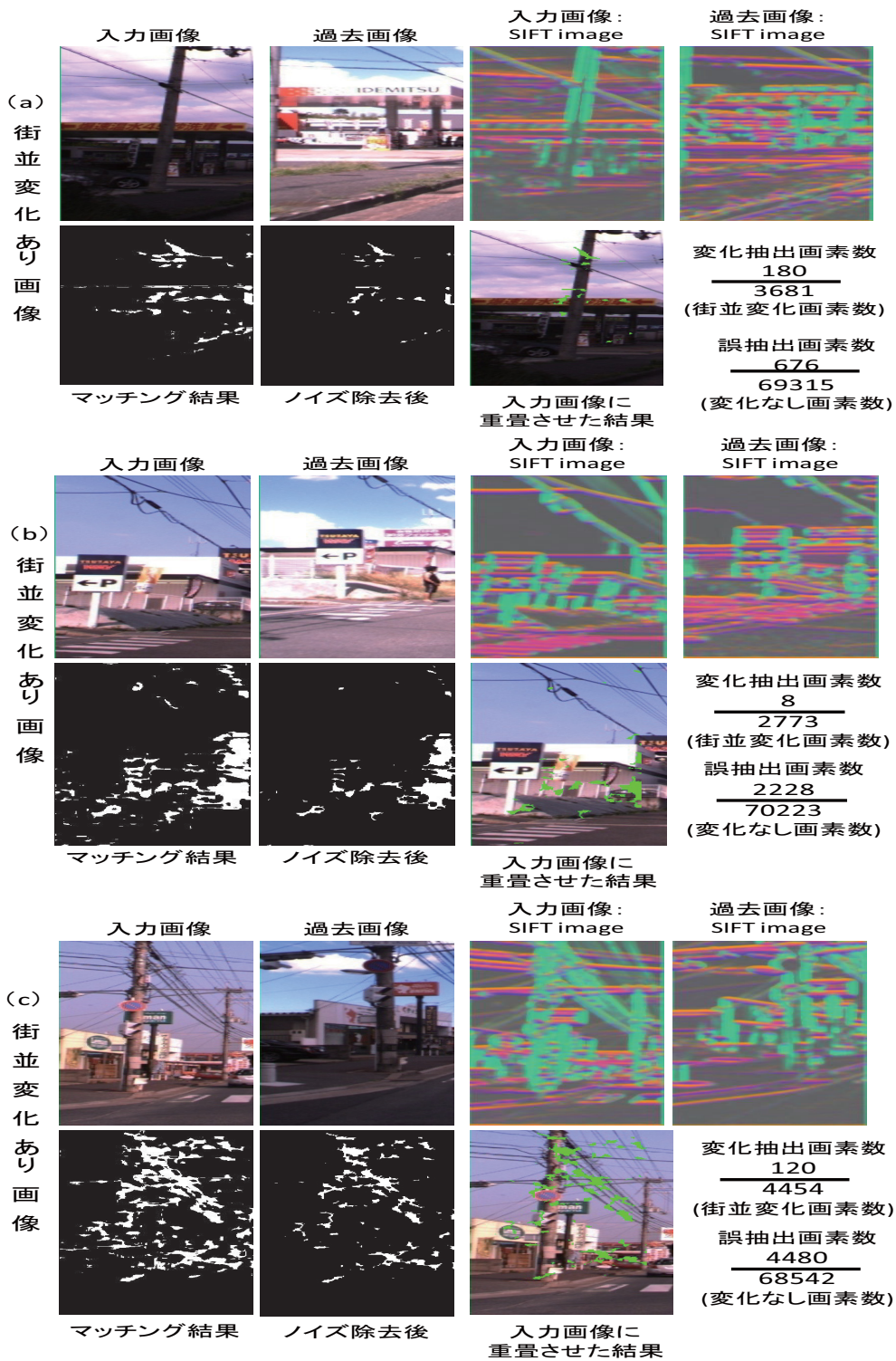


図 59 既存手法による変化領域抽出結果 (街並変化あり)

なっているため画像サイズに比例して計算コストが高くなる．提案手法ではエピソード拘束を利用することで，マッチングの際の探索範囲を限定し計算コスト削減を図っている．既存手法では高性能な計算サーバー（CPU：AMD Opteron Dual Core model 880（クロック 2.4GHz）4CPU，メモリ：2TB）環境下で，約 220sec の処理時間を要した．提案手法が汎用 PC（CPU：Intel Core2 Duo 3.00GHz，メモリ：2GB）環境下で処理時間 290sec であったことを比較すると，提案手法（同環境であれば最低でも 4 割程度の処理時間に短縮可能）による計算コスト削減が確認できた．

4.5 まとめ

本章では、異なる時刻と撮影条件下で得られた車載カメラ画像間から「変化が生じてる時間」をもとに街並の変化を検出する手法を提案した。街並変化検出における街並変化検出漏れ率は6%、街並変化誤検出率が39%となった。これは第2.3節で検討した最低精度目標である街並変化誤検出率を50%、街並変化検出漏れ率を10%以下の範囲内であり、目標精度を達成できたことが確認できた。

まず、撮影条件の変化による見えの変化以外の変化を検出するため、画素間のマッチングを行う。この画素間のマッチングではエピポラ拘束を用いることで、計算コスト削減と誤対応低減を図る。また、エピポラ拘束の推定にはSIFT特徴点の粗いマッチングを利用するが、この特徴点の対応関係が十分に得られないような場合は画素間のマッチングも困難となるため、画像全体に街並の変化が生じていると判断する。

次に、3フレームの過去画像と入力画像間で変化領域抽出を行い、この結果をもとに多数決を行うことで街並の変化を検出を実現した。以下に今後の課題をあげる。

変化領域抽出手法の高速化

大量の入力画像を高速に処理するためにはさらなる高速化が必要となる。まず、変化領域抽出手法では画素の特徴量の1つとしてSIFT特徴量を用いている。SIFT特徴量は撮影条件の変化に頑健な手法であるが、計算コストが非常に高い。そのため、SIFT特徴量計算を高速化させたSURF特徴量などの使用を検討する必要がある。また、画素間のマッチング手法ではグラフカットを用いている。グラフカットは実時間で大域解が得られる手法であるが、大量画像を処理するには十分な計算コストとは言えない。提案手法ではエピポラ拘束を用いることで、探索範囲を限定することで高速化を目指したが、さらなる高速化が必要である。

安定したエピポーラ拘束の推定

画素間マッチングはエピポーラ線上のみを探索するため、エピポーラ拘束を安定して推定する必要がある。このエピポーラ拘束を安定して推定するためには、画像から偏りなく正しい対応点関係を得ることが必要となる。提案手法では特徴点に SIFT 特徴点を用いているが、さらなる精度向上のため特徴点抽出および特徴量の記述の改善が考えられる。

まず、特徴点抽出の改善について述べる。電線上、街路樹など似たテクスチャが多く存在することによる誤対応を防ぐため、直線エッジ上の SIFT 特徴点は除くなど抽出された特徴点の評価を行う方法が考えられる。また、特徴量の記述の改善として PCA-SIFT[70] が提案されている。手法では 39×39 の正方形領域の画素に対して、画像の輝度変化の大きさを x, y 方向で計算し、その結果に対し主成分分析した結果を射影し、36 次元の特徴量を記述している。これらの方法を用いることで、対応点関係推定の精度向上が実現できるが計算コストが増加する問題も生じる。そのため大量画像を処理することを前提とした、精度と計算コストのトレードオフへの考慮が求められる。

変化領域抽出の精度向上

提案手法では注目画素が変化画素かどうかの判断を行う際、各マップごとに最大値を求め正規化を行ったマッチングエラーマップに対して閾値処理を行うため、全体的に相違度が高いといった相違度の分布によっては、マップごとに結果が異なる場合が生じる。しかしながら、全体的に相違度が大きくなるようなマップにおいては先に行う特徴点間マッチングにおいて十分な対応関係が得られない場合が多く、また対応関係が得られない場合は変化領域抽出は行われないため、これらの影響を受けることは少ないが、正規化前の値を用いて閾値処理を行うなどの考慮が必要である。

誤検出率の低下

誤検出の原因として街並以外の変化が生じる問題がある。提案手法では変化領域を「変化が生じている時間」により識別しているため、変化領域がどのような変化であるかは認識していない。そのため、街路樹・雲など自然物体の変化に対しても街並の変化として誤検出する場合が生じる。そこで、抽出された変化領域が何の変化であるかを識別することで誤検出率を低下させることが可能となる。しかしながら、先述したとおり街並の事前知識はなく多種多様である。そこで形状やテクスチャが既知である車両や人を学習データを用いて抽出された変化領域からそれらの短期変化を識別する方法が考えられるが、物体認識の識別率も現状は完全ではなく検出漏れを増加させる可能性もある。提案手法の目的は検出漏れなく、人手による画像解析のコストを削減するかが問題となるため、検出漏れを増加させることなく誤検出率の低下を目指す必要がある。

さらに誤検出の原因として著しい照明条件の変化があげられる。ドライブレコーダーで利用されているカメラは低解像度であり、さらに逆光などにより看板や建物の色によっては完全に白飛びしてしまい、マッチングが困難な場合がある。誤検出率低下のためにも、輝度差が著しく変化している場合はマッチングを行わないなど照明条件に対するなんらかの評価を行う必要がある。

5. 結論

本論文では高頻度のデジタル地図情報更新を実現するため、単眼カメラと普及型 GPS を搭載した多数の一般車両による情報収集を行うデジタル地図情報更新システムの枠組み提案と、システム実現に必要な低品質の大量画像を解析する画像解析手法を開発した。

5.1 本論文の主たる成果

本論文では地図更新効率化のための新しいデジタル地図情報更新システムの枠組みとシステムを実現するために必要な画像解析手法を提案した。このデジタル地図情報は、カーナビゲーションシステム以外にもエリアマーケティングなどの経済分野、また環境シミュレーション・防災情報システムなどの生活・環境分野への転用も可能となる。さらに、システム実現を図るため、システムで得られる大量の低品質の画像を効率的に処理するため、位置情報の誤差に対応するための撮影条件が著しく異なる画像間の対応付けと、画像間の比較による変化検出を行う画像解析手法の提案を行い、高度道路交通システムと画像解析という2つの分野に対して応用的に貢献した。

以下では本論文の主たる成果についてまとめる。

デジタル地図情報更新システムの枠組み

まず2章では、単眼カメラと普及型 GPS を搭載した多数の一般車両による情報収集を行うシステムの枠組み提案とシステム実現に必要な画像解析要素技術の検討を行った。

既存の手法では、情報収集に高価な装置や安定した撮影条件が必要であるため、高頻度の地図更新を実現することは不可能であった。そこで本論文では、安価で既に普及が進んでいる単眼カメラと GPS を搭載した多数の一般車両を観測車両に利用することで、低コストかつ高頻度な街並観測を可能とするデジタル地図情報更新システムの枠組みを提案した。さらに、処理の効率化を図るため大量に得られる画像を自動的に解析する手法の検討を行った。

2 段階マッチングによる車載カメラ画像間の対応付け

3 章では車載カメラ画像間の対応付け手法の提案を行った。街並の変化を検出するためには、新しく得られた入力画像と同一地点をもっとも近い位置から撮影した画像をデータベースから選択する必要がある。しかしながら、撮影時の位置情報には前後 15m 程度の誤差に加え、著しい撮影条件の違いにより、簡単な画像間の対応付けでは適応できない。そこで、2 段階のマッチングによる画像間の対応付け手法を提案した。第 1 段階では撮影条件の違いに対応するための画像補正を行った時系列画像からパノラマ画像を作成し、マッチングを行うことで撮影条件・撮影位置の大きく異なる画像間のマッチングを可能とした。

第 2 段階では、第 1 段階のマッチングにより探索範囲を限定して、SIFT 特徴点による対応付けを行うことで、マッチングの精度向上を行った。実験の結果、提案手法によりさまざまな撮影条件の異なる車載カメラ画像間の対応付けに成功した。

複数画像からの変化領域抽出による街並の変化検出

4 章では、異なる撮影条件下で撮影された車載カメラ画像からの変化領域の検出手法の提案を行った。地図更新のための変化検出では画像間の比較を行うが、撮影条件が異なり、さらに、対象となる街並に関する事前知識もないため、画像を一つの特徴として捉えて変化検出を行うことは困難である。そこで、事前知識なしに対象である街並の変化を検出するため、提案手法ではまず画素間の密なマッチングを行うことで画像から変化領域を抽出する。この変化領域には移動物体などから生じる常時変化と、駐車車両などによる短期変化、街並の変化による長期変化が含まれる。そこで「変化が生じる時間」により、この変化領域を識別する。そのため、3 フレームの過去画像からの変化領域抽出結果による投票により、街並の変化検出を行った。実験の結果、街並変化検出漏れは 6%、街並変化誤検出が 39% となり、提案手法により想定するシステムにおいて求められる誤検出、検出漏れの目標 (最低精度目標である街並変化誤検出率 50%、街並変化検出漏れ 10%) を達成し、画像解析の効率化を実現した。

5.2 今後の課題と展望

本論文が提案した街並情報更新の効率化により、「街並変化に迅速に対応できるカーナビゲーションシステム」から「正確な街並の解析に基づいた道路・都市計画」などの社会応用まで幅広い技術の実現が期待される。以下に、各章における今後の課題について述べる。

デジタル地図情報更新システムの枠組み

システムの実用化の際には、更新処理だけではなくデータ収集においても大量のデータを効率的に解析するための技術が必要となる。

想定システムでは観測車両とサーバの通信は無線通信の利用を想定しているが、無線通信帯域には限界があり、サーバが全ての車両からデータを受信しようとすると、データの質が劣化する。たとえ質のよいデータが受信できても、全国を走行している全ての車両からの全撮影データを適当な時間（地図更新に許される時間）で処理することは不可能である。そこで画像の対応付け、変化検出、地図更新のための変化対象認識に適した画像を送信する車両を選択し、その画像のみを受信・処理するような技術の提案が必要である。

2段階マッチングによる車載カメラ画像間の対応付け

提案手法では特性の異なる段階的なマッチングによる車載カメラ画像間の対応付けを行っているが、この対応付けの精度は次の処理である街並変化検出の精度に大きな影響を与える。そこで、画像間の対応付けの精度向上のため、パノラマ画像の改善と車両の揺れの正確な補正が必要となる。また、照明条件の著しい変化が生じても対応付けを成功させるために、パノラマ画像からエッジ特徴量を抽出する際の閾値は自動的に調節されるべきである。しかしながら、本システムでは異なる車両から同じ場所の多数の画像から得られるため、全ての撮影画像を必ずしも用いる必要はない。そのため、提案手法では撮影された入力画像が過去画像とのマッチングに適しているかどうか、またはマッチングの結果が信頼できるものであるかを分析する必要がある。

複数画像からの変化領域抽出による街並の変化検出

さらなる地図情報更新処理の効率化を目指すためには、街並変化誤認識率の低下が重要な課題となる。そこで提案手法では識別困難な、街路樹や雲といった自然物による変化の除去が必要となる。しかしながら、先述したとおり街並の事前知識はなく多種多様である。そこで形状やテクスチャが既知である車両や人を学習データを用いて抽出された変化領域からそれらの短期変化を識別する方法が考えられるが、物体認識の識別率も現状は完全ではなく検出漏れを増加させる可能性もある。提案手法の目的は検出漏れなく、人手による画像解析のコストを削減するかが問題となるため、検出漏れを増加させることなく誤検出率の低下を目指す必要がある。

謝辞

はじめに，本研究を進めるにあたり，一度企業に勤めながらも至らない私を修士課程から博士課程にいたるまで，暖かく見守りご指導くださいました，知能情報処理学講座教授木戸出正継先生に厚く御礼申し上げます。

また，副指導教官としてご助言を頂きました，像情報処理学講座教授千原國宏先生には，有益な助言とご指導を頂きました。深く感謝致します。

名古屋大学教授村瀬洋先生には，研究当初よりご指導頂くことが私の希望でありました。お忙しい中，有益な助言とご指導を頂きましたことを深く感謝致します。

次に，修士課程よりご指導を頂いた知能情報処理学講座准教授浮田宗伯先生に深く感謝致します。分野に関して無知なだけでなく，至らない私をご指導頂きました。また，浮田先生のお陰でCMUに滞在するという貴重な体験もさせて頂きました。

また，同講座助教波部齊先生にも，助言とご指導を頂きましたことを深く感謝致します。

同講座助教松原崇充先生にはご多忙の中，貴重なお時間を割き有益な助言とご指導を頂きましたことを深く感謝致し，御礼申し上げます。松原先生のおかげで，研究の楽しさを知り，異なる分野への興味を持つことができました。

また，武庫川女子大学教授野村典子先生には学部時代に熱心にご指導頂き，さらに卒業後も有益な助言を頂きましたことを深く感謝致します。今後も，野村先生は私の目標であり憧れです。

さらに本研究を進めるにあたり多大なるご協力を頂きました知能情報処理学講座の諸氏に感謝の意を表します。諸先輩方，同輩，後輩の皆さまの励ましがなければ，現在の私はありませんでした。

最後に，再びの学生生活を見守り，支えてくれた家族・友人に感謝します。特に両親には最後に多大な心労をかけてしまいました。これからは，今までの恩を少しでも返していけるよう今以上に精進いたします。

参考文献

- [1] Yukio Ogawa, Kazuaki Iwamura, and Kakumoto Shigeru. Extracting object information from aerial images : A map-based approach. *IEICE transactions on information and systems*, Vol. 83, No. 7, pp. 1450–1457, 2000.
- [2] Katsuyuki Kamei, Wayne Hoy, Takashi Tamada, and Kazuo Seo. Modeling of urban scenes by aerial photographs and simply reconstructed buildings. *IEICE transactions on information and systems*, Vol. 83, No. 7, pp. 1441–1449, 2000.
- [3] Weiming Li, Xiaoming Li, Yihong Wu, and Zhanyi Hu. A novel framework for urban change detection using vhr satellite images. In *International Conference on Pattern Recognition*, Vol. 2, pp. 312–315, 2006.
- [4] Z C Shi and Ryouzuke Shibasaki. High accurate positioning and mapping in urban area using laser range scanner. In *IEEE Intelligent Vehicle Symposium*, pp. 125–132, 2001.
- [5] Ioannis Stamos and Peter K. Allen. 3-d model construction using range and image data. In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 531–536, 2000.
- [6] Z C Shi and Ryouzuke Shibasaki. Automated building extraction from digital stereo imagery. In *Automatic Extraction of Man-Made Objects from AerialAnd Space Images*, pp. 119–128, 1997.
- [7] Nico Cornelis, Kurt Cornelis, and Luc Van Gool. Fast compact city modeling for navigation pre-visualization. In *IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp. 1339–1344, 2006.
- [8] Junji Sato, Tomokazu Takahashi, Ichiro Ide, and Hiroshi Murase. Change detection in streetscapes from gps coordinated omni-directional image se-

- quences. In *International Conference on Pattern Recognition*, Vol. 4, pp. 935–938, 2006.
- [9] 加藤浩仁, 石黒浩, Matthew Barth. Town digitizing : 全方位カメラとgpsによる実時間街モデリング. 情報処理学会研究報告. CVIM, No. 4, pp. 111–118, 2001.
- [10] Google. Google streetview. <http://maps.google.com/help/maps/streetview/>.
- [11] 佐藤准嗣, 高橋友和, 井手一郎, 村瀬洋. 市街地映像マップの構築のための車載カメラ映像間対応付け. 画像の認識・理解シンポジウム 講演予稿集, pp. 341–346, 2005.
- [12] Agarwala Aseem, Agrawala Maneesh, Cohen Michael, Salesin David, and Szeliski Richard. Photographing long scenes with multi-viewpoint panoramas. *ACM Transactions on Graphics*, Vol. 25, pp. 853–861, July 2006.
- [13] J. Yu Zheng and S. Tsuji. Panoramic representation for route reconstruction by a mobile robot. *International Journal of Computer Vision*, Vol. 9, pp. 55–76, 1992.
- [14] Jiang Yu Zheng. Digital route panoramas. *IEEE Multimedia*, Vol. 10, No. 3, pp. 57–67, 2003.
- [15] Huijing Zhao and Ryosuke Shibasaki. Reconstructing textured cad model of urban environment using vehicle-borne laser range scanners and line cameras. In *International Workshop on Computer Vision Systems*, Vol. 2095, pp. 284–295, 2001.
- [16] M. Shi and J.Y. Zheng. Spatial resolution analysis of route panorama. In *International Conference on Image Processing*, pp. II: 311–314, 2003.
- [17] BOLLES R. C. Epipolar-plane image analysis : An approach to determining structure from motion. *International Journal of Computer Vision*, Vol. 1, No. 1, pp. 7–55, 1987.

- [18] Takeshi Mikami, Thanda Oo, Shintaro Ono, Hiroshi Kawasaki, Yutaka Oh-sawa, and Katsushi Ikeuchi. Distortion-free fusion of multiple video camera images using epi analysis. *IEICE transactions on information and systems*, Vol. 89, No. 6, pp. 1336–1347, 2006-06-01.
- [19] Notomi Mikito, Ozawa Shiro, and ZEN Heitou. Modeling of urban scene by motion analysis. *IEICE transactions on Electronics*, Vol. 81, No. 5, pp. 872–879, 1998.
- [20] Katsushi Ikeuchi, Masao Sakauchi, Hiroshi Kawasaki, and Imari Sato. Constructing virtual cities by using panoramic images. *International Journal of Computer Vision*, Vol. 58, pp. 237–247, August 2004.
- [21] C. Harris and M. Stephens. A combined corner and edge detector. In *The Fourth Alvey Vision Conference*, pp. 147–151, 1988.
- [22] David G. Lowe. Object recognition from local scale-invariant features. In *International Conference on Computer Vision*, 1999.
- [23] R. Bellman and R. Kalaba. *Dynamic Programming and Modern Control Theory*. Academic Press, 1965.
- [24] H. Sakoe and S. Chiba. Dynamic programming optimization for spoken word recognition. In *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. ASSP-26, pp. 43–49, 1978.
- [25] Bruce D. Lucas and Takeo Kanade. An iterative image registration technique with an application to stereo vision. In *International Joint Conference on Artificial Intelligence*, pp. 674–679, 1981.
- [26] Brown Matthew and Lowe David G. Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision*, Vol. 74, No. 1, pp. 59–73, 2007.

- [27] 野口裕之, 向川康博, 越後富夫, 八木康史. ルートパノラマ作成のための線分抽出. 画像の認識・理解シンポジウム 講演予稿集, pp. 1554–1559, 2005.
- [28] Lowe David G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, Vol. 60, pp. 91–110, November 2004.
- [29] Martin A. Fischler and Robert C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, Vol. 24, pp. 381–395, 1981.
- [30] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, pp. 1330–1334, 2000.
- [31] Akira Hayashi, Yuko Mizuhara, and Nobuo Suematsu. Embedding time series data for classification. In *Machine Learning and Data Mining in Pattern Recognition*, Vol. 3857, pp. 356–365, 2005.
- [32] Shum Heung-Yeung and He Li-Wei. Rendering with concentric mosaics. In *Annual Conference on Computer graphics and interactive techniques*, pp. 299–306, 1999.
- [33] Benno Heigl, Reinhard Koch, Marc Pollefeys, Joachim Denzler, and Luc Van Gool. Plenoptic modeling and rendering from image sequences taken by a hand-held camera. In *Symposium fur Mustererkennung*, pp. 94–101, 1999.
- [34] Steedly Drew, Pal Chris, and Szeliski Richard. Efficiently registering video into panoramic mosaics. In *IEEE International Conference on Computer Vision*, Vol. 2, pp. 1300–1307, 2005.
- [35] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *European Conference on Computer Vision*, pp. 404–417, 2006.

- [36] Sudipta N. Sinha, Jan michael Frahm, Marc Pollefeys, and Yakup Genc. Gpu-based video feature tracking and matching. In *Workshop on Edge Computing Using New Commodity Architectures*, 2006.
- [37] Vincent Lepetit and Pascal Fua. Keypoint recognition using randomized trees. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, p. 2006, 2006.
- [38] Breiman Leo. Random forests. *Machine Learning*, Vol. 45, pp. 5–32, 2001.
- [39] Simon Baker, Stefan Roth, Daniel Scharstein, Michael J. Black, J.P. Lewis, and Richard Szeliski. A database and evaluation methodology for optical flow. In *International Conference on Computer Vision*, pp. 1–8, 2007.
- [40] Sudipta N. Sinha, Philippos Mordohai, and Marc Pollefeys. Multi-view stereo via graph cuts on the dual of an adaptive tetrahedral mesh. In *International Conference on Computer Vision*, pp. 1–8, 2007.
- [41] Noah Snavely, Steven M. Seitz, and Richard Szeliski. Photo tourism: Exploring photo collections in 3d. In *ACM Transactions on Graphics*, pp. 835–846, 2006.
- [42] Vladimir Kolmogorov and Ramin Zabih. Multi-camera scene reconstruction via graph cuts. In *European Conference on Computer Vision*, pp. 82–96, 2002.
- [43] Yuen Jenny Liu Ce, Torralba Antonio, Sivic Josef, and Freeman William T. Sift flow: Dense correspondence across different scenes. In *European Conference on Computer Vision*, pp. 28–42, 2008.
- [44] Masafumi Noda, Tomokazu Takahashi, Daisuke Deguchi, Ichiro Ide, Hiroshi Murase, Yoshiko Kojima, and Takashi Naito. Recognition of road markings from in-vehicle camera images by a generative learning method. In *IAPR Conference on Machine Vision Applications*, pp. 514–517, 2009.

- [45] Masanari Takaki and Hironobu Fujiyoshi. Traffic sign recognition using sift features. *IEEJ Transactions on Electronics, Information and Systems*, Vol. 129, No. 5, pp. 824–831, 2009-05-01.
- [46] Yaser Sheikh, Omar Javed, and Takeo Kanade. Background subtraction for freely moving cameras. In *International Conference on Computer Vision*, pp. 1219–1225, 2009.
- [47] Kobi Levi and Yair Weiss. Learning object detection from a small number of examples: the importance of good features. In *IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp. 53–60, 2004.
- [48] C. Stauffer and W. Grimson. Adaptive background mixture models for real-time tracking. In *IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 2, 1999.
- [49] 議田賢治, 坪内貴之, 菅谷保之, 金谷健一. 移動ビデオカメラ画像からの運動物体の抽出. 情報処理学会研究報告. CVIM, 第 2004 巻, pp. 41–48, 2004.
- [50] Yuji Yamauchi and Hironobu Fujiyoshi. People detection based on co-occurrence of appearance and spatiotemporal features. In *International Conference on Pattern Recognition*, 2008.
- [51] D. M. Gavrila and S. Munder. Multi-cue pedestrian detection and tracking from a moving vehicle. *International Journal of Computer Vision*, Vol. 73, No. 1, pp. 41–59, 2007.
- [52] Leibe Bastian, Schindler Konrad, Cornelis Nico, and Van Gool Luc. Coupled object detection and tracking from static cameras and moving vehicles. *IEEE transactions on pattern analysis and machine intelligence*, Vol. 30, pp. 1683–1698, 2008.
- [53] Pedro F. Felzenszwalb and Daniel P. Huttenlocher. Efficient belief propagation for early vision. *International Journal of Computer Vision*, Vol. 70, pp.

41–54, October 2006.

- [54] Yuri Boykov, Olga Veksler, and Ramin Zabih. Markov random fields with efficient approximations. In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 648–655, 1998.
- [55] Vladimir Kolmogorov and Ramin Zabih. What energy functions can be minimized via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, pp. 65–81, 2004.
- [56] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2004.
- [57] Kenichi Kanatani. *Geometric Computation for Machine Vision*. Oxford University Press, 1993.
- [58] Kenichi Kanatani. *Statistical Optimization for Geometric Computation: Theory and Practice*. Dover Publications, 1996.
- [59] 奥富正敏 (編集) . デジタル画像処理. CG – ARTS 協会, 2004.
- [60] Masao Shimizu and Masatoshi Okutomi. Sub-pixel estimation error cancellation on area-based matching. *International Journal of Computer Vision*, Vol. 63, pp. 207–224, 2005.
- [61] Yuri Boykov and Vladimir Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, No. 9, 2004.
- [62] Hiroshi Ishikawa and Davi Geiger. Occlusions, discontinuities, and epipolar lines in stereo. In *European Conference on Computer Vision*, pp. 232–248, 1998.

- [63] Kim Junhwan, Kolmogorov Vladimir, and Zabih Ramin. Visual correspondence using energy minimization and mutual information. In *International Conference on Computer Vision*, pp. 1033–, 2003.
- [64] Yuri Boykov, Olga Veksler, and Ramin Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, pp. 1222–1239, 2001.
- [65] Andrew Blake, Carsten Rother, M. Brown Patrick Perez, and Philip H. S. Torr. Interactive image segmentation using an adaptive gmmrf model. In *European Conference on Computer Vision*, pp. 428–441, 2004.
- [66] Yuri Boykov and Gareth Funka-lea. Graph cuts and efficient n-d image segmentation. *International Journal of Computer Vision*, Vol. 70, pp. 109–131, 2006.
- [67] L. Ford and D. Fulkerson. *Flow in Networks*. Princeton University Press, 1962.
- [68] Andrew V. Goldberg and Robert E. Tarjan. A new approach to the maximum flow problem. In *ACM Symposium on Theory of Computing*, pp. 136–146, 1986.
- [69] Yuri Boykov, Olga Veksler, and Ramin Zabih. Fast approximate energy minimization via graph cutsh. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 7, pp. 1222–1239, 2001.
- [70] Yan Ke and Rahul Sukthankar. Pca-sift: A more distinctive. In *IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp. 506–513, 2004.

研究業績一覧

査読付学術論文

- 1 Naoko Enami ,Norimichi Ukita ,Masatsugu Kidode ,Image Matching with a Car-mounted Camera robust to Changes in Imaging Conditions ,International Journal of Pattern Recognition and Artificial Intelligence, Vol.23, No.7, pp.1369-1395, Nov,2009.(論文第2,3章に対応)

査読付国際会議

- 1 Naoko Enami ,Norimichi Ukita ,Masatsugu Kidode ,Image Matching Robust to Changes in Imaging Conditions with a Car-mounted Camera , The 2nd ACM/IEEE International Conference on Distributed Smart Cameras, Sep, 2008. (論文第3章に対応)
- 2 Naoko Enami ,Norimichi Ukita , Masatsugu Kidode , Image Rectification for Robust Matching of Car-mounted Camera Image , The 11th IAPR Conference on Machine Vision Applications, May, 2009.(論文第3章に対応)

査読なし国内会議

- 1 榎並直子,浮田宗伯,木戸出正継,撮影条件の変化に頑健な車載カメラ画像間の対応付け,画像センシングシンポジウム(SSII),6月,2008年.(論文第3章に対応)
- 2 榎並直子,浮田宗伯,木戸出正継,領域マッチングによる車載カメラ画像からの変化領域の抽出,パターン認識・メディア理解研究会(PRMU),10月,2010年.(論文第4章に対応)