

修士論文

輝度情報と奥行情報の整合性を利用した  
エネルギー最小化に基づく  
移動計測された三次元点群からの移動物体領域の検出

高部 篤志

2016年3月14日

奈良先端科学技術大学院大学  
情報科学研究科

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  
修士(工学) 授与の要件として提出した修士論文である。

高部 篤志

審査委員：

横矢 直和 教授	(主指導教員)
小笠原 司 教授	(副指導教員)
佐藤 智和 准教授	(副指導教員)
河合 紀彦 助教	(副指導教員)

# 輝度情報と奥行情報の整合性を利用した エネルギー最小化に基づく 移動計測された三次元点群からの移動物体領域の検出\*

高部 篤志

## 内容梗概

近年、屋外実環境の三次元モデルを用いたアプリケーションとして、バーチャルアースシステムや、ドライビングシミュレータなどが利用されている。これらのアプリケーションに用いられる三次元モデルは、レーザレンジファインダ、カメラ、GPS・IMU等を組み合わせた車載システムを用いた屋外実環境の計測に基づき生成されるが、計測結果に移動物体が含まれる場合、生成モデルに欠損等の歪みが発生するという問題がある。これに対し、歪みの無いモデルを生成することを目的として、計測データからの移動物体検出・除去を行う手法に関する様々な研究が行われている。従来手法の一つとして、複数の異なる地点で計測・撮影された三次元点群や画像群を入力とし、三次元点群を複数フレームに投影し対応点の輝度情報の整合性を検証することで三次元点群中の移動物体を検出する手法が提案されている。この手法は、輝度情報のみを用いているため、移動物体と背景で輝度が類似している場合には、高精度に移動物体を検出することは難しい。一方、同一経路を複数回走行し撮影した画像群に対して時間方向と空間方向の位置合わせを行うことで、移動物体の除去を行う手法が提案されているが、同一の経路に対しての計測に多くの時間が必要となり、撮影コストが大きいという課題が残されている。

---

\*奈良先端科学技術大学院大学 情報科学研究科 修士論文, NAIST-IS-MT1451064, 2016年3月14日.

本研究では，一回の移動計測により得られる同期の取れた，三次元点群，画像群およびカメラの位置姿勢情報を入力とし，輝度情報と奥行情報の整合性を統合したエネルギー関数を最小化することで三次元点群中の移動物体検出を行う手法を提案する．提案手法では，静止物体の表面の明るさや形状は短時間では変化しないことに着目し，基準フレームにおいて観測した三次元点群を他のフレーム上に投影し，複数フレーム間の輝度値と奥行値の差に基づき移動物体らしさを評価する．ここでは，実際の計測データにおける輝度値および奥行値の差と，手動でラベル付けした移動物体との関係から，移動物体らしさを評価する関数を設計し，これをエネルギー関数のデータ項として用いる．また，各点の隣接点との奥行値や輝度値の差に基づく平滑化項を設計する．輝度情報と奥行情報を統合したエネルギー関数を最小化するラベル付けを求めることによって，移動物体の検出精度の向上を図る．実験では，車載型の全方位レーザーライダーユニット (Velodyne)，カメラ，GPS・IMU により計測した屋外環境データを用いることで，提案手法の有効性を検証する．

## キーワード

移動物体検出，三次元点群，画像群，グラフカット

# Moving object detection from a point cloud measured by an on-vehicle system based on energy minimization using photometric and depth consistencies\*

Atsushi Takabe

## Abstract

3D models of outdoor environments have been used for several applications such as a virtual earth system and a driving simulator. Such 3D models are constructed on the basis of real-world measurement using an on-vehicle system equipped with laser rangefinders, cameras, and GPS/IMU. One problem here is that moving objects in the measured data lead to inaccurate 3D models when modeling outdoor environments including dynamic objects. To solve this problem, several methods to detect or remove moving objects from the data of outdoor environments have been proposed. One of the methods detects 3D points on moving objects considering photometric consistency calculated by projecting 3D points onto omnidirectional images captured at different positions. Since this method uses only photometric consistency, it is difficult to detect points on moving objects whose luminance values are similar to those on background static objects. Another method removes moving objects from a set of images captured at near positions by aligning them. One problem in this method is that a vehicle with a measurement system must drive along the same path several times, which

---

\*Master's Thesis, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT1451064, March 14, 2016.

increases the measurement cost.

This study proposes a moving object detection method for point clouds by minimizing an energy function based on photometric and depth consistencies assuming that input data consist of synchronized point clouds, images, and camera poses from a single sequence captured with a moving on-vehicle system. Our proposed method is based on the idea that the surface brightness and shape of static objects do not change for a short time. The likelihood of moving objects for 3D points is evaluated on the basis of the differences of depth and luminance values calculated by projecting 3D points obtained at a target frame onto other frames. We design a function that evaluates the likelihood of moving objects from the relationship between the ground truth labels given with manual operations and the differences of depth and luminance values between several frames, and use it for a data term of the energy function. A smoothness term of the energy function is designed on the basis of the difference of depth and luminance values between each point and its neighboring points. The proposed method improves the detection precision of points on moving objects by minimizing the energy function in which both depth and photometric information are simultaneously considered. The effectiveness of the proposed method is verified by experiments using data measured in outdoor environments by an on-vehicle system with a LiDAR unit (Velodyne), an omnidirectional camera, and a GPS/IMU.

**Keywords:**

Detection of moving objects, 3D point cloud, image, graph cut

# 目次

1. はじめに	1
2. 移動物体・除去に関する従来研究および本研究の位置づけと方針	4
2.1 移動物体検出・除去に関する従来研究	4
2.1.1 一回の移動計測データを入力とする手法	4
2.1.2 複数回の移動計測データを入力とする手法	9
2.2 本研究の位置づけと方針	10
3. 移動計測された屋外環境計測データからの 移動物体検出	12
3.1 エネルギー関数の定義	12
3.2 輝度情報の整合性	13
3.3 奥行情報の整合性	15
3.4 移動物体らしさ関数の定義	16
4. 屋外環境計測データからの移動物体検出実験	18
4.1 実験条件	18
4.2 移動物体らしさ関数の決定	34
4.3 移動物体領域の検出実験結果と定量的評価	34
4.3.1 輝度情報と奥行情報の整合性の確認	34
4.3.2 移動物体検出精度の定量的評価	36
4.4 考察	59
5. まとめ	62
謝辞	63
参考文献	64

## 目 次

1	車載型計測システム	2
2	移動物体により生じた不自然なテクスチャ例	2
3	Yan ら [15] の手法による移動物体の除去例	5
4	Shen ら [27] の手法による前景移動物体の除去例	6
5	Flores ら [10] の手法による歩行者除去の例	7
6	Kawai ら [12] の手法による移動物体除去の例	8
7	内山ら [13] の手法による移動物体除去の例	9
8	高橋ら [14] の手法による移動物体除去と色調統一の例	10
9	輝度情報の整合性	14
10	奥行情報の整合性	15
11	奥行画像の補間例	15
12	計測した点群と手動により指定した移動物体領域	16
13	データセット A の入力画像	20
14	データセット A の三次元点群	21
15	データセット B の入力画像 (Frame 1~Frame 10)	22
16	データセット B の入力画像 (Frame 11~Frame 20)	23
17	データセット B の三次元点群 (Frame 1~Frame 10)	24
18	データセット B の三次元点群 (Frame 11~Frame 20)	25
19	データセット C の入力画像 (Frame 1~Frame 10)	26
20	データセット C の入力画像 (Frame 11~Frame 20)	27
21	データセット C の三次元点群 (Frame 1~Frame 10)	28
22	データセット C の三次元点群 (Frame 11~Frame 20)	29
23	データセット D の入力画像 (Frame 1~Frame 10)	30
24	データセット D の入力画像 (Frame 11~Frame 20)	31
25	データセット D の三次元点群 (Frame 1~Frame 10)	32
26	データセット D の三次元点群 (Frame 11~Frame 20)	33
27	輝度誤差のヒストグラムと輝度情報に基づく移動物体らしさ	35
28	奥行誤差のヒストグラムと奥行情報に基づく移動物体らしさ	35

29	奥行情報と輝度情報に基づく整合性 (白：誤差大, 黒：誤差小) . . .	36
30	データセット B に対する移動物体領域検出結果 (Frame 1~10) . . .	38
31	データセット B に対する移動物体領域検出結果 (Frame 11~20) . . .	39
32	データセット B に対する輝度情報のみを用いた検出結果 (Frame 1 ~10) . . . . .	40
33	データセット B に対する輝度情報のみを用いた検出結果 (Frame 11 ~20) . . . . .	41
34	データセット B に対する奥行情報のみを用いた検出結果 (Frame 1 ~10) . . . . .	42
35	データセット B に対する奥行情報のみを用いた検出結果 (Frame 11 ~20) . . . . .	43
36	データセット B における 20 フレーム平均の比較グラフ . . . . .	43
37	データセット C に対する移動物体領域検出結果 (Frame 1~10) . . .	45
38	データセット C に対する移動物体領域検出結果 (Frame 11~20) . . .	46
39	データセット C に対する輝度情報のみを用いた検出結果 (Frame 1 ~10) . . . . .	47
40	データセット C に対する輝度情報のみを用いた検出結果 (Frame 11 ~20) . . . . .	48
41	データセット C に対する奥行情報のみを用いた検出結果 (Frame 1 ~10) . . . . .	49
42	データセット C に対する奥行情報のみを用いた検出結果 (Frame 11 ~20) . . . . .	50
43	データセット C における 20 フレーム平均の比較グラフ . . . . .	50
44	データセット D に対する移動物体領域検出結果 (Frame 1~10) . . .	52
45	データセット D に対する移動物体領域検出結果 (Frame 11~20) . . .	53
46	データセット D に対する輝度情報のみを用いた検出結果 (Frame 1 ~10) . . . . .	54
47	データセット D に対する輝度情報のみを用いた検出結果 (Frame 11 ~20) . . . . .	55

48	データセット D に対する奥行情報のみを用いた検出結果 (Frame 1 ～10) . . . . .	56
49	データセット D に対する奥行情報のみを用いた検出結果 (Frame 11 ～20) . . . . .	57
50	データセット D における 20 フレーム平均の比較グラフ . . . . .	57
51	奥行方向に低速移動する車両の検出結果 (データセット B Frame8)	59
52	地面上の点が誤検出されている例 . . . . .	61
53	画像全体が明るくなった例 . . . . .	61

## 表 目 次

1	各種センサの仕様 . . . . .	19
2	データセット B に対する移動物体領域検出結果の定量的評価値 . .	44
3	データセット B に対する輝度情報のみを用いた検出結果の定量的 評価値 . . . . .	44
4	データセット B に対する奥行情報のみを用いた検出結果の定量的 評価値 . . . . .	44
5	データセット C に対する移動物体領域検出結果の定量的評価値 . .	51
6	データセット C に対する輝度情報のみを用いた検出結果の定量的 評価値 . . . . .	51
7	データセット C に対する奥行情報のみを用いた検出結果の定量的 評価値 . . . . .	51
8	データセット D に対する移動物体領域検出結果の定量的評価値 . .	58
9	データセット D に対する輝度情報のみを用いた検出結果の定量的 評価値 . . . . .	58
10	データセット D に対する奥行情報のみを用いた検出結果の定量的 評価値 . . . . .	58

## 1. はじめに

バーチャルアースに代表される遠隔地のテレプレゼンスシステム [1] や、ドライビングシミュレータ [2,3] 等のアプリケーションが近年普及している。これらのアプリケーションでは、固定された三次元形状と表面テクスチャを利用したモデルベースドレンダリング [4,5] や、視点位置に応じて三次元形状やテクスチャを変更するイメージベースドレンダリング [6-8] により任意の視点からの映像提示が実現されている。このような手法で用いられる屋外環境の三次元形状および画像群は、図1に示すような、レーザレンジファインダ、カメラ、GPS・IMU等を組み合わせた車載システム等を用いて取得されることが多いが、計測対象環境に人や車などの移動物体が存在した場合には、計測データに移動物体が残るため、任意の視点からの映像を生成した際に、図2に示すような不自然なテクスチャが生成されるという問題がある。この問題に対して、移動計測された点群および画像データから、あらかじめ移動物体の検出または除去を行う手法が提案されている。

従来、移動計測された屋外実環境データから移動物体を検出する手法として、一回の移動計測によって得られる複数の異なる地点で計測・撮影された三次元点群や画像群を入力とし、三次元点群を複数フレームに投影し対応点の輝度情報の整合性を検証することで三次元点群中の移動物体を検出する手法が提案されている [9]。この手法は、輝度情報のみを用いているため、移動物体と背景で輝度が類似している場合には、高精度に移動物体を検出することが難しいという問題がある。また、一回の移動撮影によって得られる画像群を入力として移動物体の除去を行う手法も提案されている [10,11]。これらの手法では、背景を平面と仮定し、前後のフレーム間の対応付けを行うことで動物体の除去を行うため、平面仮定の適用できない一般的なシーンで用いることができず、汎用性が低いといった問題点がある。これに対し、一回の移動撮影により得られる全方位画像群を入力として、平面仮定を用いずに移動物体の除去を行う手法が提案されている [12]。この手法は、まず入力画像群から環境の三次元形状を復元し、自由視点画像生成に基づいて基準フレームの前後数フレームを基準フレームでの見え方に変換する。次に、生成された基準フレーム視点の画像群を比較することで移動物体の背景を取得することにより、画像上の移動物体を除去する。この手法では、平面仮定を用



図 1 車載型計測システム



図 2 移動物体により生じた不自然なテクスチャ例

いないため、汎用的であるが、その結果は三次元復元の精度に大きく依存する。

他方、同一経路を複数回移動撮影した全方位画像データを入力として、計測データを位置合わせし統合することによって動物体の存在しない全方位画像を取得する手法も提案されている [13,14]。これらの手法では、一時的に駐停車された自動車など、一回の撮影中には動かない動物体に対しても異なる時刻で撮影された画像を用いて除去できるという利点がある。しかし、同一経路を複数回撮影するために膨大な撮影コストがかかるといった問題がある。

本研究では、一回の移動計測により得られる同期の取れた三次元点群、画像群

およびカメラの位置姿勢情報を入力とし、輝度情報と奥行情報の整合性に基づき定義されるエネルギー関数を最小化することでフレーム毎に三次元点群中の移動物体領域の検出を行う手法を提案する。提案手法は、平面仮定を用いることなく、また三次元復元などの精度には依存せずに移動物体領域を検出できるという特長を持つ。

以下、2章では、移動物体検出・除去に関する従来研究および本研究の位置付けと方針について述べる。3章では、移動計測された屋外実環境測定データを用いて三次元点群の移動物体領域を検出する提案手法について述べる。4章では、車載型の全方位レーザーライダーユニット (Velodyne)、カメラ、GPS・IMUにより計測した屋外環境データを用いて実験を行い、提案手法の有効性を示す。最後に、5章でまとめと今後の展望、課題について述べる。

## 2. 移動物体・除去に関する従来研究および本研究の位置づけと方針

本研究の目的は、一度の移動計測により得られる三次元点群中から移動物体領域の検出を行うことである。本章では入力データから、移動物体の検出、または移動物体等の特定の物体領域の除去を行う関連研究を概観し、従来研究に対する本研究の位置づけと方針について述べる。

### 2.1 移動物体検出・除去に関する従来研究

入力データからの移動物体の検出、または移動物体等の特定の物体領域の除去を行う手法は入力データの種類によって以下のように分類できる。

- 一回の移動計測データを入力とする手法
  - － 三次元点群および画像群を入力として用いる手法
  - － 画像群を入力として用いる手法
- 同一地点に対する複数回の計測データを入力とする手法

以下、各手法について詳述する。

#### 2.1.1 一回の移動計測データを入力とする手法

[三次元点群および画像群を入力データとして用いる手法]

三次元点群および画像群を入力データとして用いる手法として、移動計測された点群および画像群から、移動物体の検出、または除去を行う手法が提案されている。Kanatani ら [9] は、一回の移動計測によって得られる三次元点群および経路上の異なる地点で撮影された画像群を入力とし、三次元点群を複数フレームに投影し対応点の輝度値の分散を検証することで三次元点群中の移動物体を検出する手法を提案している。この手法は、輝度情報のみを用いているため、移動物体

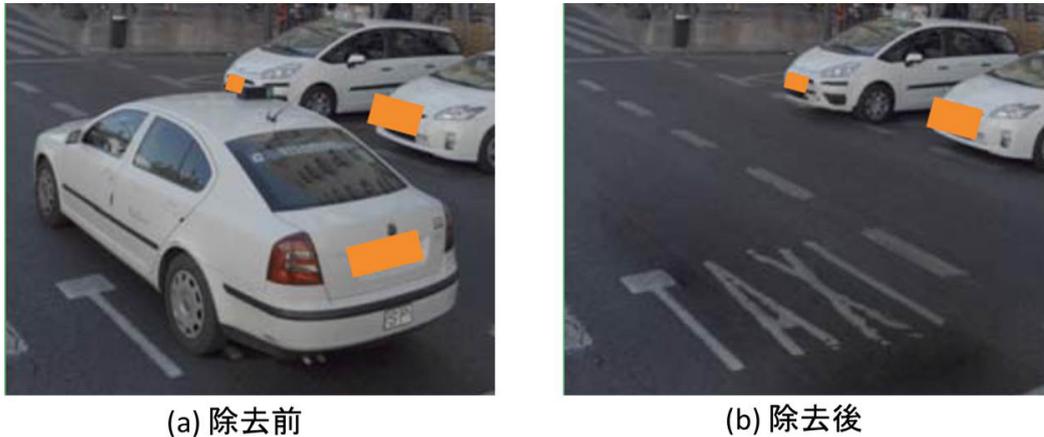


図 3 Yan ら [15] の手法による移動物体の除去例

とその背景とで輝度が類似している場合には，高精度に移動物体を検出することは難しい．また，移動物体は道路上に存在するという仮定を用いており，適用可能範囲が限定されている．Yan ら [15] は一回の移動計測データから移動物体を検出し，除去する手法を提案している (図 3)．この手法はまず，三次元点群に対してセグメンテーションを行い，基準フレームとその直後の 1 フレームでの各セグメントの大きさとの輝度値をもとに三次元点群に対してトラッキングベースの移動物体検出を行う．移動物体と判定されたセグメントの輪郭を画像上に投影し，GlabCut [16] を用いて画像上の移動物体領域を検出する．その後，位置合わせを行った他のフレームから移動物体の存在しないテクスチャを選択し，画像上の移動物体領域を上書きすることで移動物体の除去を行う．この手法では，移動計測が可能なレーザレンジファインダで取得された疎な三次元点群に対してトラッキングを行うため，トラッキングが失敗しやすく，検出精度がトラッキング結果に大きく依存するという問題点がある．Huang ら [17]，Premebida ら [18] は，機械学習を用いて屋外環境中の移動物体の検出を行っている．これらの手法は，画像と点群の両方を用いて学習を行い，移動物体の位置と向きを識別する．しかし，乗り物や人などの学習された特定の物体のみしか検出できず，未知の移動物体への対応は難しい．また，三次元空間をグリッドに分割し，各グリッドの物体の有無を表す占有グリッド [19] を用いた移動物体検出手法も提案されている [20-25]．一般的に，レーザレンジスキャナによるラスタスキャンで得られる三次元点群は

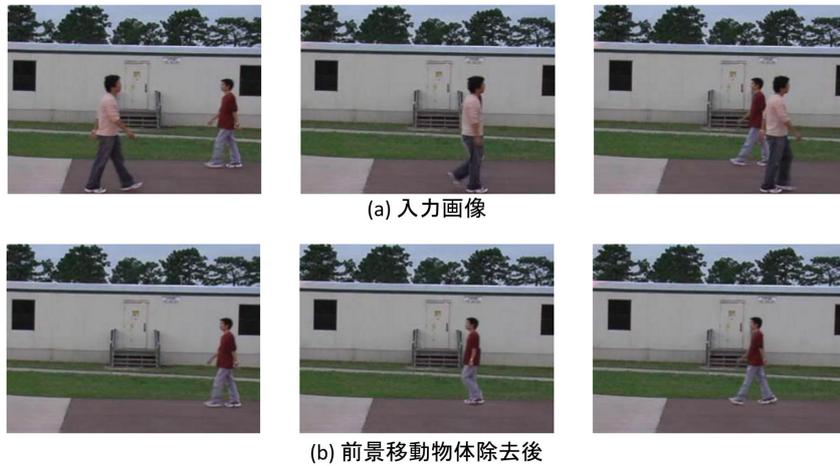


図 4 Shen ら [27] の手法による前景移動物体の除去例

疎であるため，フレーム毎に必ずしも静止物体上の同一箇所が計測されるとは限らない．このため，例えば，グリッドのサイズが小さすぎると，あるフレームで計測された箇所が次のフレームで計測されない場合などにおいて，静止物体が移動物体として誤検出される．すなわち，移動物体の検出結果がグリッドのサイズに大きく依存するという問題がある．

#### [画像群を入力データとして用いる手法]

画像群を入力データとして用いる手法として，一般的な単眼カメラで撮影された映像を対象とした手法と全方位カメラで撮影された全方位映像を対象とした手法が提案されている．一般的なカメラを用いて移動撮影された画像群を入力とする手法として，讓田ら [26] は，移動撮影したカメラの移動距離が小さく，画像群における撮影対象シーンがカメラから十分遠方にあるという仮定のもと，射影変換でフレーム間の対応付けを行うことで移動物体を除去している．また，Shen ら [27] は，パンチルトズームカメラで撮影された画像群を入力として，讓田ら [26] の手法と同様に対象シーンの静止物体領域を射影変換によって対応付け，移動物体の背景画像を生成する手法を提案している．この手法は，初期フレームでユーザが移動物体領域を手動で指定し，Mean-Shift [28] を用いてこれを追跡する．これにより，除去対象の背景に存在する等速直線運動をする移動物体は除去せずに，前景の移動物体のみの除去を行うことができる (図 4)．しかし，これらの手法は，



(a) 歩行者検出された入力画像



(b) 歩行者除去後の画像

図 5 Flores ら [10] の手法による歩行者除去の例

画像間で射影変換による位置合わせができないようなカメラの動きには、原理的に対応できない。これに対して、Matsushita ら [29] の手法では、背景の形状やカメラの動きに制約を置かずに、オプティカルフローを用いることで前後フレームを対応付け、対象の除去を行う。しかし、除去対象が移動物体の場合には、手動で毎フレーム対象領域を指定する必要がある。

全方位カメラを用いて撮影した全方位画像群を入力とした手法も提案されている [10–12, 30]。堀ら [30] は、一台の固定された全方位カメラによって撮影された全方位画像群を入力として、移動物体の除去を行っている。この手法では、輝度値の出現頻度を画素ごとに参照することで、移動物体の除去を行っている。しかしこの手法はカメラが固定されていることを前提としているため、移動撮影された画像群への適用は難しい。一方、一台の全方位カメラを用いて移動撮影した全方位画像群を入力とした手法として、Flores ら [10]、Kawai ら [11] は、除去対象の背景が平面であるという仮定を用いることで対象の除去を行う手法を提案している。Flores らの手法では、Google Street View における歩行者の背景は平面であると仮定し、入力画像上で Leibi ら [31] の手法によって歩行者を検出し、除去を行う。歩行者を検出したフレームに隣接するフレームの画像に対して射影変換を施し、検出した歩行者の領域に上書きすることで、歩行者の存在しない全方位画像群を生成している (図 5)。Kawai ら [11] は、全方位画像群において、全方位カメラの死角領域を除去の対象とし、その領域にはほぼ平面の地面が存在するとい



(a)入力画像



(b)移動物体除去後の画像

図 6 Kawai ら [12] の手法による移動物体除去の例

う仮定のもと、その領域を他のフレームのテクスチャを使って補間する手法を提案している。これらの手法では平面仮定を用いているため、汎用性が低いといった問題点が挙げられる。これに対して、Kawai ら [12] は、一回の移動撮影により得られる全方位画像群を入力として、平面仮定を用いずに移動物体の除去を行う手法を提案している。図 6 に Kawai ら [12] の手法による移動物体除去の例を示す。この手法は、まず入力画像群から Structure from Motion [32] と Multi-view Stereo [33] で環境の三次元形状を復元し、生成した全方位奥行き画像に基づいて基準フレームの前後数フレームを基準フレームでの見え方に変換する。次に、生成された基準フレーム視点の画像群を比較することで動物体の背景を取得することにより、画像上の移動物体を除去する。この手法では、平面仮定を用いないため汎用的であるが、その結果は三次元復元の精度に大きく依存する。



(a) 入力画像



(b) 移動物体除去後

図 7 内山ら [13] の手法による移動物体除去の例

### 2.1.2 複数回の移動計測データを入力とする手法

内山ら [13], 高橋ら [14] は同一経路を複数回撮影する手法を提案している。これらの手法は、異なる時刻に撮影された複数の画像間の位置合わせを行い、それらを統合することで移動物体領域を除去する手法である。図7に内山ら [13] の手法による移動物体除去の例を示す。内山ら [13] の手法では、まず同一経路を複数回走行して得られた全方位画像群のフレーム間で基準フレームに撮影位置に近い複数の画像を選択し、非剛体レジストレーションにより画素の対応付けを行う。次に、補正した画像群をサブウィンドウ単位に領域分けし、サブウィンドウ単位でエネルギーを最小化するように部分画像を選択し、統合することで移動物体の存在しない全方位画像群を生成する。このエネルギー関数は、ベクトルメディアンフィルタ [34] に基づき、部分画像に移動物体が存在するかどうかの尤もらしさを表す項と、隣接する部分画像間の連続性を考慮した項で構成される。高橋ら [14] は、車載全方位カメラで異なる時刻に密に撮影された画像群を入力として、線形濃度変換パラメータの推定と移動物体候補領域の推定の二つの処理を交互に繰り返しながら、移動物体除去を行いながら対象シーンの色調統一を行っ



図 8 高橋ら [14] の手法による移動物体除去と色調統一の例

ている (図 8 参照). これらの手法は一時的に駐停車中の車など, 一回の撮影中には動かない移動物体に対しても, 異なる時刻に撮影された画像群を用いて除去できるという利点はあるものの, 同一経路を複数回撮影しなければならないため, 広域な範囲の計測データを対象とした場合には, 膨大な撮影コストがかかるといった問題がある.

## 2.2 本研究の位置づけと方針

前節で概観したように, 屋外実環境計測データを様々なアプリケーションに利用する上で, 移動物体が三次元モデル生成や映像提示の妨げになるなどの問題点を解決するために, 自動車や歩行者などの移動物体を入力データから検出および除去することが求められており, それを実現する手法がすでに多く提案されている. 従来研究において, 屋外環境の一回の移動計測により得られる三次元点群および画像群を入力とする手法では, 検出精度の向上のために移動物体の検出範囲を限定していたり, 未知の移動物体への対応が困難であるといった問題点があった. また, 三次元点群に対するトラッキングを用いた手法では, 移動計測が可能なレーザレンジファインダで取得された疎な三次元点群に対してトラッキングを

行うため、トラッキングが失敗しやすく、検出精度がトラッキング結果に大きく依存するという問題点がある。一方、移動撮影された画像群を入力とする手法では、多くの手法で平面仮定を用いたり、カメラ移動を制限しており、汎用性が低い。平面仮定やカメラの移動に制限を用いない場合には撮影された画像のすべてのフレームに対して移動物体領域を手動で指定する必要があることや、結果が三次元復元の精度に依存するという点が挙げられる。これに対して、同一経路を複数回移動撮影する手法では、平面仮定などは用いずに、かつ一時的に駐停車中の移動物体なども検出除去できるという利点はあるものの、撮影コストが高いといった問題点があった。このように従来手法では、汎用性と撮影コストがトレードオフの関係になっていた。

本研究では、一回の移動計測により得られる同期の取れた三次元点群、画像群およびカメラの位置姿勢情報を入力とし、輝度情報と奥行情報の整合性を統合したエネルギー関数を最小化することでフレーム毎に三次元点群中の任意の移動物体領域の検出を行う手法を提案する。提案手法では、輝度情報だけではなく、奥行情報も用いたエネルギー関数を最適化することによって、検出範囲を限定することなく、またトラッキングの精度等には依存せずに移動物体領域を検出できる。

### 3. 移動計測された屋外環境計測データからの 移動物体検出

提案手法では、レーザレンジファインダ、カメラ、GPS・IMUを組み合わせた車載システムを用いて移動計測することによって得られた、同期のとれた三次元点群、画像群、カメラ位置姿勢情報を入力として、三次元点群中の移動物体領域を検出する。なお、三次元点群はレーザレンジファインダによるラスタスキャンで全周計測され、各点の隣接関係は既知であるものとする。また、計測された全周三次元点群と画像群はフレーム単位で同期されているものとする。

本研究では、静止物体の表面の明るさや形状は短時間では変化しないことに着目し、複数フレーム間の奥行値と輝度値の整合性に基づくエネルギー関数を定義し、それをグラフカット [35] を用いて最小化することで、各フレームの点群に含まれる移動物体上の点を検出する。また、実際の計測データにおける複数フレーム間の奥行値および輝度値の整合性と、手動でラベル付けした移動物体との関係から、移動物体らしさを評価する関数をあらかじめ設計し、これをエネルギー関数に用いることで、検出精度を向上させる。以下、3.1 節ではエネルギー関数の定義、3.2 節で輝度情報の整合性、3.3 節では奥行情報の整合性、3.4 節では移動物体らしさを評価する関数について述べる。

#### 3.1 エネルギー関数の定義

提案手法では、各計測点に対する移動物体・静止物体のラベルを決定する問題を、各点の移動物体らしさおよび各点の隣接点との関係を考慮したエネルギーの最小化問題として定式化する。具体的には、基準フレームにおいて各計測点の移動物体または静止物体を表すラベル  $X$  をパラメータとするエネルギー関数  $E$  を以下のように定義する。

$$E(X) = \sum_{v \in V} g_v(X_v) + \kappa \sum_{(u,v) \in N} h_{uv}(X_u, X_v) \quad (1)$$

ここで、 $\kappa$  は係数、 $V$  は基準フレームにおいて計測された点の集合、 $N$  は隣接する 2 点の組の集合である。このエネルギー関数は、第一項がデータ項であり、観

測データに基づくラベルの尤もらしさを表す。また、第二項は隣接するラベルの平滑化項である。

エネルギー関数  $E$  の第一項  $g_v$  は、基準フレームにおいて計測した各点  $v$  をその前後数フレームに投影したときの、フレーム間の輝度値や奥行値の差に基づいて算出する。具体的には、奥行情報の整合性に基づく点  $v$  の移動物体らしさを  $M_{D,v}$ 、輝度情報の整合性に基づく移動物体らしさを  $M_{P,v}$  としたとき、 $g_v$  を以下のように定義する。

$$g_v(X_v) = \begin{cases} (1 - M_{D,v}) + \alpha(1 - M_{P,v}) & (X_v : \text{移動物体}) \\ M_{D,v} + \alpha M_{P,v} & (X_v : \text{静止物体}) \end{cases} \quad (2)$$

ただし、 $\alpha$  は重み係数である。この項では、移動物体らしさが高い場合にはエネルギー最小化によって移動物体のラベルが選択されやすくなり、移動物体らしさが低い場合は静止物体のラベルが選択されやすくなる。

エネルギー関数  $E$  の第二項  $h_{uv}$  は、隣接するラベルに対する平滑化項であり、ここでは各点  $v$  とその隣接点  $u$  の奥行値  $d_v, d_u$ 、および点  $v, u$  を観測したフレーム  $n$  に投影した点  $p_{v,n}, p_{u,n}$  の画素値  $I(p_{v,n}), I(p_{u,n})$  (HSV 表色系の明度  $V$ ) に基づき以下のように定義する。

$$h_{u,v}(X_u, X_v) = \begin{cases} 0 & (X_v = X_u) \\ \frac{1}{|d_u - d_v| + |I(p_{u,n}) - I(p_{v,n})| + \epsilon} & (\text{otherwise}) \end{cases} \quad (3)$$

ここで、 $\epsilon$  は式 (3) の分母が 0 とならないための定数である。この項は、隣接した点と同じラベルを持つ場合に 0 を返し、隣接した点が違うラベルを持つ場合には奥行値や輝度値が近いほど大きい値を返す。したがって、最小化により、隣接した点の奥行値や輝度値が近いほど同じラベルが選択されやすくなる。

以下、式 (2) で用いる各点  $v$  の輝度情報の整合性に基づく移動物体らしさ  $M_{P,v}$ 、奥行情報の整合性に基づく移動物体らしさ  $M_{D,v}$  の算出方法について述べる。

### 3.2 輝度情報の整合性

図 9 に示すように、基準フレームにおける移動物体上の観測点  $v$  を異なる時刻に撮影した複数フレームの画像上に投影した場合、異なる物体の領域に投影され

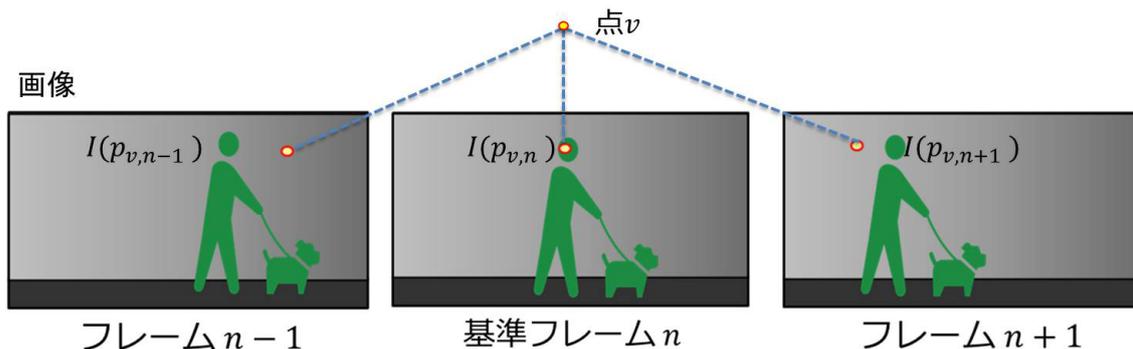


図 9 輝度情報の整合性

るため、画像上に投影された画素の輝度値は一致しないことが多い。そこで、レーザレンジファインダによって測定された三次元点を各入力画像上へ投影し、各投影点のフレーム間での輝度値の差を検証することで各点に対する移動物体らしさを評価する。

具体的には、図9に示すように、基準フレーム  $n$  において測定された三次元点  $v$  を基準フレーム  $n$  の画像上に投影した画素  $p_{v,n}$  の画素値  $I(p_{v,n})$  と、点  $v$  をフレーム  $n+i$  ( $-m \leq i \leq m$ ) の画像上に投影したときの画素  $p_{v,n+i}$  の画素値  $I(p_{v,n+i})$  を用いて、輝度情報の整合性に基づく移動物体らしさ  $M_{P,v}$  を以下のように算出する。

$$M_{P,v} = F_P \left( \max_i (I_{err}(i)) \right) \quad (4)$$

ただし、輝度誤差  $I_{err}(i)$  は以下のように定義する。

$$I_{err}(i) = \begin{cases} 0 & (D_{v,n+i} - d(p_{v,n+i}) > Th_d) \\ |I(p_{v,n}) - I(p_{v,n+i})| & (otherwise) \end{cases} \quad (5)$$

基準フレーム  $n$  に観測された点  $v$  が、他のフレームでは遮蔽される場合があるため、 $Th_d$  を実験的に求めた閾値とし、後述する各フレームで取得した点群から得られる奥行画像に点  $v$  を投影した画素  $p_{v,n+i}$  の奥行値  $d(p_{v,n+i})$  と、点  $v$  をフレーム  $n+i$  に投影して計算される奥行値  $D_{v,n+i}$  を用い、 $D_{v,n+i} - d(p_{v,n+i}) > Th_d$  を満たすフレームは遮蔽領域とみなし輝度誤差  $I_{err}(i)$  を 0 とする。また、式(4)中の  $F_P(\cdot)$  ( $0 \leq F_P(\cdot) \leq 1$ ) は、あらかじめ各点の輝度誤差と手動でラベリングされた移動物体との対応関係に基づいて定義された、輝度情報の整合性に基づく移動物体らしさを評価する関数であり、3.4節で詳述する。

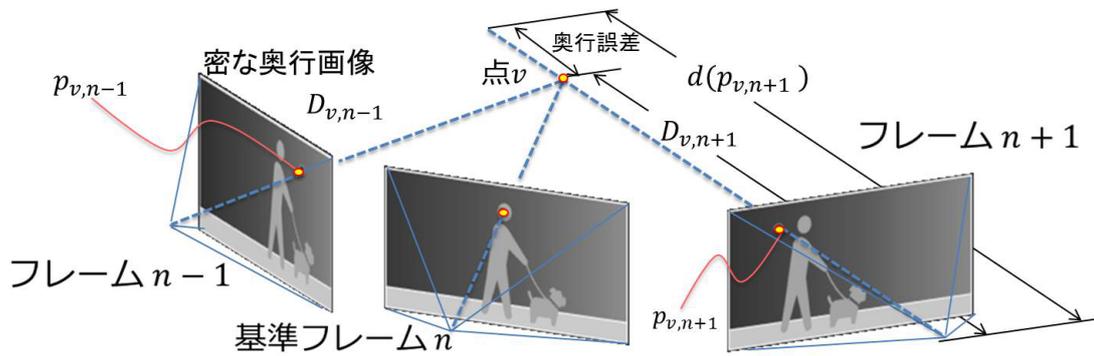
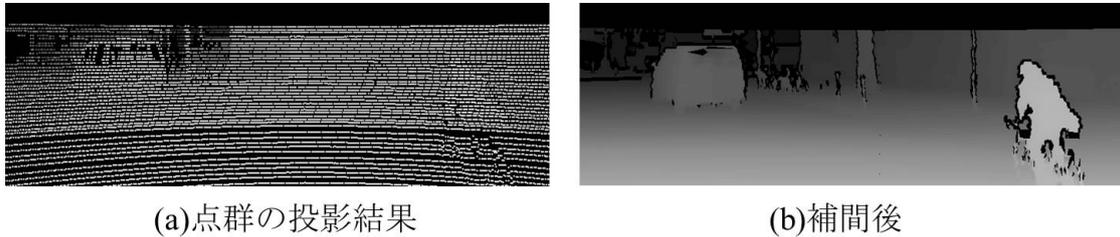


図 10 奥行情報の整合性



(a)点群の投影結果

(b)補間後

図 11 奥行画像の補間例

### 3.3 奥行情報の整合性

前節で述べた輝度情報の整合性と同様に、基準フレーム  $n$  における移動物体上の観測点  $v$  を異なる時刻に取得した他のフレーム  $n+i$  ( $i \neq 0$ ) の画像上に投影した場合においては、各フレームで取得した点群から得られる奥行値と、点  $v$  をそのフレームに投影し計算される奥行値が一致しない (図 10 参照). 本研究では、各フレームにおいて取得された三次元点からそのフレームの密な奥行画像を生成しておき、基準フレームの点群をその前後数フレームの奥行画像上に投影し奥行値の誤差を検証することで、移動物体らしさを評価する.

具体的には、図 11 に示すように、各入力視点において取得された三次元点群から、隣接する点を結んだ面を生成し、奥行を補間することで入力視点毎に密な奥行画像を生成する. 次に図 10 に示したように、基準フレーム  $n$  において測定された三次元点  $v$  をその前後  $m$  フレームの密な奥行画像に投影し、フレーム  $n+i$  ( $-m \leq i \leq m$ ) における点  $v$  までの奥行を  $D_{v,n+i}$ , 点  $v$  をフレーム  $n+i$  に

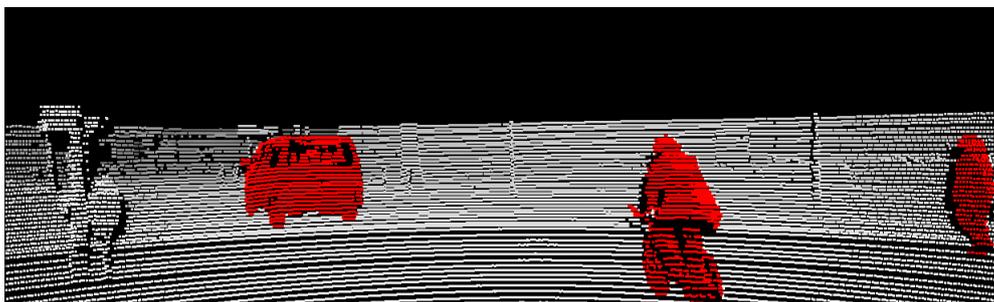


図 12 計測した点群と手動により指定した移動物体領域

投影した画素  $p_{v,n+i}$  の奥行値を  $d(p_{v,n+i})$  とし、奥行情報の整合性に基づく移動物体らしさ  $M_{D,v}$  を以下のように算出する.

$$M_{D,v} = F_D \left( \max_i (D_{err}(i)) \right) \quad (6)$$

ただし、奥行誤差  $D_{err}(i)$  は以下のように定義する.

$$D_{err}(i) = \begin{cases} 0 & (D_{v,n+i} - d(p_{v,n+i}) > Th_d) \\ d(p_{v,n+i}) - D_{v,n+i} & (otherwise) \end{cases} \quad (7)$$

輝度情報の整合性と同様に、遮蔽領域と考えられる  $D_{v,n+i} - d(p_{v,n+i}) > Th_d$  を満たすフレームでは奥行誤差  $D_{err}(i)$  を 0 とする. ここで,  $F_D(\cdot) (0 \leq F_D(\cdot) \leq 1)$  は, あらかじめ各点の奥行誤差と手動でラベリングされた移動物体との対応関係に基づいて定義された, 奥行情報の整合性に基づく移動物体らしさを評価する関数であり, 次節で詳述する.

### 3.4 移動物体らしさ関数の定義

本研究では, 移動物体の検出精度を高めるために, 実際の計測結果における, 式 (5), (7) で算出される輝度誤差および奥行誤差と, 図 12 に示すような手動でラベリングされた移動物体との対応関係に基づいて, 輝度の整合性に基づく移動物体らしさ関数  $F_P(\cdot)$  および奥行の整合性に基づく移動物体らしさ関数  $F_D(\cdot)$  を決定する.

具体的にはまず、実際の計測結果において、図 12 に示すように手動でラベリングを行った真値を作成する。次に、移動物体領域および静止物体領域それぞれにおいて輝度誤差および奥行誤差のヒストグラムを作成する。最後に、輝度誤差および奥行誤差の各ヒストグラムにおいて、階級ごとに各領域の度数の比を計算し、これをルックアップテーブルに格納したものを移動物体らしさとして用いる。

## 4. 屋外環境計測データからの移動物体検出実験

提案手法の有効性を評価するために、屋外環境を移動計測することによって得られた、同期のとれた三次元点群、画像群、カメラの位置姿勢情報を用いて、三次元点群中の移動物体領域を検出する実験を行った。

### 4.1 実験条件

本実験では、提案手法を用いた移動物体の自動検出システムを作成し、車載システムを用いて移動計測することによって得られた公開データセット [36] を用いた定量的評価を行った。このデータセットには実験車両に取り付けられた全周の計測が可能なレーザレンジファインダ (Velodyne HDL-64E) によって計測された三次元点群、外部カメラ (Point Grey Flea 2 (FL2-14S3C-C)) によって取得された画像群、GPS・IMU(OXTS RT 3003) により取得されたカメラの位置姿勢情報が含まれている。各種センサの仕様を表1に示す。

ここでは、図13, 14に示すデータセットAの連続する10フレームの三次元点群に対して手動で移動物体をラベリングし、移動物体らしさ関数の算出を行った。また、これにより得られた移動物体らしさ関数を用いて、図15~18に示すデータセットBの連続する20フレーム、図19~22に示すデータセットCの連続する20フレーム、図23~26に示すデータセットDの連続する20フレームに対して提案手法を適用し三次元点群中の移動物体領域の検出を行った。それぞれのデータセットの特徴として、データセットAは歩行者、自転車、自動車の存在する環境の計測データであり、データセットBは複数の自動車が存在する環境の計測データ、データセットCはカメラの奥行方向に移動する自動車の存在する環境の計測データ、データセットDは、低速で移動計測を行った、複数の歩行者が存在する環境の計測データである。なお、移動物体領域の検出実験では、基準フレームの前後5フレームを用いて奥行・輝度情報に基づく整合性を計算し、式(1)で用いるパラメータ $\kappa$ 、式(2)で用いるパラメータ $\alpha$ は予備実験を行いそれぞれ設定する。以下では、まずデータセットAを用いた移動物体らしさ関数の決定について述べ、次にデータセットB, C, Dを用いた移動物体検出の定量的評価について

表 1 各種センサの仕様

Velodyne HDL-64E	
点群取得間隔	0.1sec
取得点数：全周計測一回	約 13 万点
角度分解能	0.09°
俯角範囲	0°～26.8°
距離計測誤差	±20mm
計測可能範囲	≤ 120m
Point Grey Flea 2	
撮影間隔	0.1sec
解像度 (pixel)	1242 × 375
OXTS RT 3003	
データ取得間隔	0.1sec
GPS・IMU 分解能	0.02m / 0.1°

述べる。最後に、データセット B, C, D に対してエネルギー関数  $E$  として、輝度情報または奥行情報のみを用いて移動物体検出を行った場合の結果と提案手法を比較することで、提案手法の有効性を評価する。

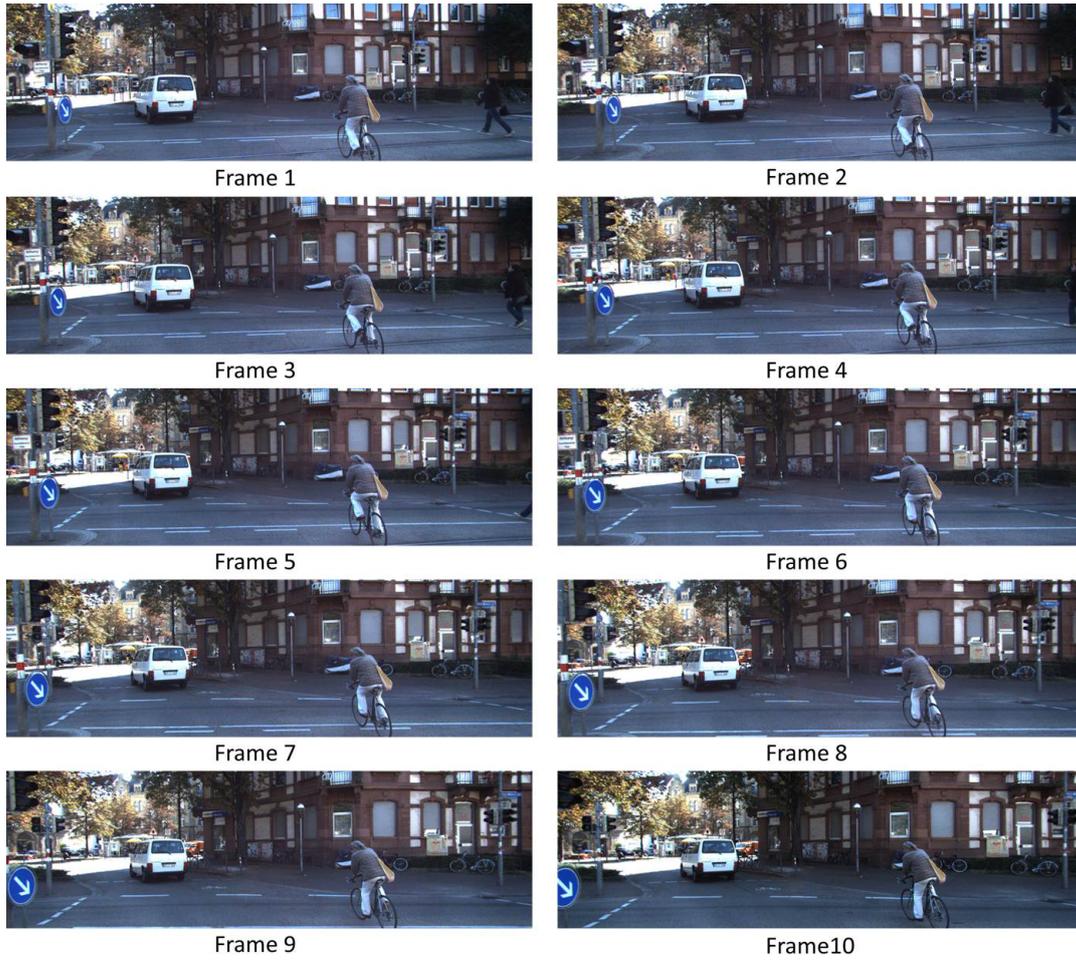


図 13 データセット A の入力画像

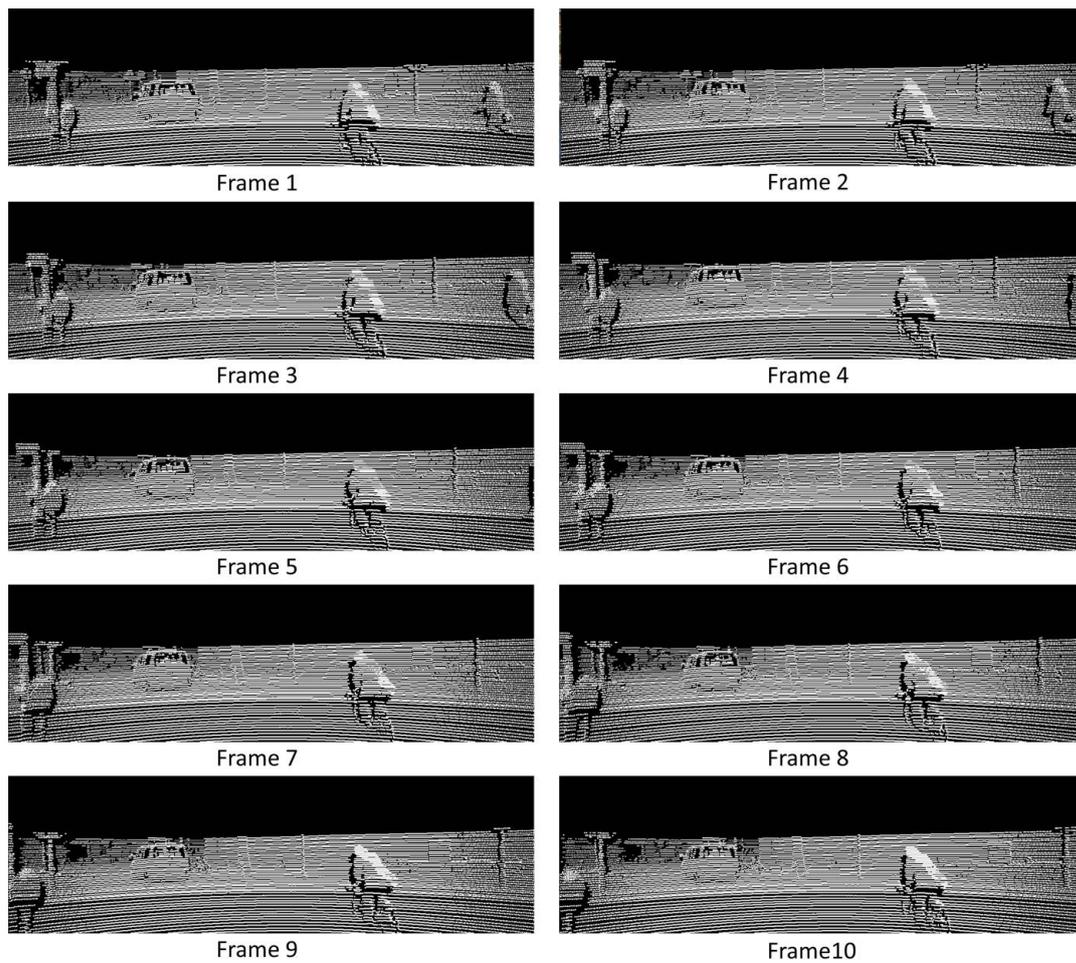


図 14 データセット A の三次元点群

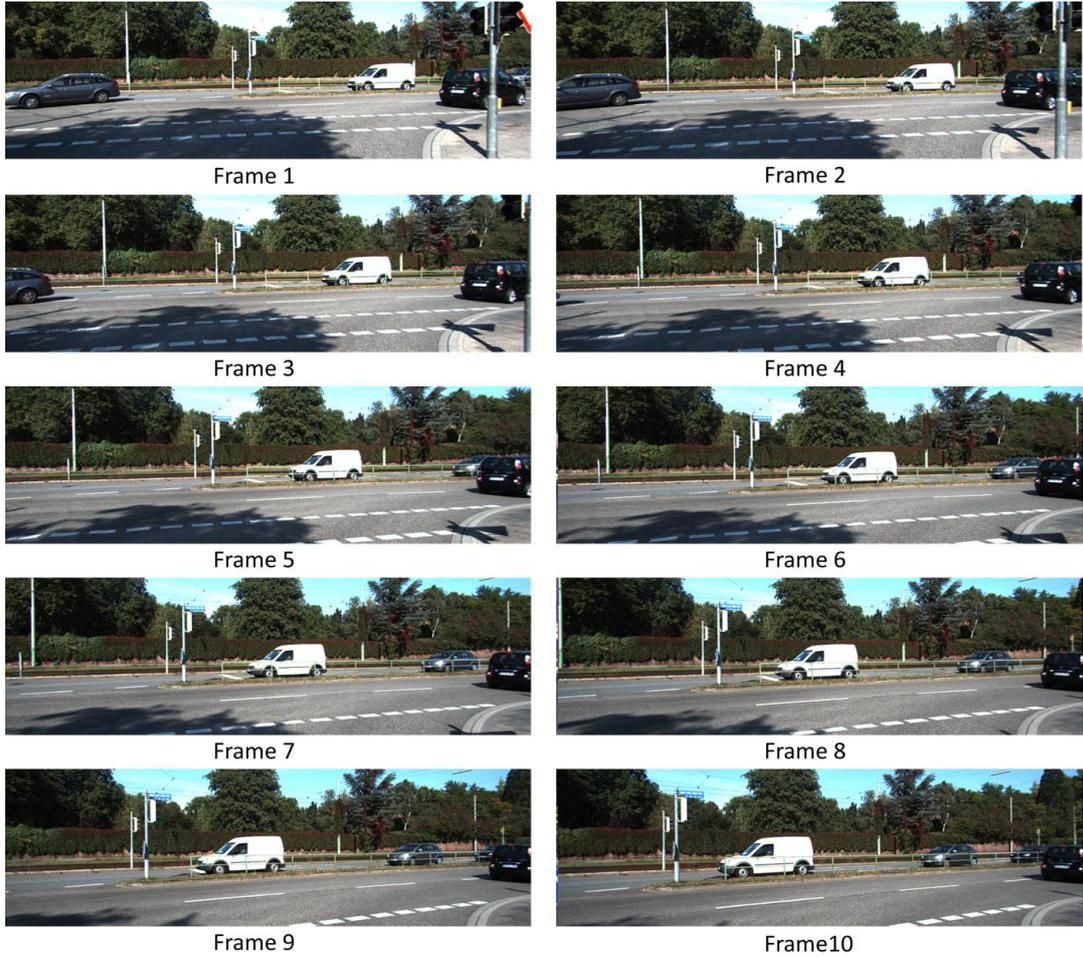


図 15 データセット B の入力画像 (Frame 1~Frame 10)

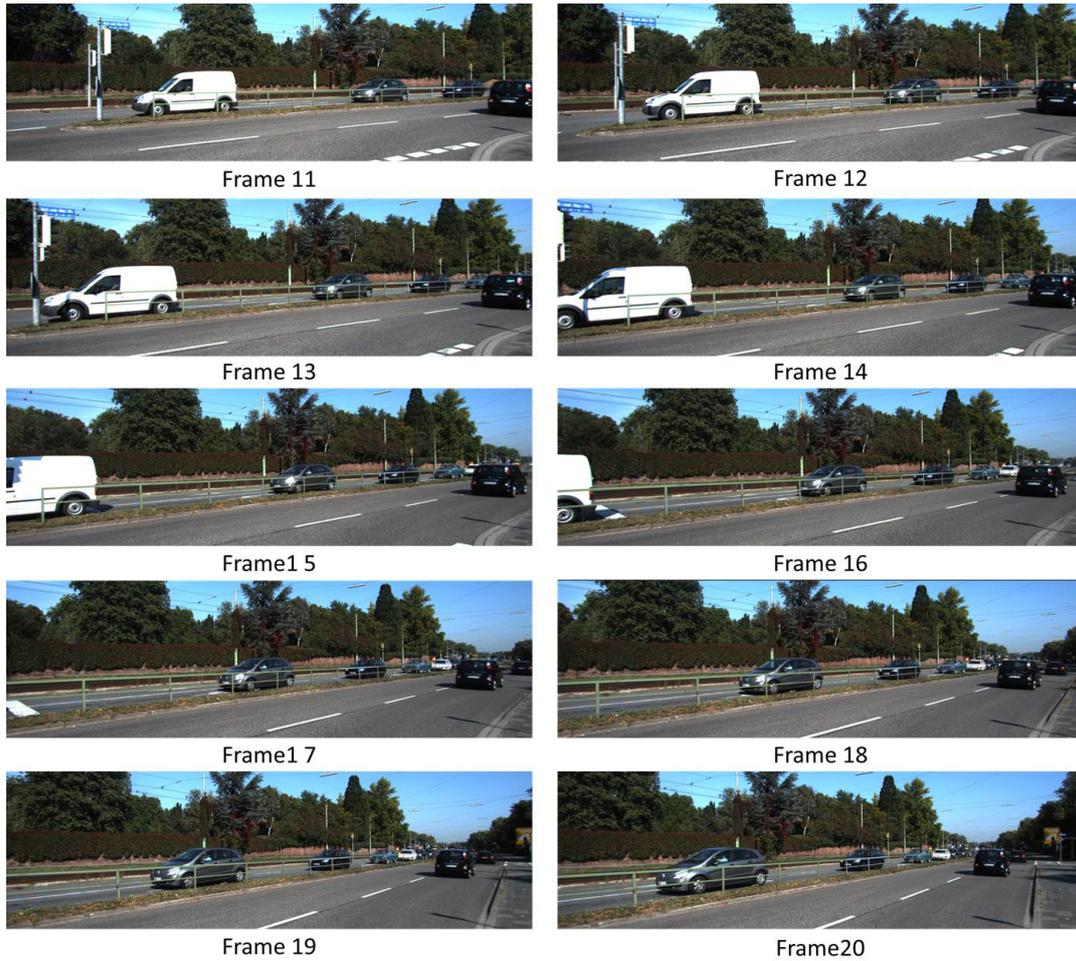


図 16 データセット B の入力画像 (Frame 11~Frame 20)

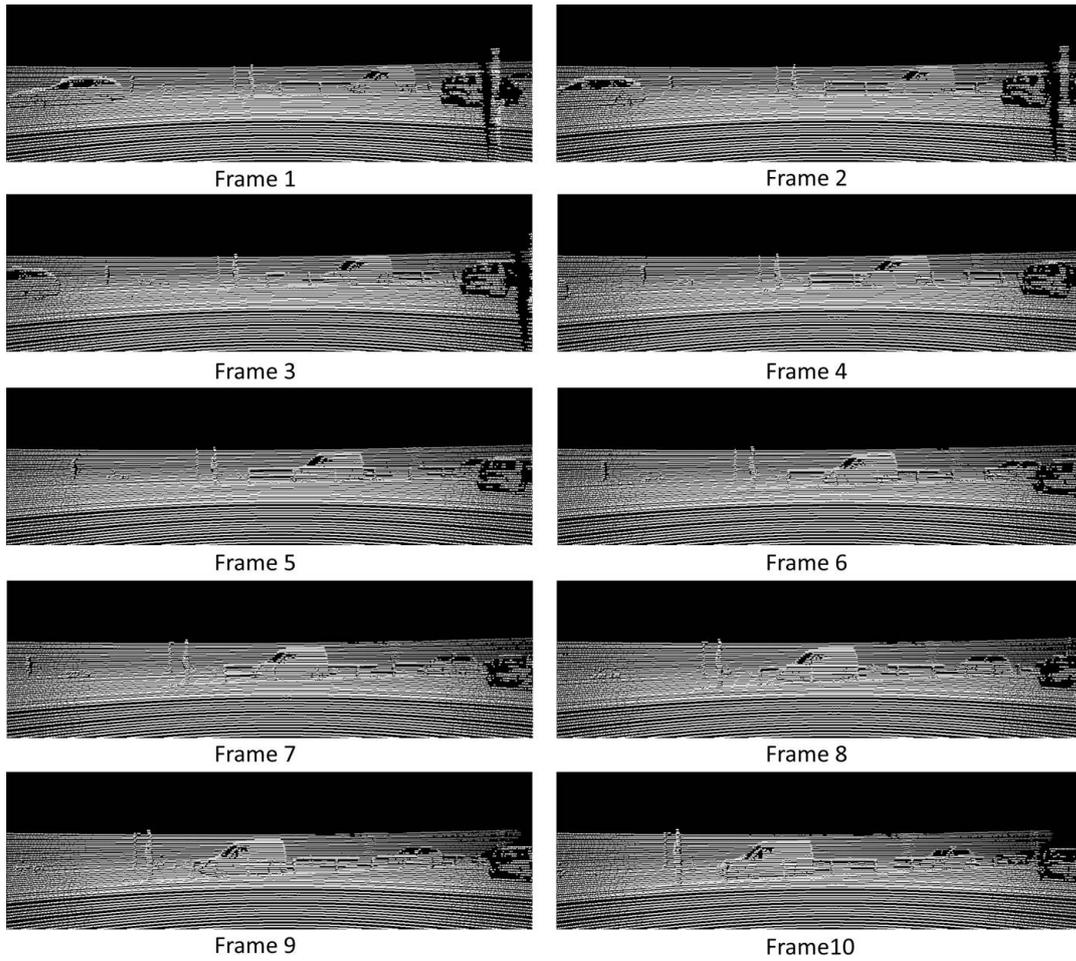


図 17 データセット B の三次元点群 (Frame 1~Frame 10)

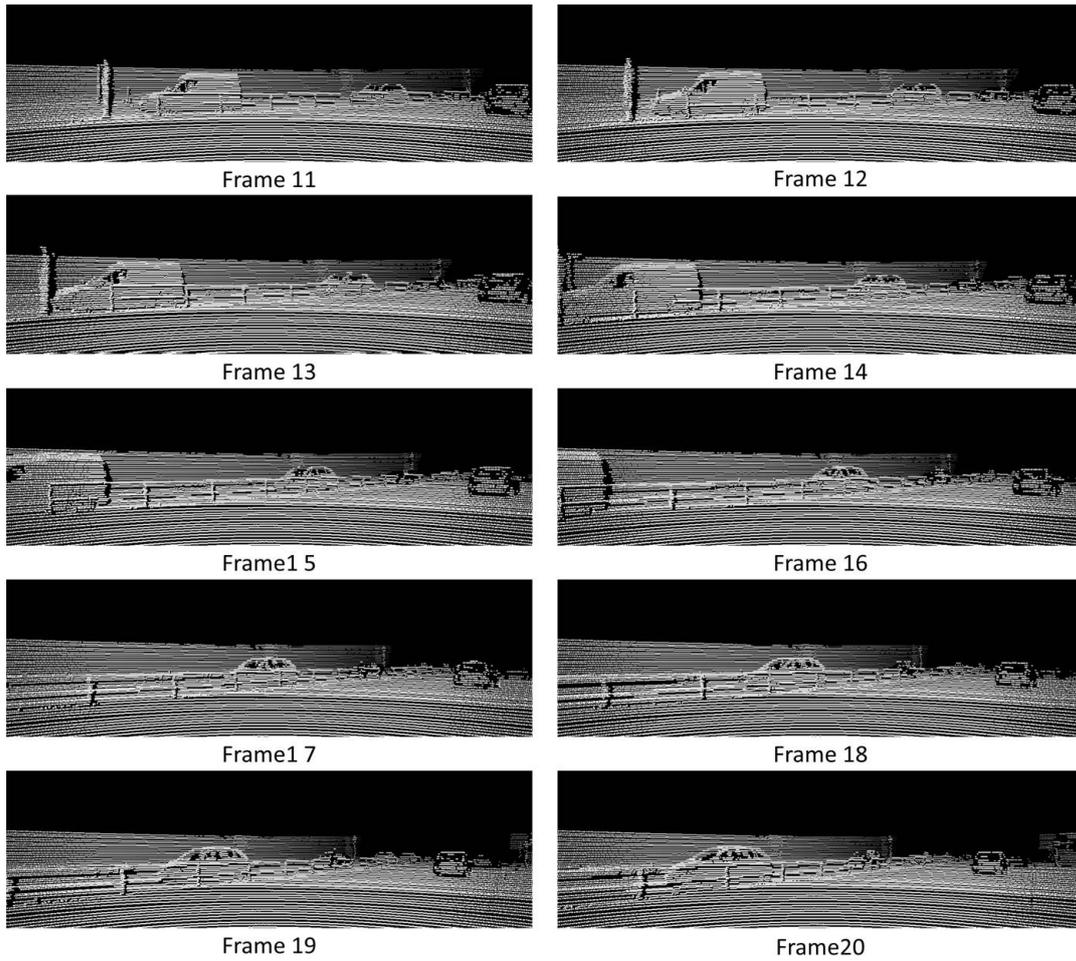


図 18 データセット B の三次元点群 (Frame 11~Frame 20)

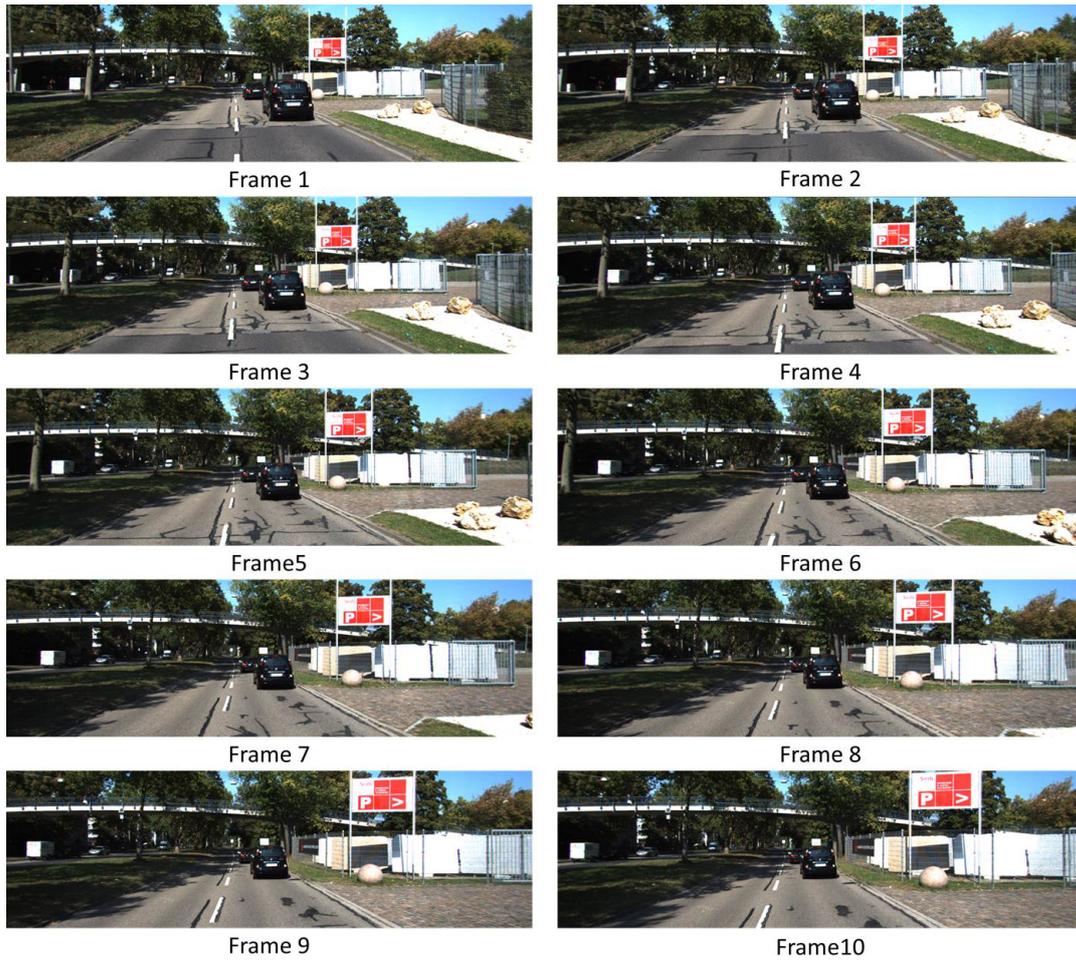


図 19 データセット C の入力画像 (Frame 1~Frame 10)



図 20 データセット C の入力画像 (Frame 11~Frame 20)

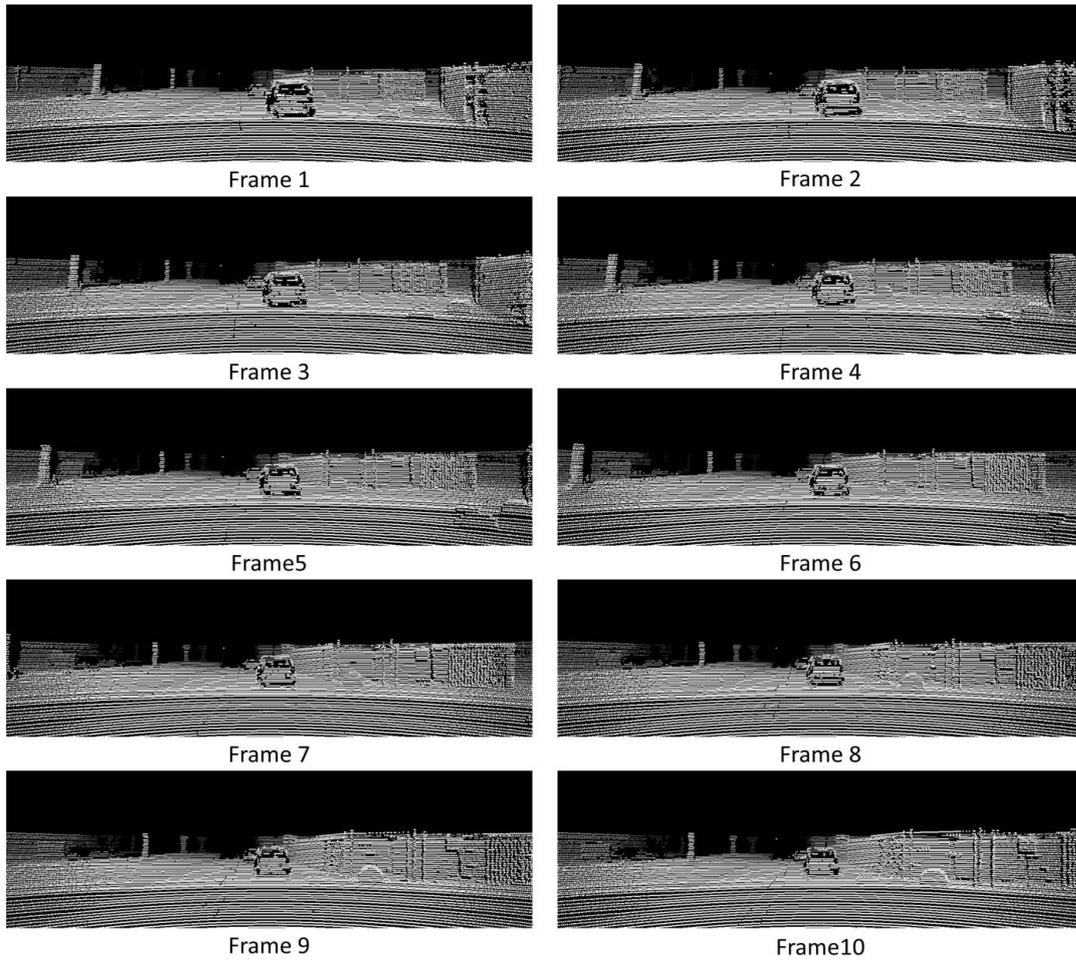


図 21 データセット C の三次元点群 (Frame 1~Frame 10)

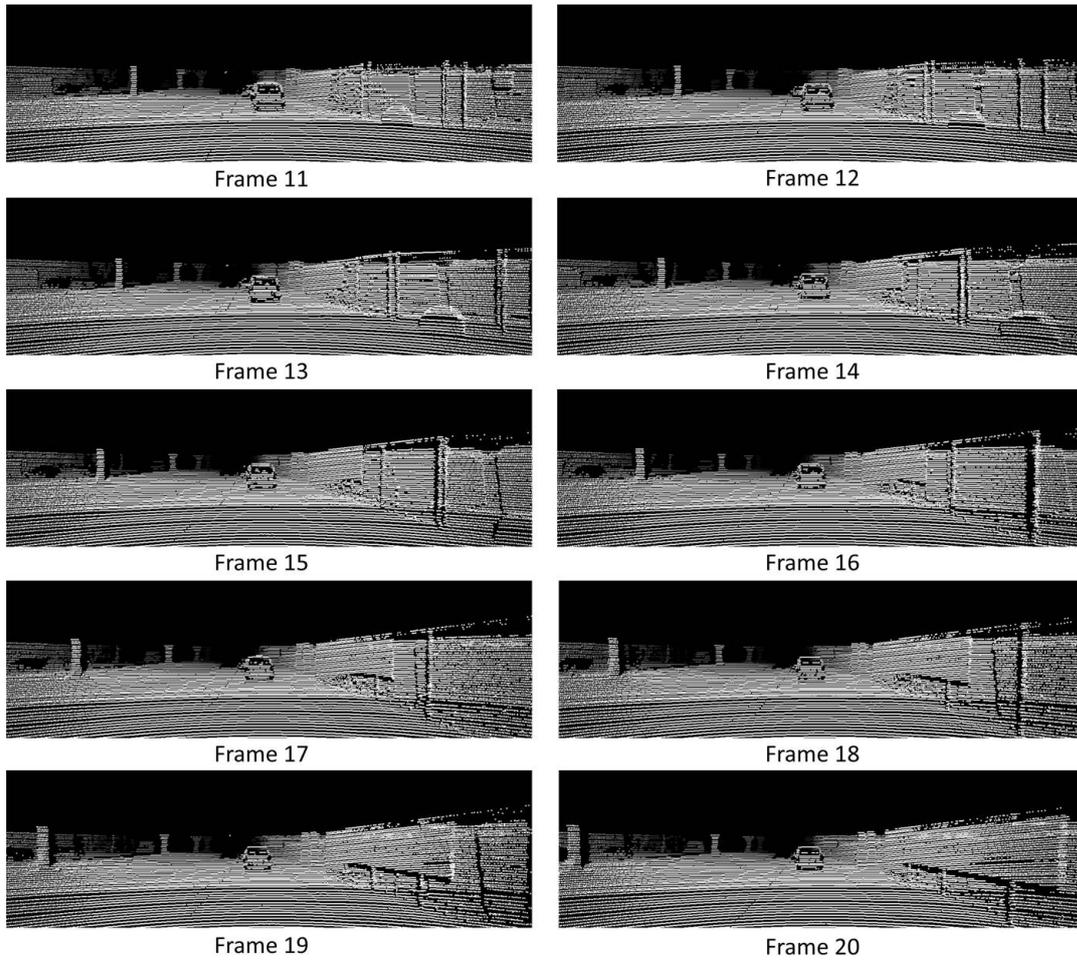


図 22 データセット C の三次元点群 (Frame 11~Frame 20)

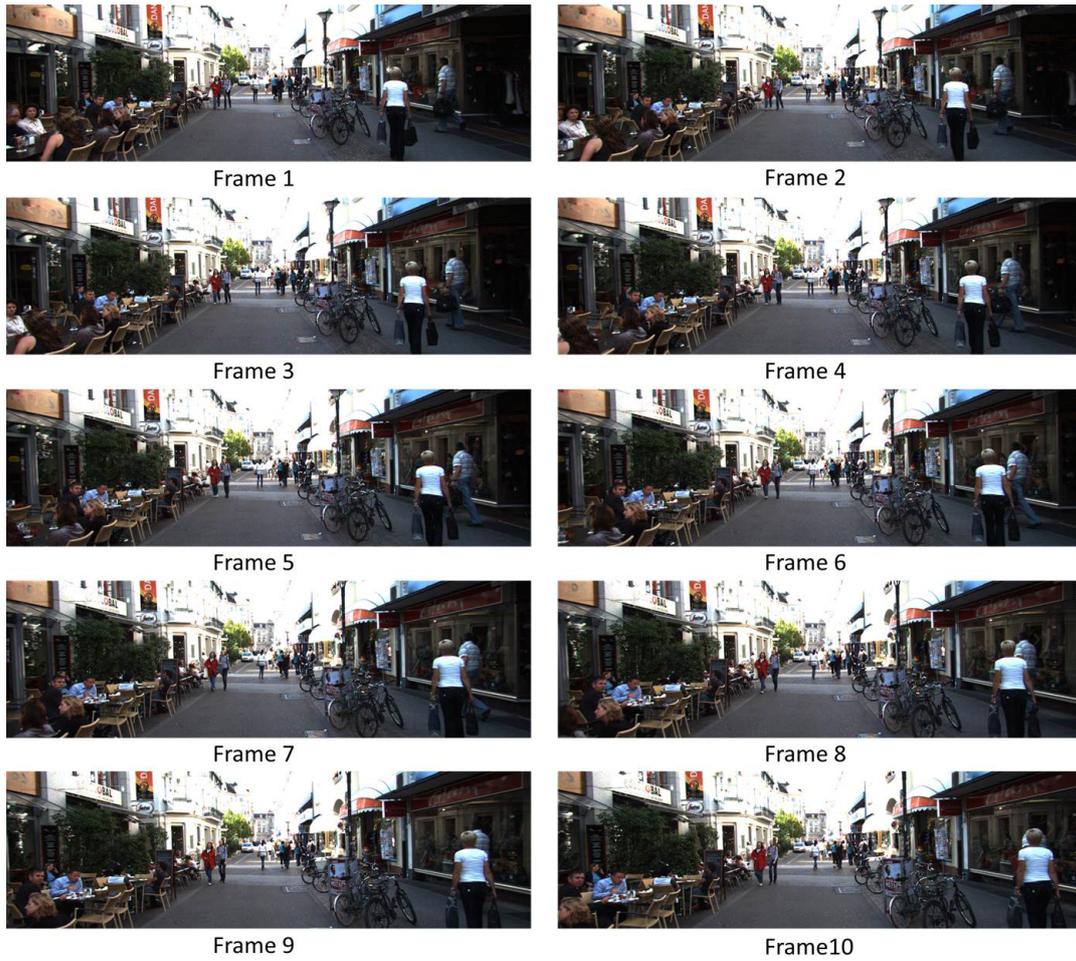


図 23 データセット D の入力画像 (Frame 1~Frame 10)

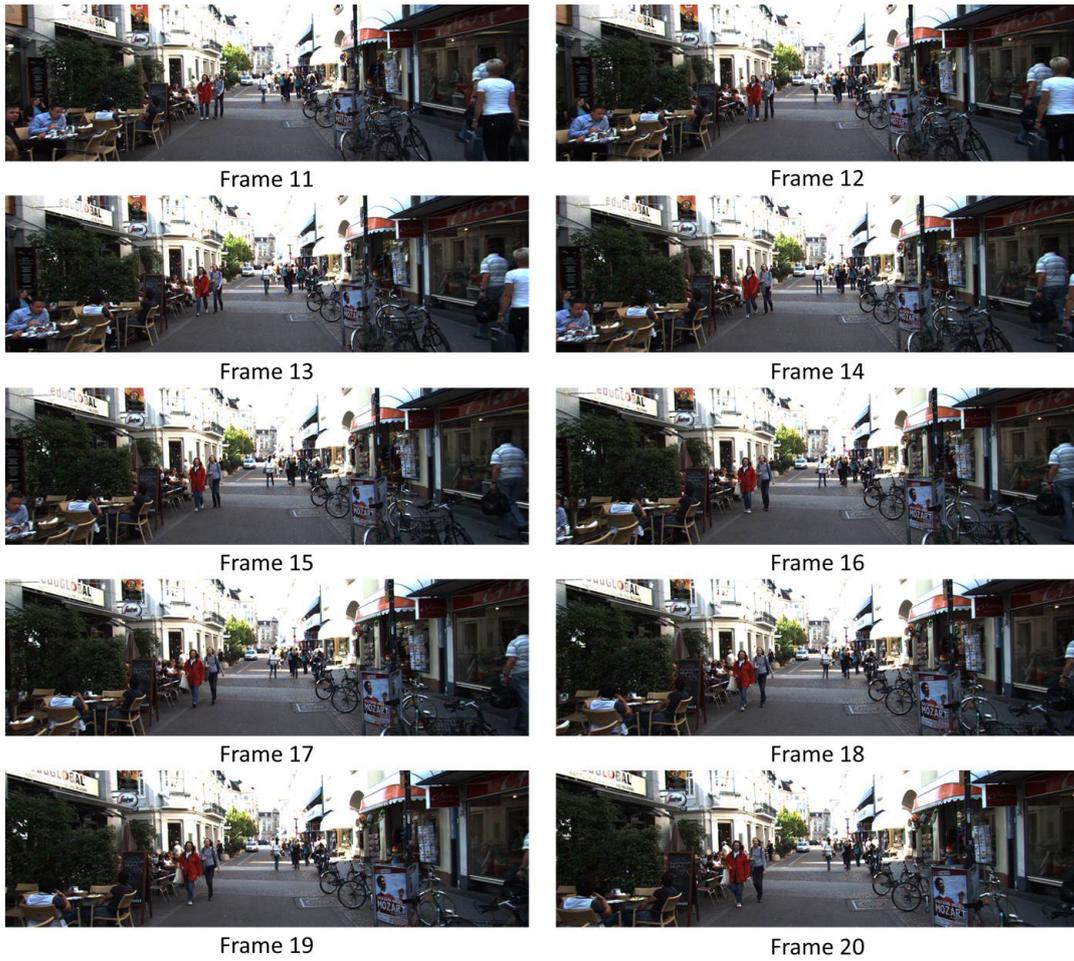


図 24 データセット D の入力画像 (Frame 11~Frame 20)

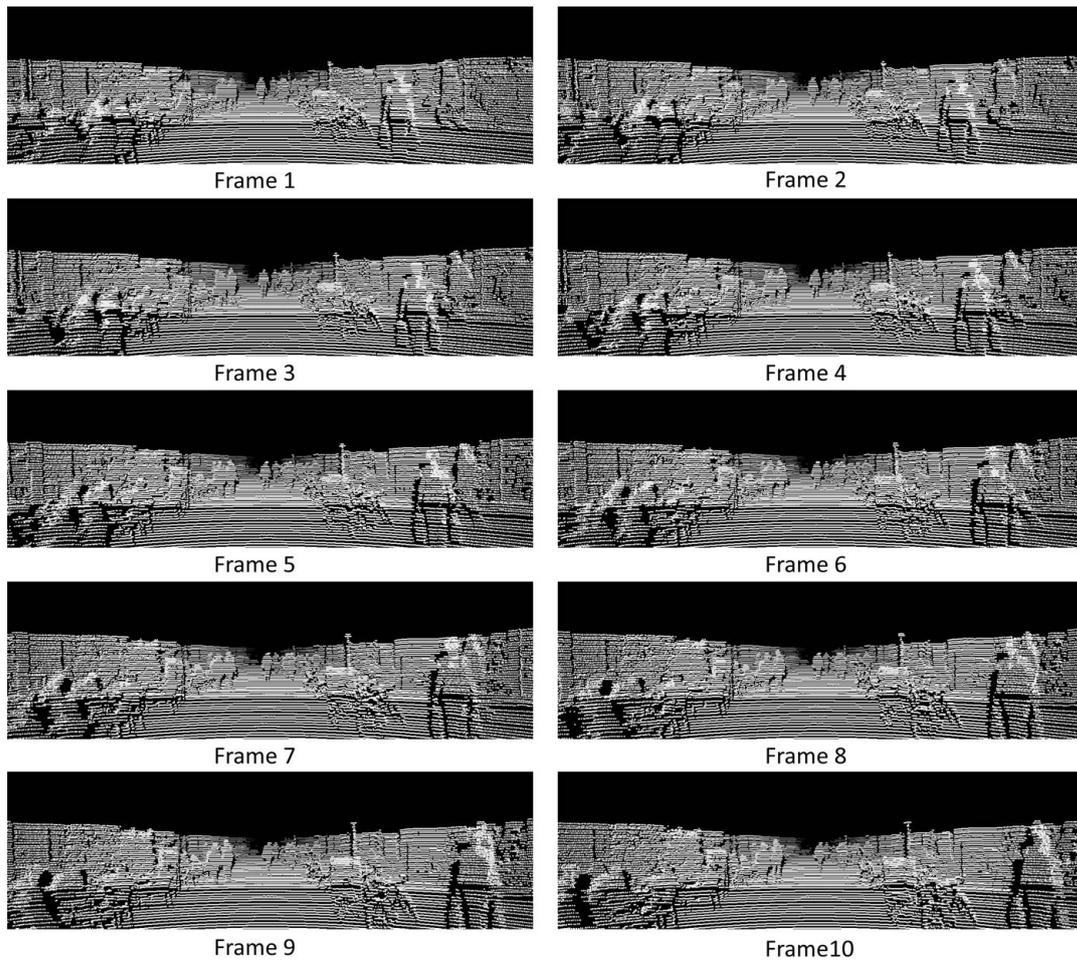


図 25 データセット D の三次元点群 (Frame 1~Frame 10)

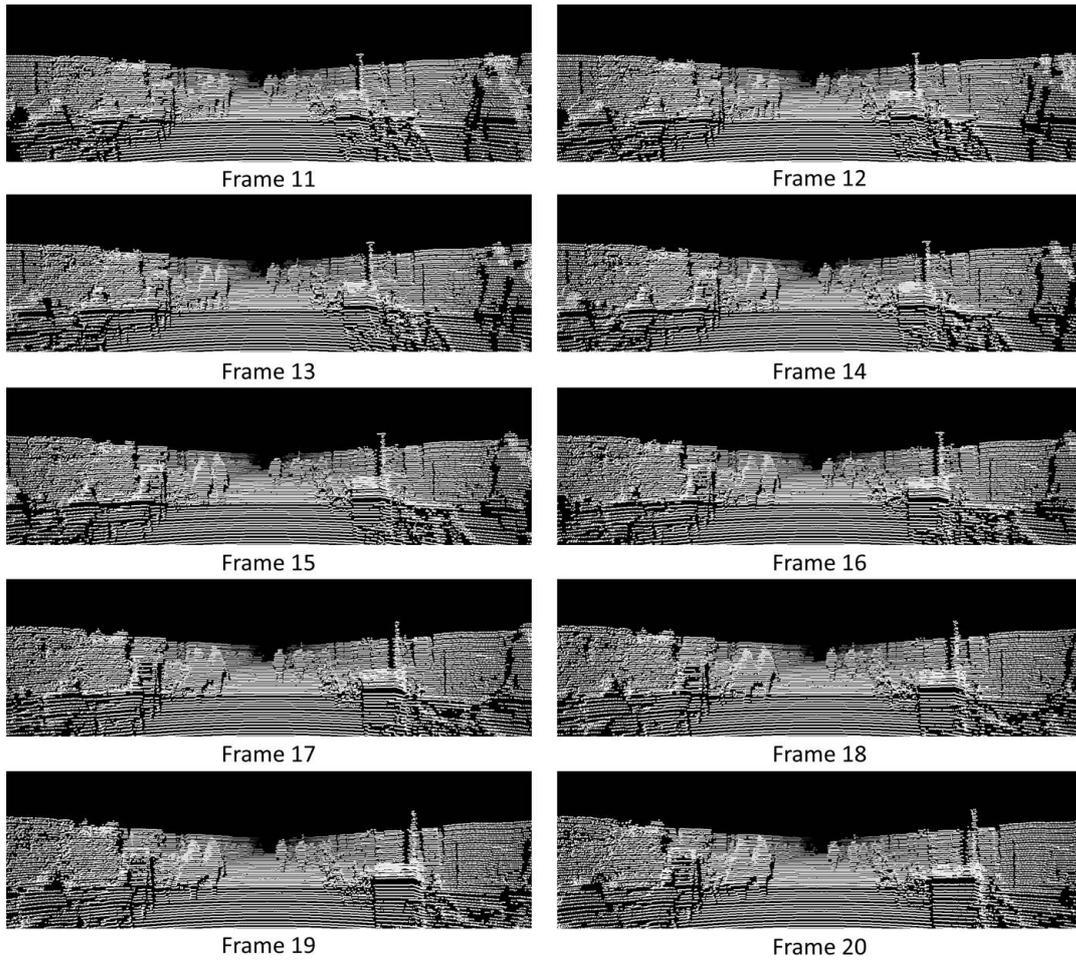


図 26 データセット D の三次元点群 (Frame 11~Frame 20)

## 4.2 移動物体らしさ関数の決定

移動物体領域および静止物体領域それぞれの輝度誤差および奥行誤差のヒストグラムとそれに基づき計算された輝度情報および奥行情報に基づく移動物体らしさを、それぞれ図 27, 図 28 に示す. 輝度誤差のヒストグラムは, 輝度値 (0~255) の階級を 1 ごとに, 奥行誤差のヒストグラムは, 階級を 0.1m ごとに分けた上で, それぞれの階級の度数から移動物体と静止物体の比率を計算した. また, これを階級軸に対してガウシアンフィルタをかけることで平滑化し, 輝度情報・奥行情報に基づく移動物体らしさ関数  $F_P$ ,  $F_D$  をそれぞれ作成した. なお, 輝度誤差のヒストグラムにおいて, 80 以上は移動物体らしさを 0.9, 奥行誤差のヒストグラムにおいて, 8m 以上は移動物体らしさを 0.9 とした.

## 4.3 移動物体領域の検出実験結果と定量的評価

本実験では, データセット B の連続する 10 フレームのデータを用いて, まず提案手法による輝度情報と奥行情報の整合性を確認する. 次に, 前節で決定した移動物体らしさ関数を用いて, データセット B, C, D の連続する 10 フレームにおける三次元点群中の移動物体領域の検出を行い, 検出率の定量的評価を行う.

### 4.3.1 輝度情報と奥行情報の整合性の確認

図 15 に示すように, データセット B では, 移動する複数の車両が存在する. このデータに対して, まず輝度情報と奥行情報の整合性を確認する. 図 17 に示すデータセット B の Frame 1 の点群について, 輝度誤差および奥行誤差の最大値を 1 に正規化することで画像化した結果を図 29 に示す. 図 29 より, 移動物体領域で双方の誤差が大きくなっていることが確認できる. また, 遠方の地面上の点や, 細い柱状物体などの静止物体上の点でも誤差が生じているのが確認できる.

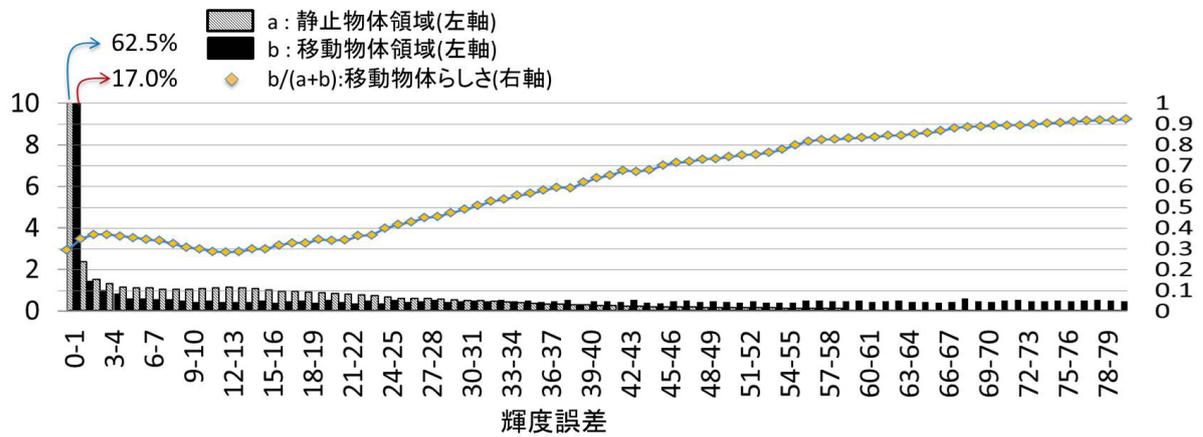


図 27 輝度誤差のヒストグラムと輝度情報に基づく移動物体らしさ

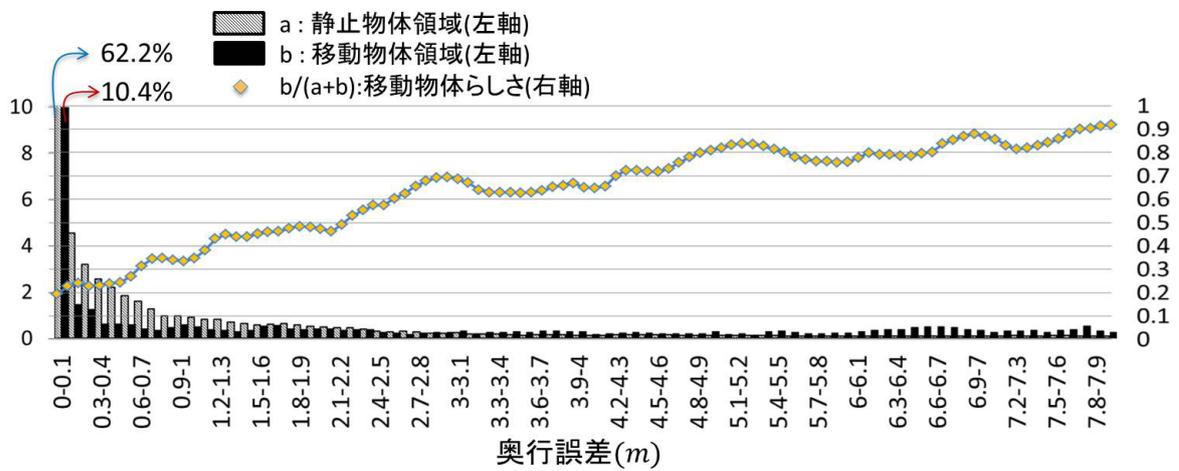
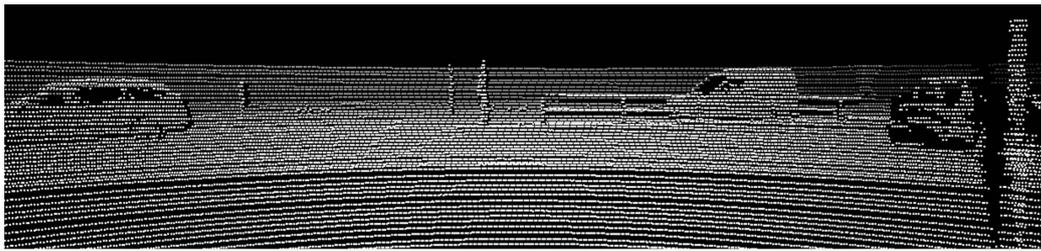
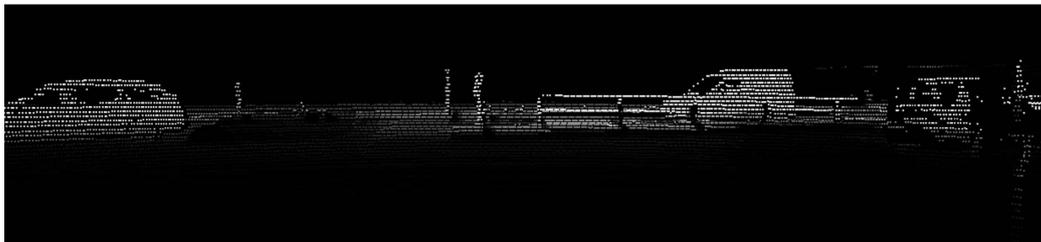


図 28 奥行誤差のヒストグラムと奥行情報に基づく移動物体らしさ



(a)データセットBの三次元点群の奥行値



(b)奥行情報の誤差マップ



(c)輝度情報の誤差マップ

図 29 奥行情報と輝度情報に基づく整合性 (白：誤差大，黒：誤差小)

#### 4.3.2 移動物体検出精度の定量的評価

本項では、エネルギー関数  $E$  を用いたエネルギー最小化に基づく移動物体の検出を行い、手動でラベルを付した真値と比較することで定量的評価を行う。具体的には、真値で移動物体領域に属する点の判定結果が移動物体領域であった点の数を TP(True Positive)、真値で移動物体領域に属する点の判定結果が静止物体領域であった点の数を FN(False Negative) とし、以下の TPR(True Positive Rate) を移動物体領域の検出率とする。

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (8)$$

また、真値で静止物体領域に属する点の判定結果が移動物体領域であった点の数を FP(False Positive) とし、真値で静止物体領域に属する点の判定結果が静止物体領域であった点の数を TN(True Negative) とし、以下の ACC(Accuracy) を識

別精度とする。

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

図 30～35 および表 2, 3, 4 にデータセット B に対する提案手法による移動物体検出, エネルギー関数に輝度情報のみを用いた移動物体検出, エネルギー関数に奥行情報のみを用いた移動物体検出の結果とその定量的評価値をそれぞれ示す。式 (1) で用いるパラメータ  $\kappa$  は, 予備実験からそれぞれ 12, 0.4, 1.6 とし, 提案手法による移動物体検出において, 式 (2) で用いるパラメータ  $\alpha = 2.5$  とした。表 2, 3, 4 の比較結果を図 36 に示す。また, 図 36 において, 有意水準 5% で有意差の認められたものを \* で示してる。

図 37～42 および表 5, 6, 7 にデータセット C に対する提案手法による移動物体検出, エネルギー関数に輝度情報のみを用いた移動物体検出, エネルギー関数に奥行情報のみを用いた移動物体検出の結果とその定量的評価値をそれぞれ示す。式 (1) で用いるパラメータ  $\kappa$  は, 予備実験からそれぞれ 12, 0.4, 1.6 とし, 提案手法による移動物体検出において, 式 (2) で用いるパラメータ  $\alpha = 2.5$  とした。表 5, 6, 7 の比較結果を図 43 に示す。また, 図 43 において, 有意水準 5% で有意差の認められたものを \* で示してる。

図 44～49 および表 8, 9, 10 にデータセット D に対する提案手法による移動物体検出, エネルギー関数に輝度情報のみを用いた移動物体検出, エネルギー関数に奥行情報のみを用いた移動物体検出の結果とその定量的評価値をそれぞれ示す。式 (1) で用いるパラメータ  $\kappa$  は, 予備実験からそれぞれ 12, 0.4, 0.2 とし, 提案手法による移動物体検出において, 式 (2) で用いるパラメータ  $\alpha = 2.5$  とした。表 8, 9, 10 の比較結果を図 50 に示す。また, 図 50 において, 有意水準 5% で有意差の認められたものを \* で示してる。

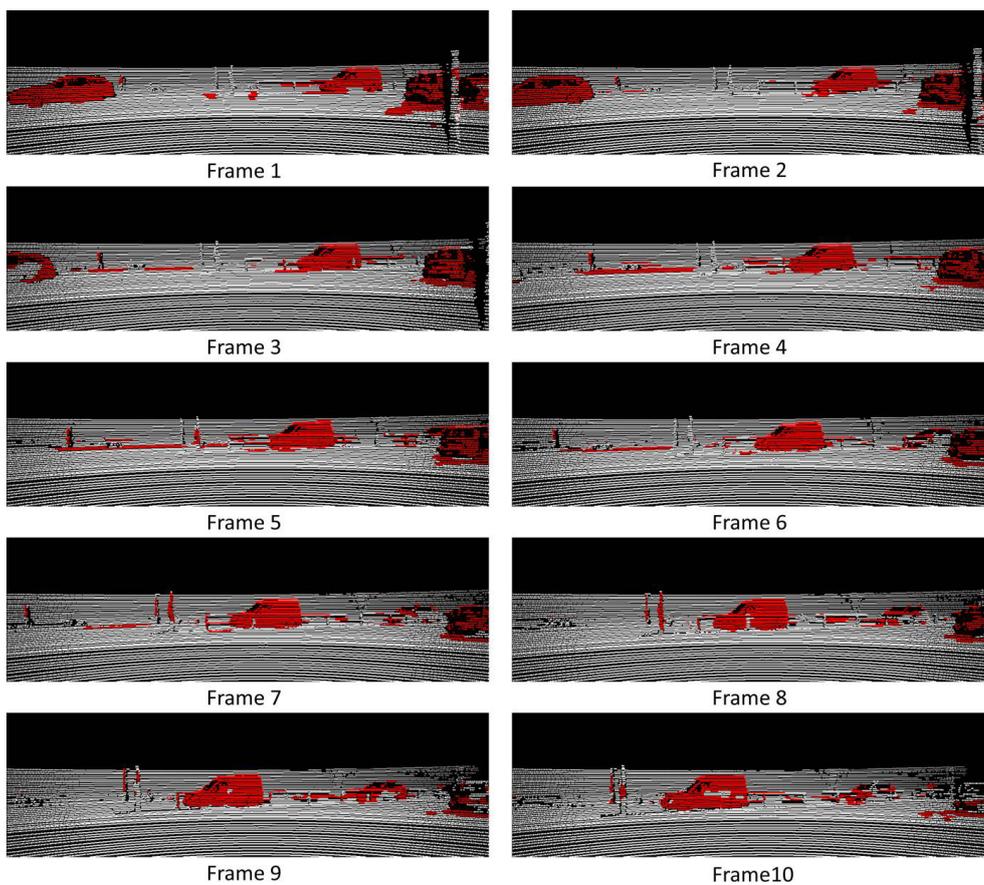


図 30 データセット B に対する移動物体領域検出結果 (Frame 1~10)

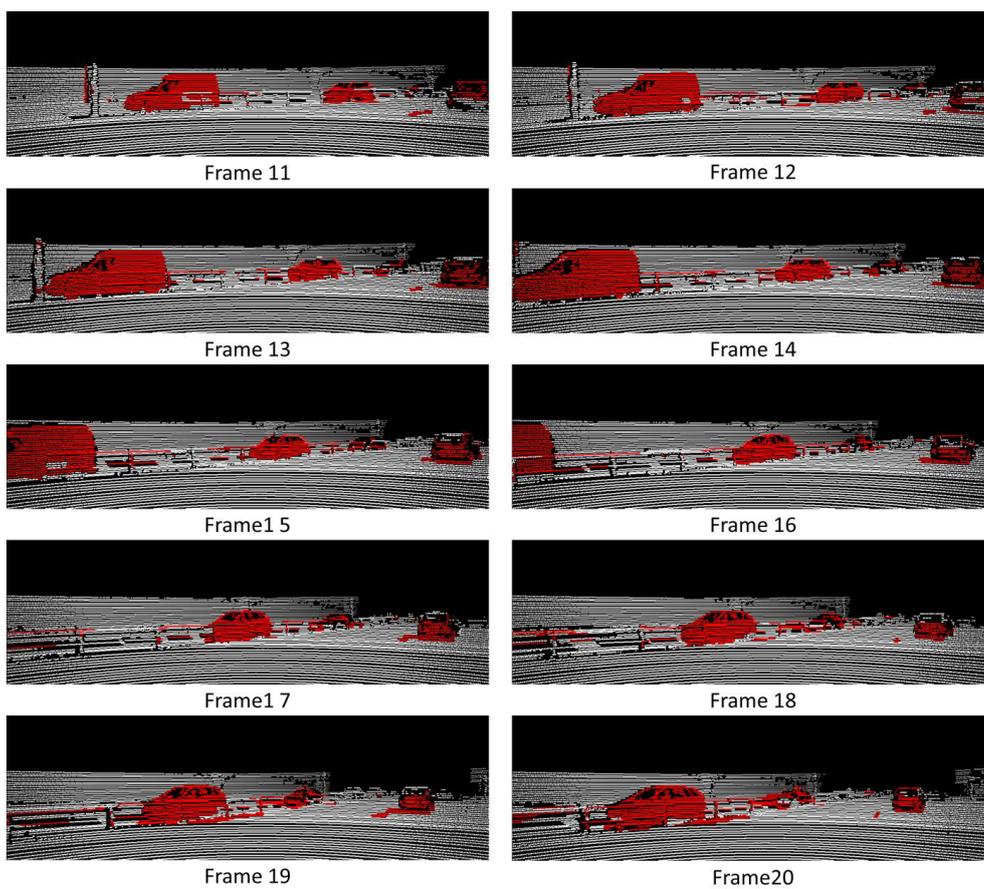


図 31 データセット B に対する移動物体領域検出結果 (Frame 11~20)

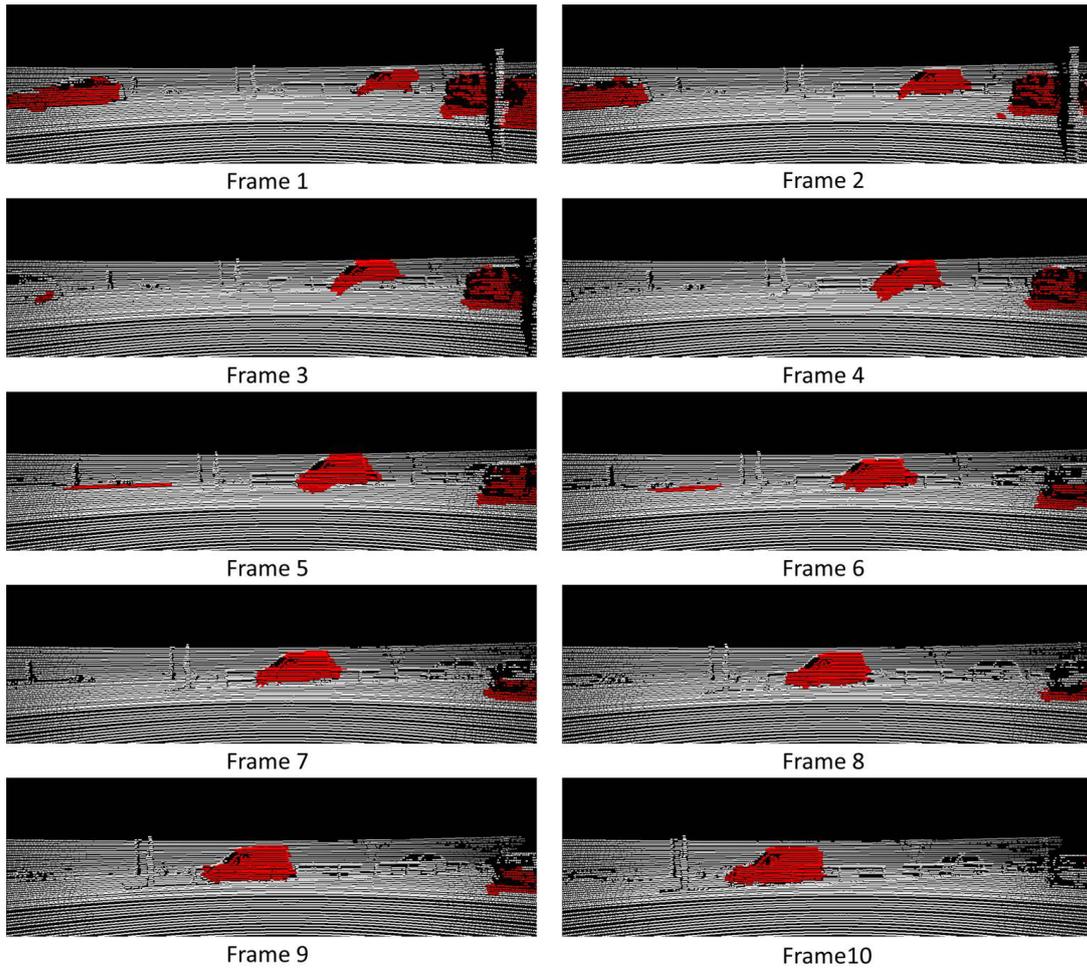


図 32 データセット B に対する輝度情報のみを用いた検出結果 (Frame 1~10)

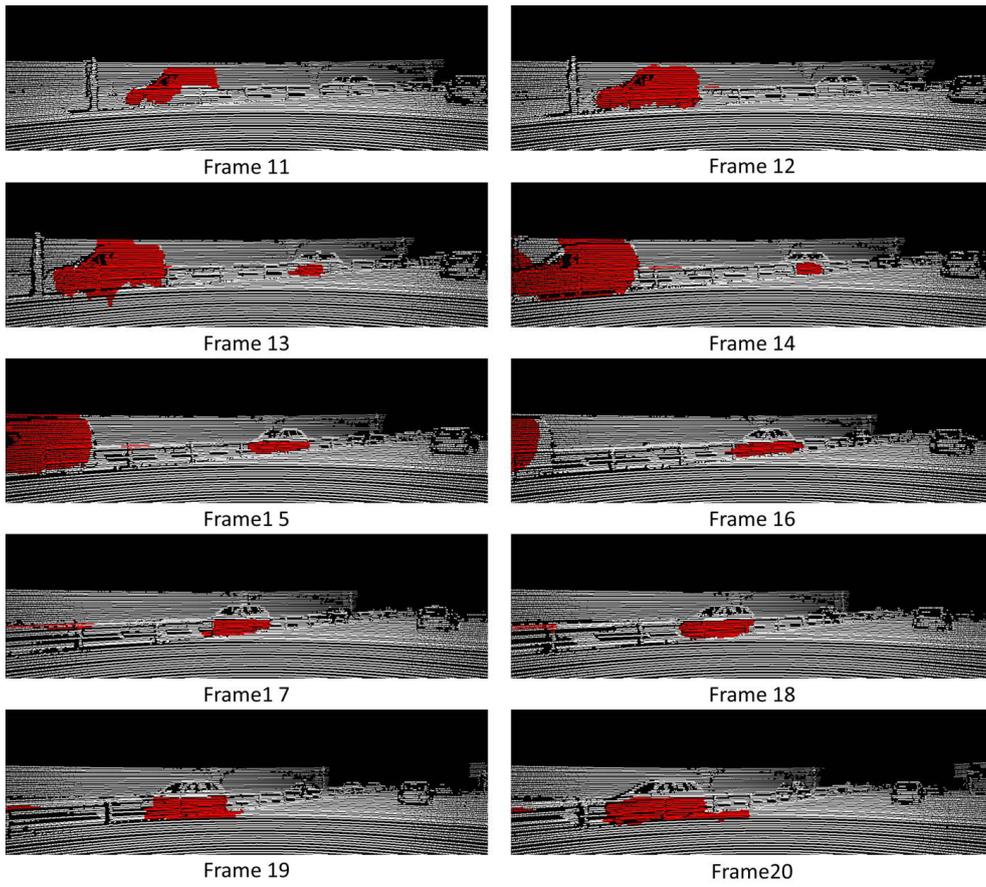


図 33 データセット B に対する輝度情報のみを用いた検出結果 (Frame 11~20)

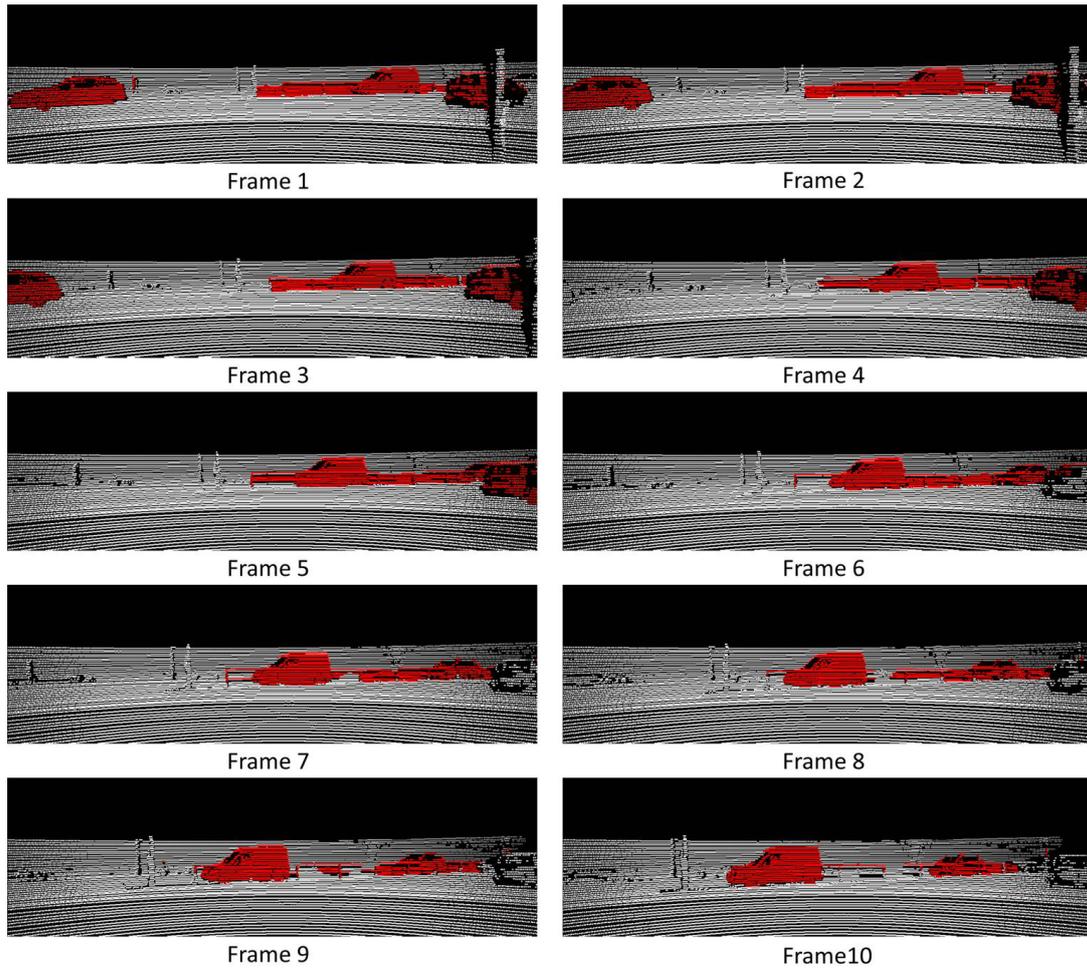


図 34 データセット B に対する奥行情報のみを用いた検出結果 (Frame 1~10)

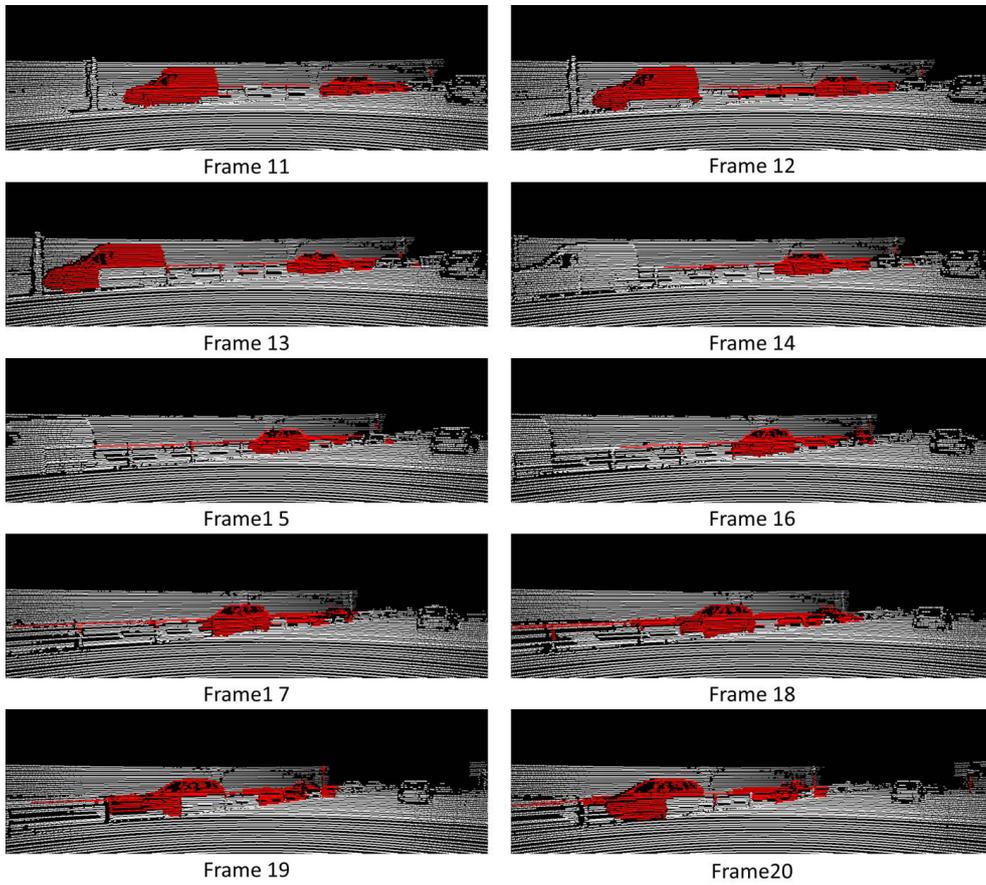


図 35 データセット B に対する奥行情報のみを用いた検出結果 (Frame 11~20)

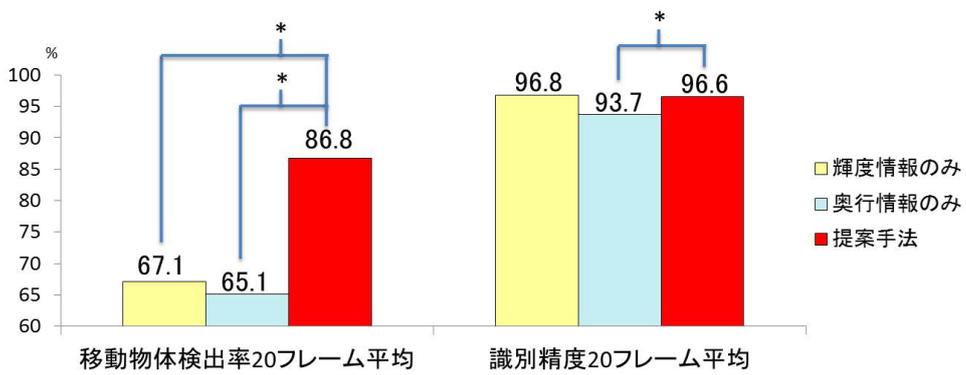


図 36 データセット B における 20 フレーム平均の比較グラフ

表 2 データセット B に対する移動物体領域検出結果の定量的評価値

Frame	1	2	3	4	5	6	7	8	9	10
TPR	0.897	0.889	0.788	0.898	0.885	0.847	0.818	0.855	0.830	0.753
ACC	0.966	0.979	0.964	0.956	0.956	0.966	0.965	0.972	0.968	0.953
11	12	13	14	15	16	17	18	19	20	Average
0.855	0.876	0.952	0.917	0.896	0.935	0.848	0.870	0.903	0.845	0.868
0.978	0.968	0.977	0.968	0.967	0.977	0.972	0.962	0.953	0.947	0.966

表 3 データセット B に対する輝度情報のみを用いた検出結果の定量的評価値

Frame	1	2	3	4	5	6	7	8	9	10
TPR	0.834	0.849	0.650	0.886	0.881	0.774	0.720	0.741	0.829	0.659
ACC	0.972	0.984	0.977	0.989	0.975	0.974	0.978	0.982	0.984	0.972
11	12	13	14	15	16	17	18	19	20	Average
0.504	0.696	0.752	0.746	0.711	0.500	0.334	0.433	0.437	0.479	0.671
0.956	0.966	0.959	0.950	0.957	0.965	0.962	0.965	0.947	0.941	0.968

表 4 データセット B に対する奥行情報のみを用いた検出結果の定量的評価値

Frame	1	2	3	4	5	6	7	8	9	10
TPR	0.874	0.850	0.849	0.848	0.887	0.706	0.735	0.779	0.799	0.771
ACC	0.944	0.939	0.950	0.958	0.959	0.956	0.954	0.961	0.963	0.969
11	12	13	14	15	16	17	18	19	20	Average
0.778	0.713	0.649	0.092	0.137	0.327	0.590	0.604	0.431	0.607	0.651
0.964	0.931	0.932	0.853	0.876	0.928	0.946	0.920	0.907	0.929	0.937

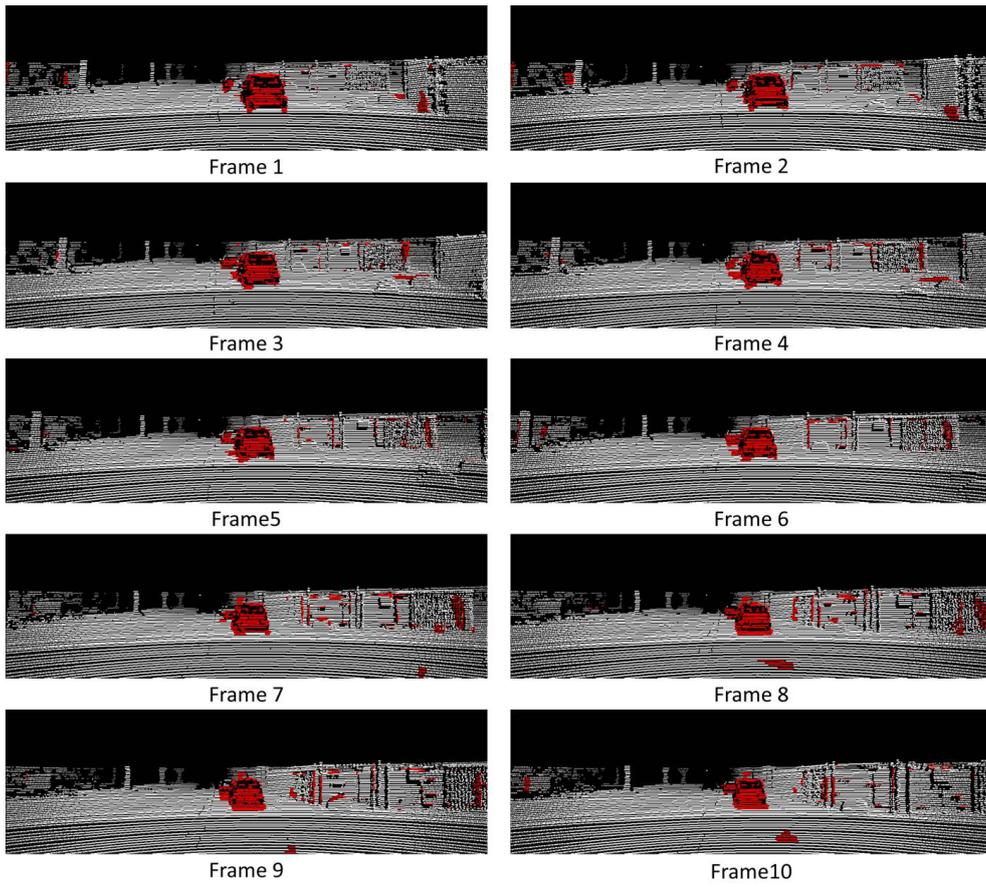


図 37 データセット C に対する移動物体領域検出結果 (Frame 1~10)

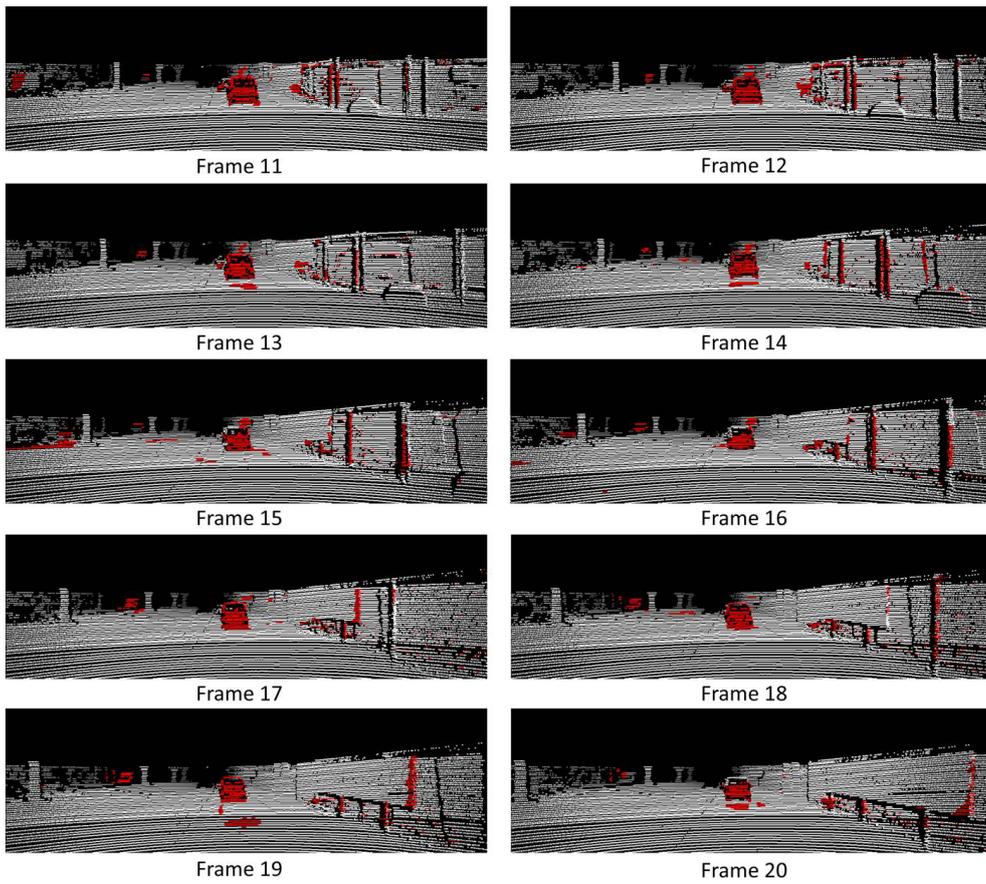


図 38 データセット C に対する移動物体領域検出結果 (Frame 11~20)

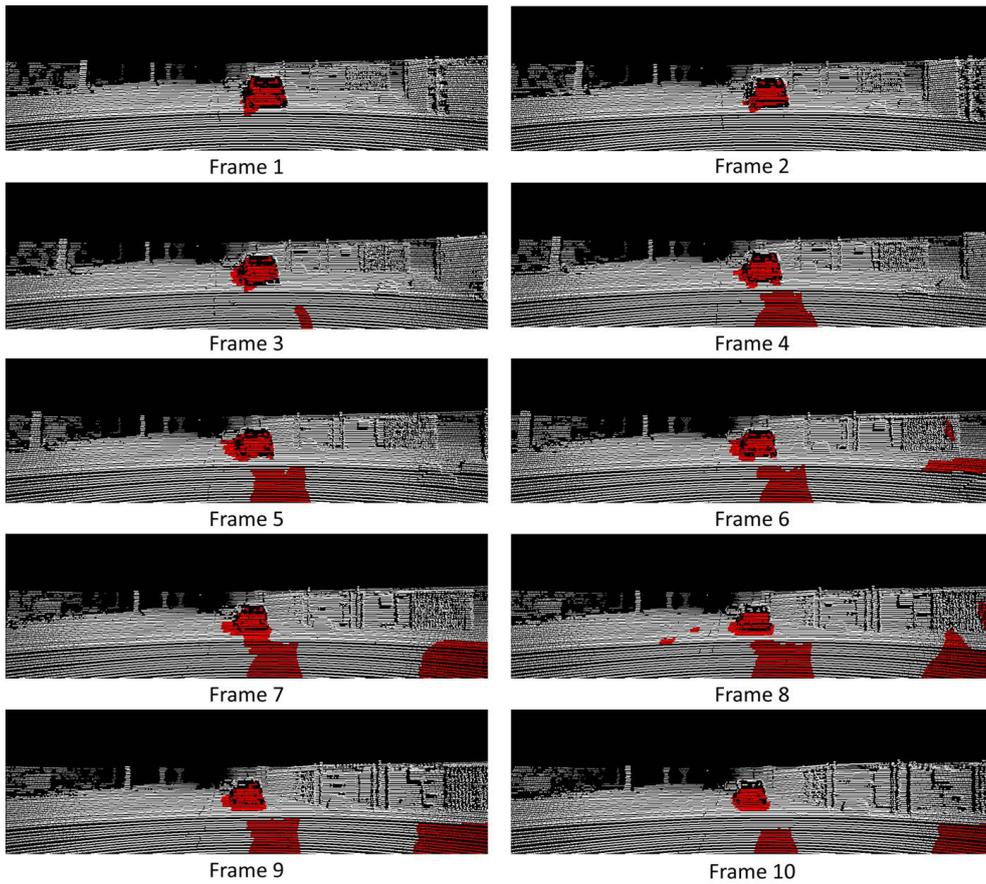


図 39 データセット C に対する輝度情報のみを用いた検出結果 (Frame 1~10)

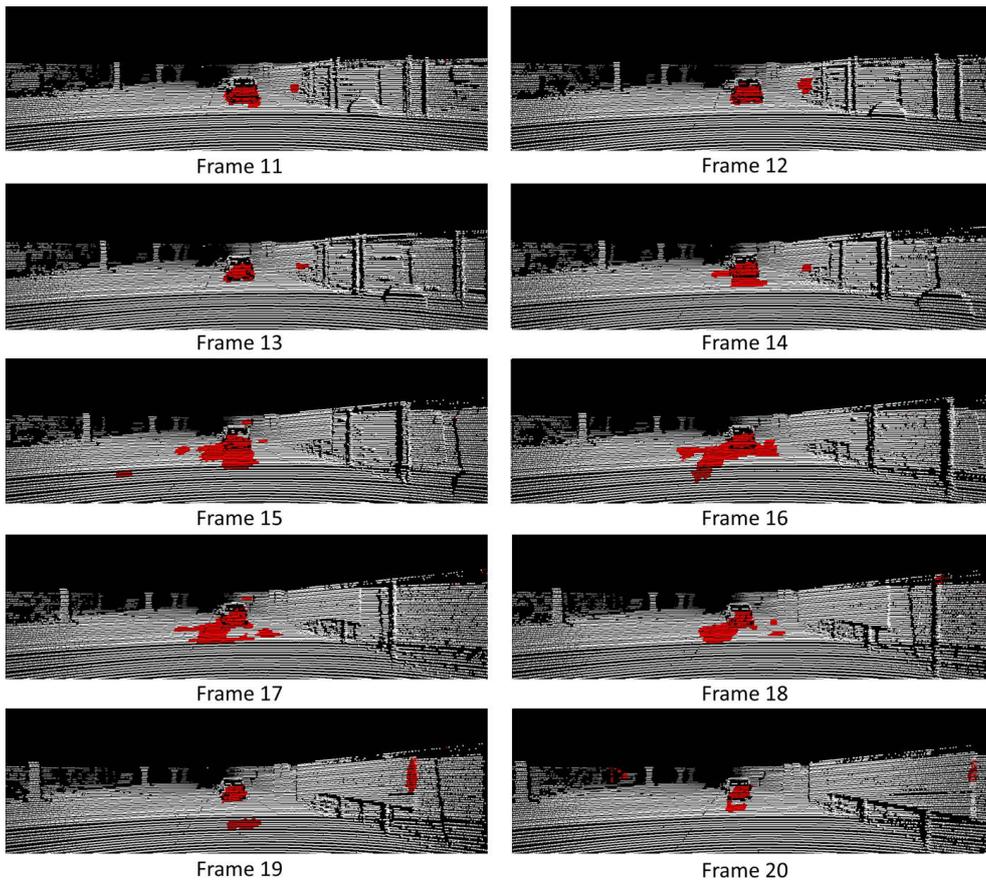


図 40 データセット C に対する輝度情報のみを用いた検出結果 (Frame 11~20)

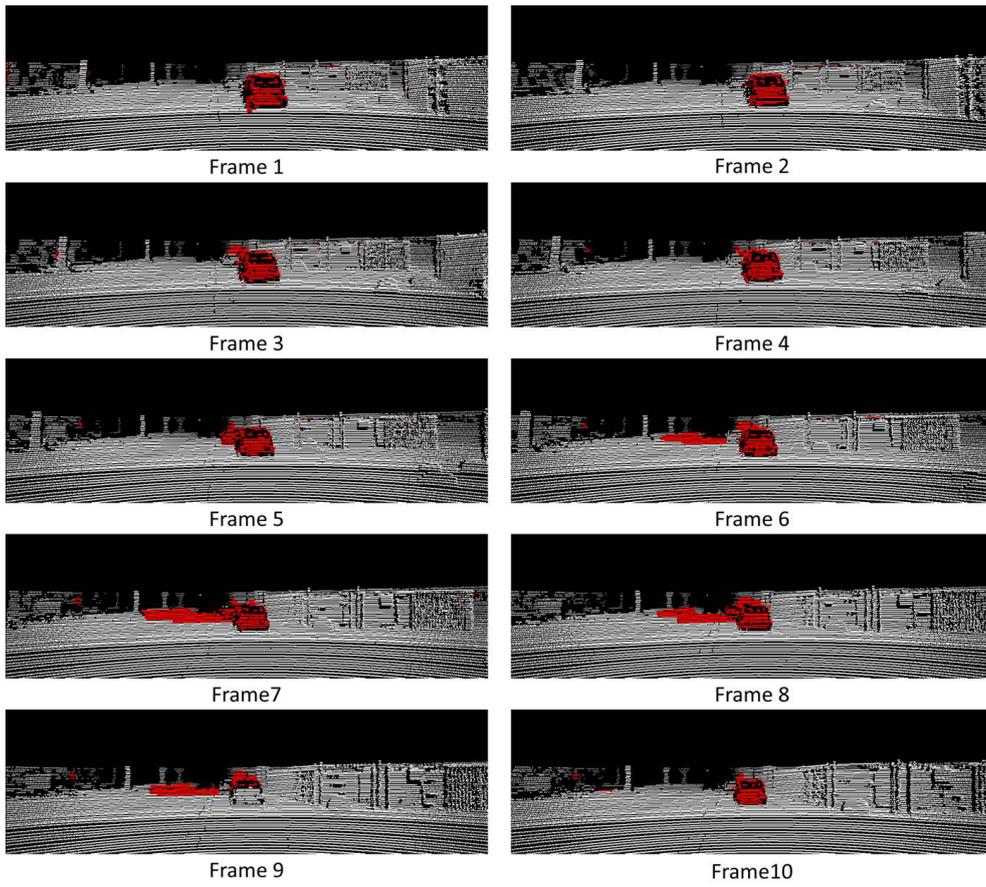


図 41 データセット C に対する奥行情報のみを用いた検出結果 (Frame 1~10)

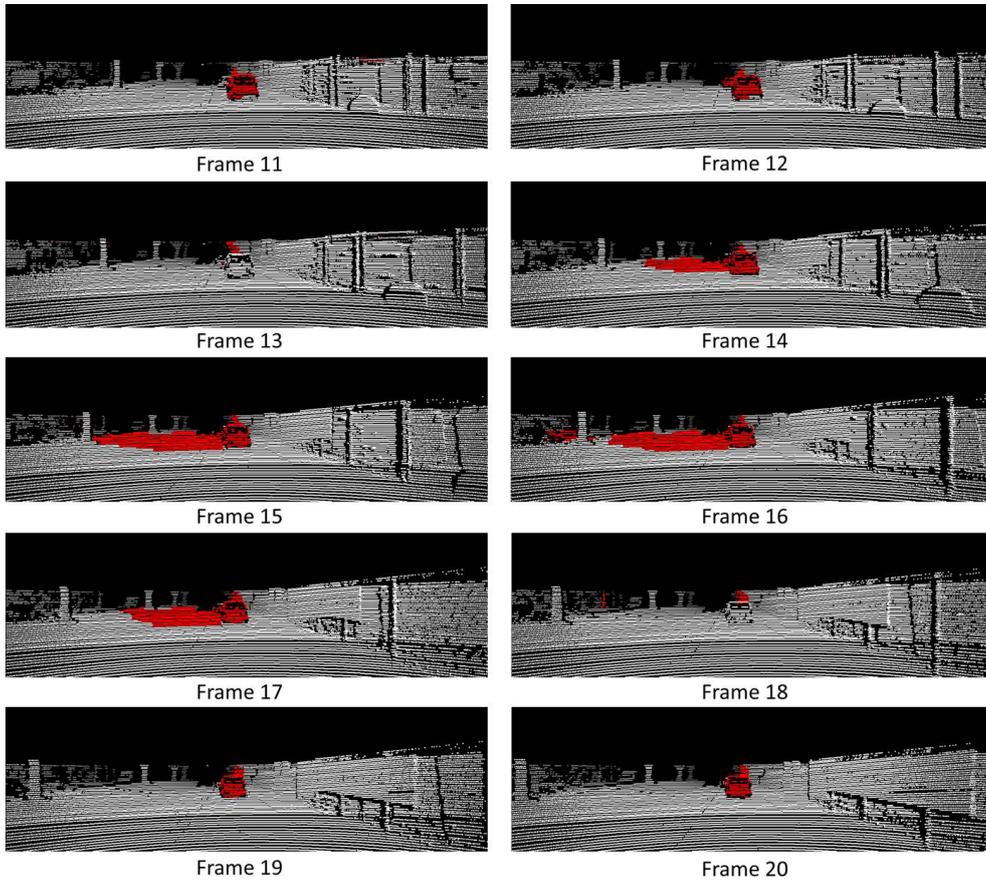


図 42 データセット C に対する奥行情報のみを用いた検出結果 (Frame 11~20)

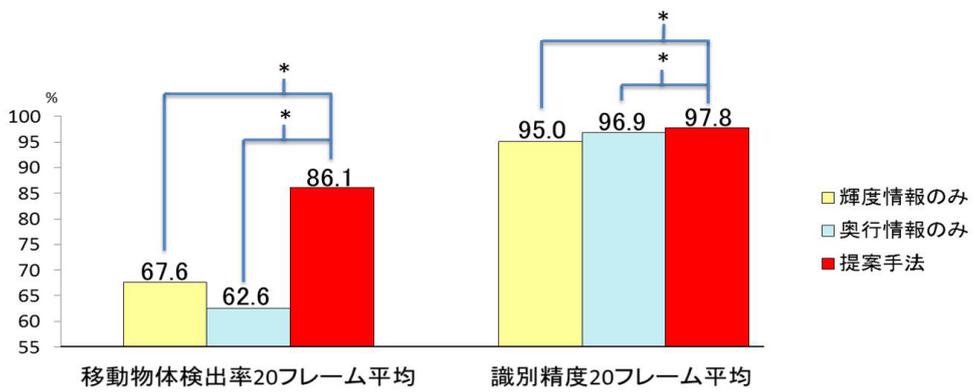


図 43 データセット C における 20 フレーム平均の比較グラフ

表 5 データセット C に対する移動物体領域検出結果の定量的評価値

Frame	1	2	3	4	5	6	7	8	9	10
TPR	0.839	0.876	0.854	0.817	0.819	0.821	0.868	0.904	0.939	0.782
ACC	0.972	0.980	0.977	0.974	0.978	0.978	0.977	0.977	0.975	0.972
11	12	13	14	15	16	17	18	19	20	Average
0.813	0.837	0.885	0.830	0.837	0.897	0.873	0.873	0.921	0.926	0.861
0.982	0.980	0.970	0.978	0.987	0.986	0.984	0.978	0.980	0.981	0.978

表 6 データセット C に対する輝度情報のみを用いた検出結果の定量的評価値

Frame	1	2	3	4	5	6	7	8	9	10
TPR	0.650	0.660	0.585	0.722	0.674	0.749	0.687	0.651	0.555	0.544
ACC	0.989	0.986	0.989	0.962	0.929	0.913	0.913	0.949	0.974	0.984
11	12	13	14	15	16	17	18	19	20	Average
0.618	0.594	0.683	0.678	0.739	0.723	0.759	0.721	0.789	0.730	0.676
0.984	0.985	0.981	0.945	0.941	0.925	0.903	0.894	0.917	0.947	0.950

表 7 データセット C に対する奥行情報のみを用いた検出結果の定量的評価値

Frame	1	2	3	4	5	6	7	8	9	10
TPR	0.614	0.701	0.020	0.739	0.756	0.749	0.702	0.012	0.634	0.684
ACC	0.989	0.993	0.977	0.958	0.921	0.916	0.935	0.978	0.990	0.992
11	12	13	14	15	16	17	18	19	20	Average
0.673	0.646	0.708	0.697	0.812	0.676	0.763	0.828	0.139	0.709	0.626
0.985	0.986	0.985	0.987	0.992	0.963	0.954	0.963	0.947	0.990	0.969

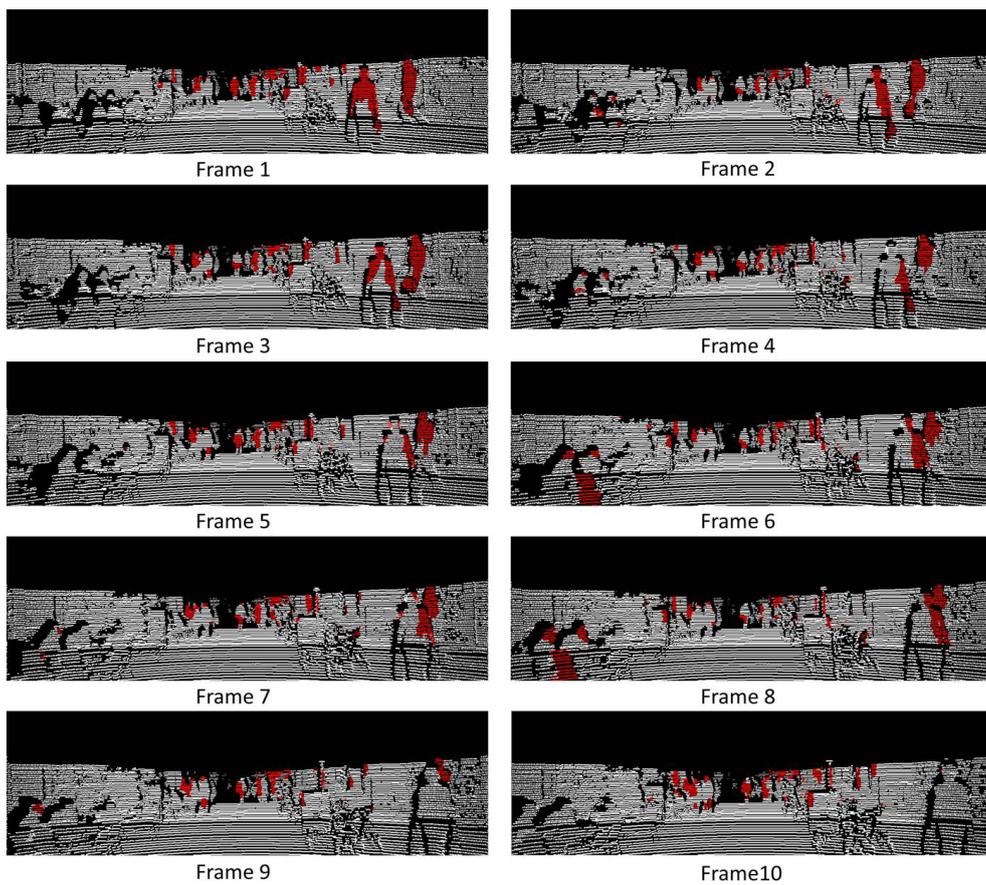


図 44 データセット D に対する移動物体領域検出結果 (Frame 1~10)

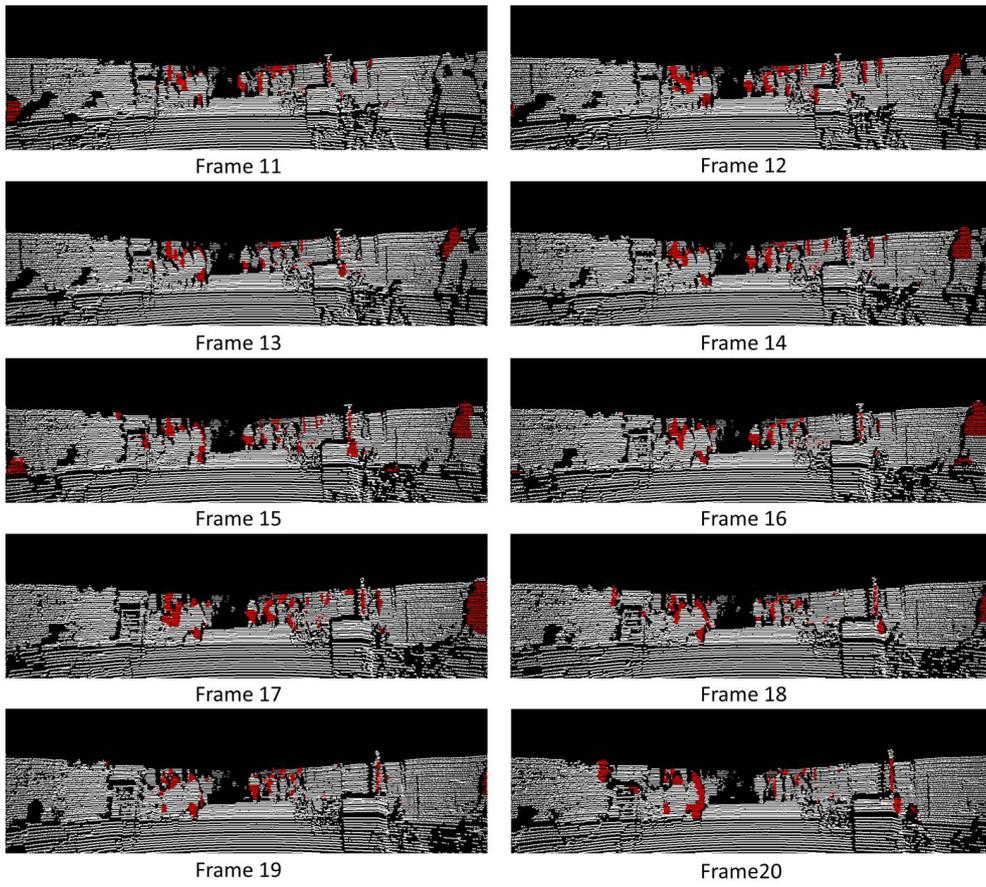


図 45 データセット D に対する移動物体領域検出結果 (Frame 11~20)

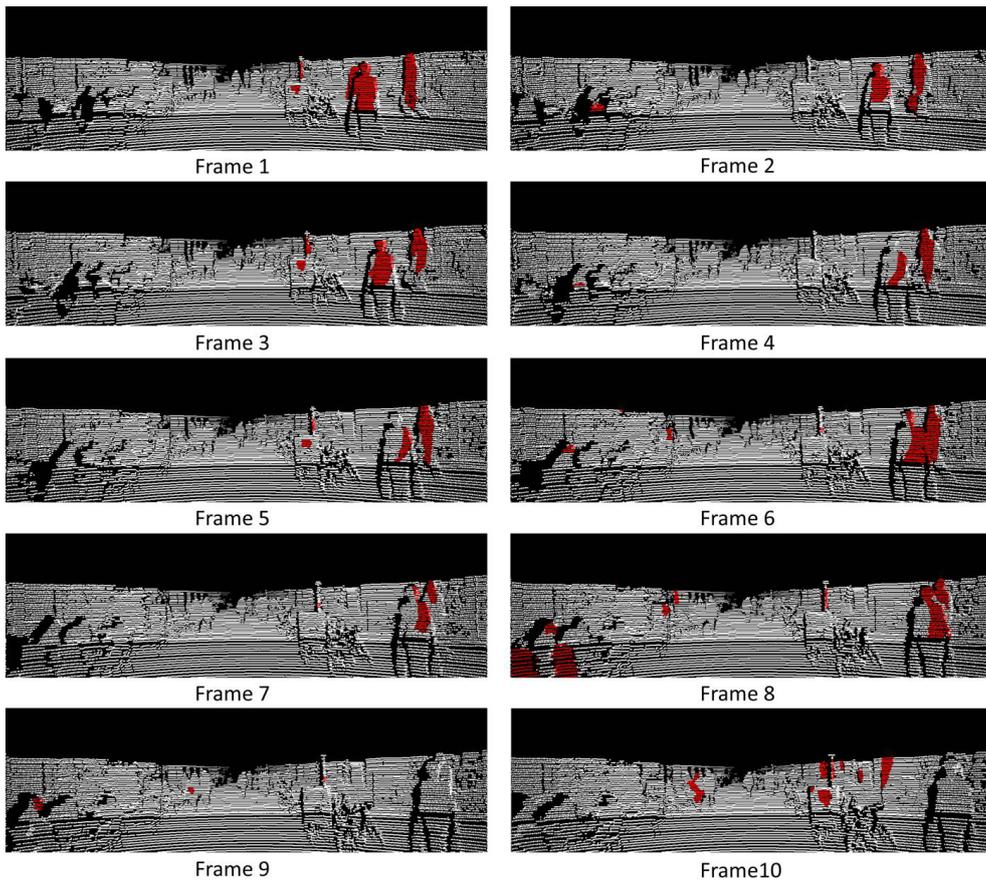


図 46 データセット D に対する輝度情報のみを用いた検出結果 (Frame 1~10)

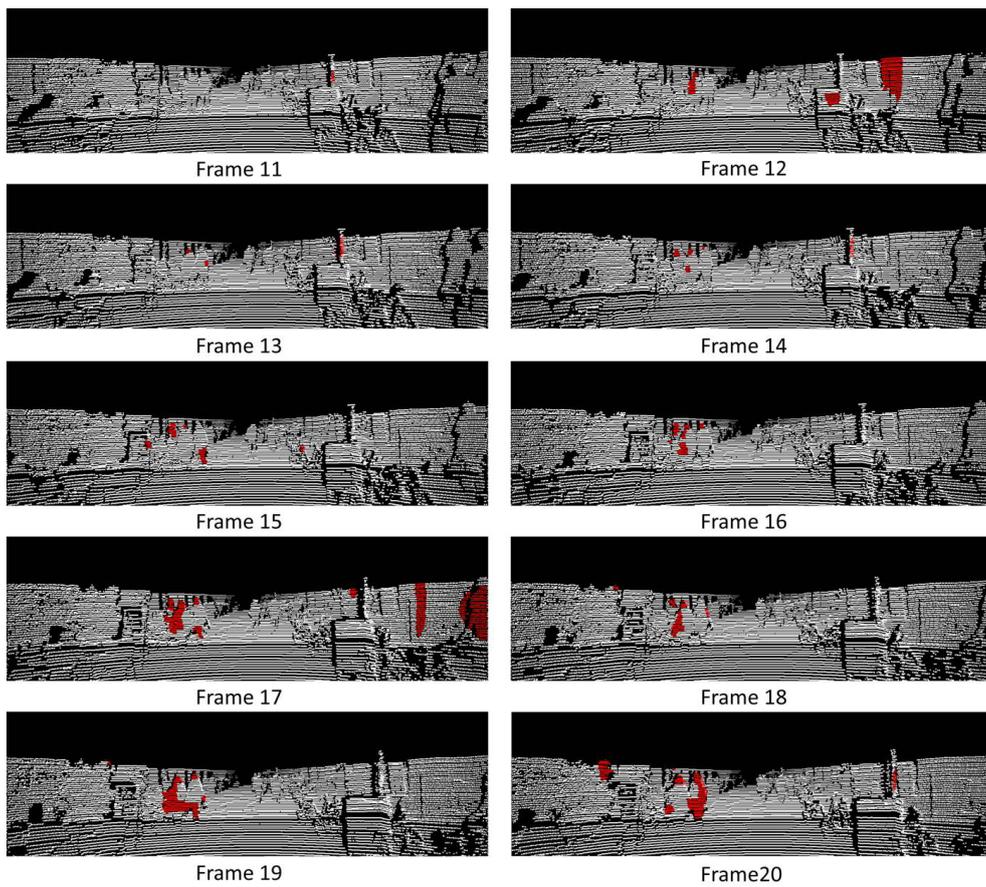


図 47 データセット D に対する輝度情報のみを用いた検出結果 (Frame 11~20)

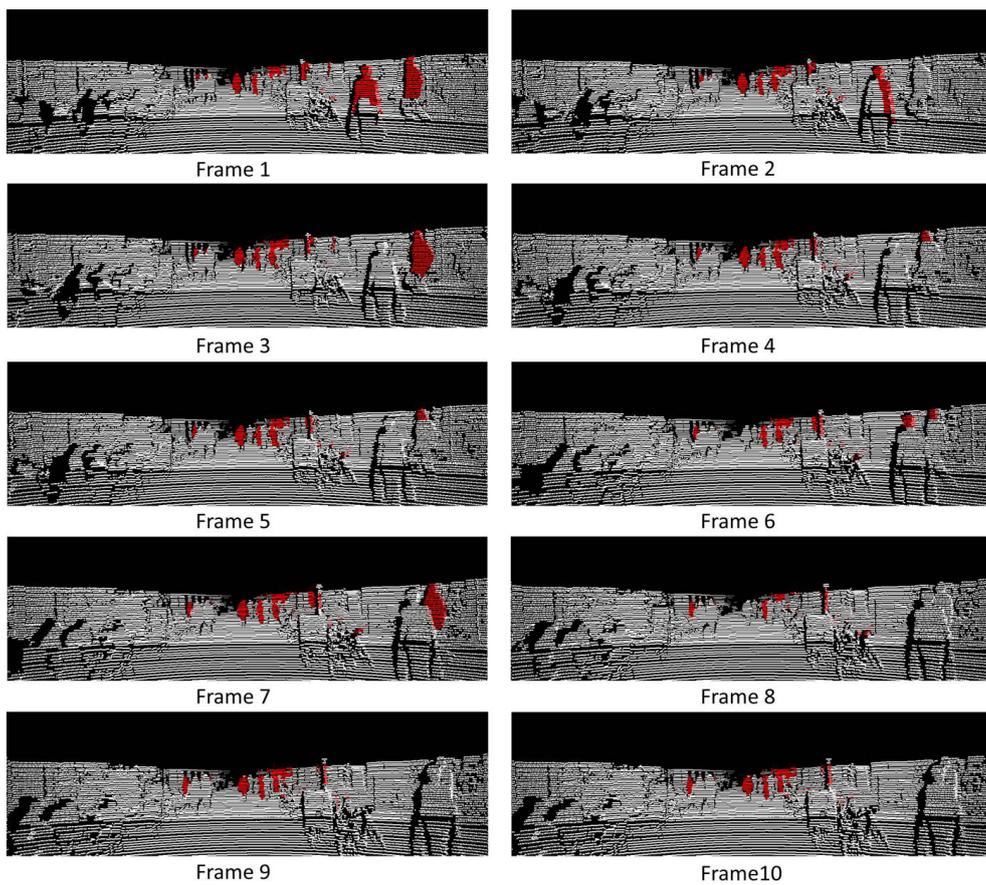


図 48 データセット D に対する奥行情報のみを用いた検出結果 (Frame 1~10)

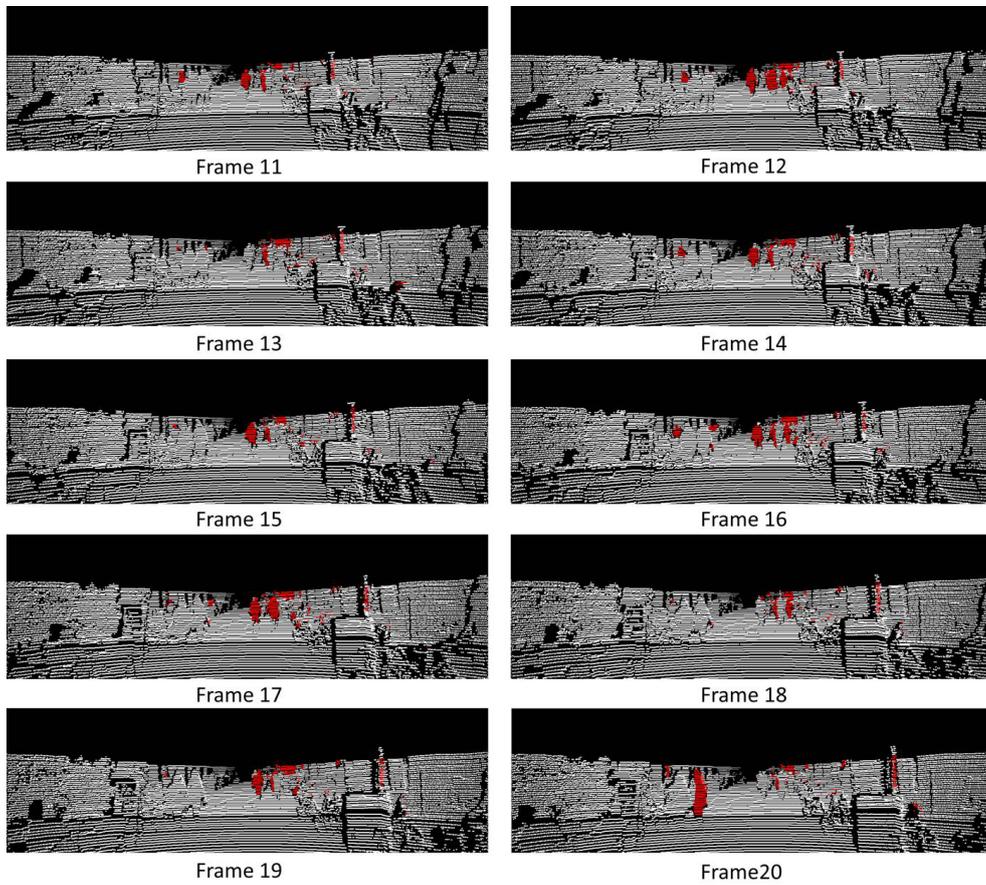


図 49 データセット D に対する奥行情報のみを用いた検出結果 (Frame 11~20)

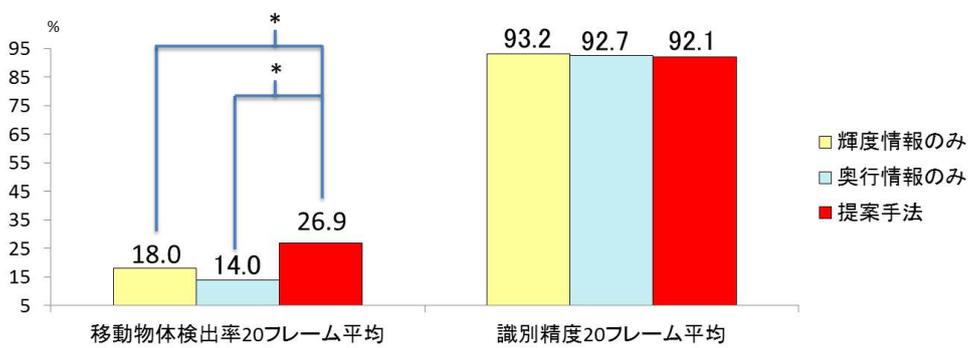


図 50 データセット D における 20 フレーム平均の比較グラフ

表 8 データセット D に対する移動物体領域検出結果の定量的評価値

Frame	1	2	3	4	5	6	7	8	9	10
TPR	0.506	0.394	0.458	0.249	0.265	0.289	0.297	0.247	0.165	0.160
ACC	0.943	0.943	0.941	0.920	0.915	0.902	0.923	0.908	0.922	0.911
11	12	13	14	15	16	17	18	19	20	Average
0.087	0.170	0.118	0.232	0.189	0.325	0.432	0.302	0.235	0.268	0.269
0.898	0.891	0.905	0.918	0.903	0.935	0.935	0.938	0.941	0.928	0.921

表 9 データセット D に対する輝度情報のみを用いた検出結果の定量的評価値

Frame	1	2	3	4	5	6	7	8	9	10
TPR	0.435	0.341	0.354	0.218	0.196	0.337	0.137	0.265	0.181	0.096
ACC	0.954	0.953	0.947	0.938	0.929	0.938	0.928	0.916	0.929	0.914
11	12	13	14	15	16	17	18	19	20	Average
0.000	0.046	0.002	0.026	0.026	0.074	0.363	0.120	0.299	0.247	0.180
0.915	0.887	0.916	0.926	0.923	0.931	0.937	0.944	0.958	0.944	0.932

表 10 データセット D に対する奥行情報のみを用いた検出結果の定量的評価値

Frame	1	2	3	4	5	6	7	8	9	10
TPR	0.399	0.159	0.270	0.112	0.109	0.085	0.233	0.052	0.141	0.167
ACC	0.946	0.935	0.940	0.923	0.918	0.912	0.925	0.923	0.932	0.931
11	12	13	14	15	16	17	18	19	20	Average
0.086	0.139	0.039	0.090	0.078	0.141	0.121	0.066	0.101	0.208	0.140
0.917	0.914	0.913	0.925	0.923	0.927	0.924	0.934	0.937	0.940	0.927

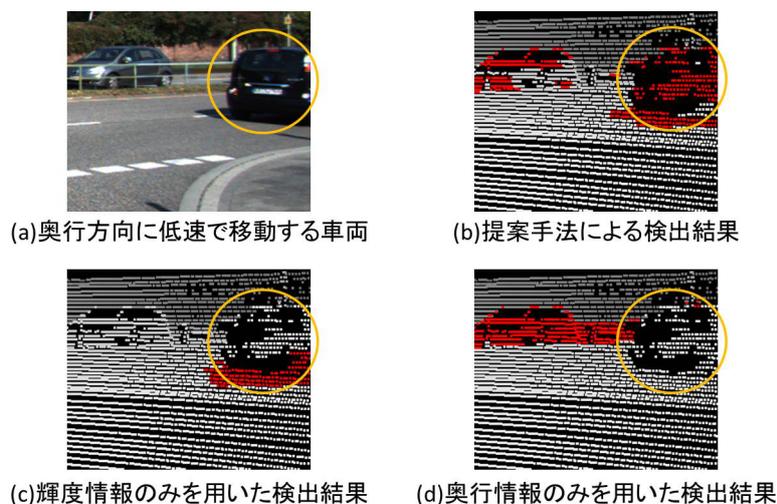


図 51 奥行方向に低速移動する車両の検出結果 (データセット B Frame8)

#### 4.4 考察

データセット B について、図 30 より、提案手法による移動物体領域の検出結果には、遠方の地面上の点や細い柱状物体上の点など、静止物体領域に属する点であるにもかかわらず移動物体領域と判定されている点が存在する。図 29 中においても、輝度誤差や奥行誤差が生じていることが確認できる。この原因としては、GPS・IMU により得られる位置・姿勢情報の精度が十分でなく、静止物体上の点が誤って異なる物体上に対応付けられるためである。また、上記の問題がない場合でも、日照条件の変化や、鏡面反射、カメラの自動露出補正機能などに起因して、同一物体の輝度値の誤差が大きくなる場合がある。このため、入力となる画像の輝度値を直接扱うのではなく、投影点の周辺のテクスチャなども用いるなど、輝度変化への対応を考える必要がある。また、エネルギー関数に輝度情報または奥行情報のみを用いた場合に対する比較実験では、図 36 のグラフに示すように、輝度情報のみを用いた場合と比較して、提案手法は移動物体の検出率が優れており有意水準 5% で有意差が認められた。また、奥行情報のみを用いた場合と比較して、提案手法は移動物体の検出率と識別精度が優れており有意水準 5% で有意差が認められた。これは、図 51 に示すように、カメラの奥行方向に低速で移動するような、輝度値や奥行値の変化が少ない移動物体上の点でも、提案手法では双方の情報を用いることで検出が可能なためと考えられる。識別精度に

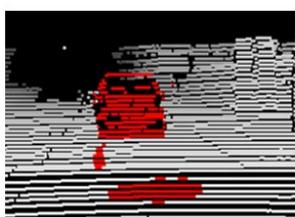
については輝度情報のみを用いた検出結果が優れているが、前述のように、GPS・IMUにより得られる位置・姿勢情報の精度が十分ではないことが原因として挙げられる。

データセット C は、カメラの奥行方向に移動する車両が存在する。図 37 より、提案手法による移動物体領域の検出結果において、移動物体が効果的に検出できていることがわかる。また、エネルギー関数に輝度情報または奥行情報のみを用いた場合に対する比較実験では、図 43 のグラフに示すように、提案手法は移動物体の検出率と識別精度が優れており有意水準 5% で有意差が認められた。これは、移動物体がカメラの奥行方向に移動するために、輝度情報のみや奥行情報のみの場合、移動物体の背景が遮蔽され続けることにより検出が困難となるためと考えられる。識別精度についても、奥行情報や輝度情報のみを用いた場合、本来分割されるべき領域で輝度値や奥行値が類似しているために誤検出が増加し、識別精度が低下しているものと考えられる。また、提案手法による移動物体領域検出結果と、輝度情報のみを用いた検出結果の両方で、図 52 に示すように、地面上の点が誤検出されている。これは、図 53 に示すように、画像左側の白い壁が近づくにつれて、反射光の影響で画像全体が明るくなり、対象フレームとその前後フレームで輝度値の差が生じたためであると考えられる。

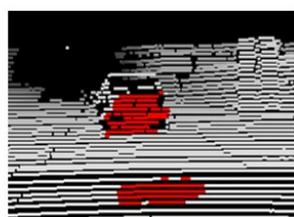
データセット D は、低速で移動する歩行者が存在する。また、カメラの移動も低速であるため、移動物体の背景が遮蔽され続ける。そのため、図 44 より、カメラと移動物体の移動の双方が低速である場合は、提案手法においても移動物体領域の検出が困難であることがわかった。また、データセット B, C で検出できた、カメラの奥行方向に移動する物体は、移動物体の移動速度は低速だがカメラの移動速度が速いことや、移動物体とカメラの双方の移動が高速であったことなどに起因して検出できたと考えられる。また、エネルギー関数に輝度情報または奥行情報のみを用いた場合に対する比較実験では、図 50 のグラフに示すように、提案手法と同様に移動物体の検出率が低いことが確認できる。しかし、提案手法が移動物体検出率で優れているのは、移動物体の輪郭付近の点と輝度値や奥行値の類似している点を同じ領域と判定できているためであると考えられる。



対象フレーム(Frame19)画像

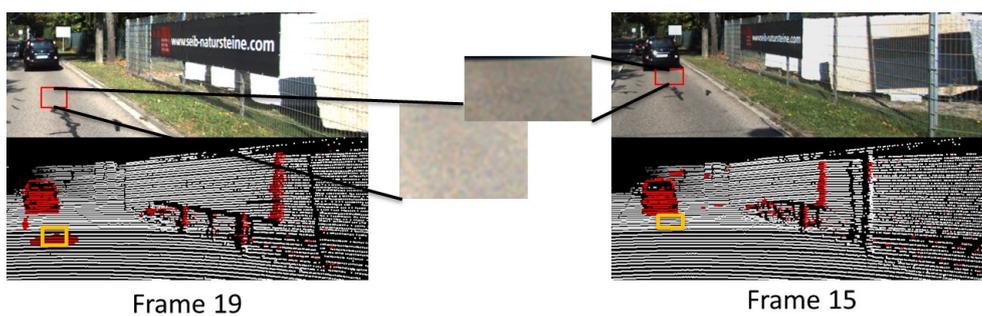


提案手法による結果



輝度情報のみを用いた結果

図 52 地面上の点が誤検出されている例



Frame 19

Frame 15

図 53 画像全体が明るくなった例

## 5. まとめ

本論文では、一回の移動計測により得られる同期の取れた、三次元点群、画像群およびカメラの位置姿勢情報を入力とし、三次元点群中から移動物体上の点を検出する手法を提案した。本研究では、輝度情報と奥行情報の整合性を統合したエネルギー関数を最小化することで、計測コストが少なく、カメラ移動や背景形状の制約を緩和した、汎用的な移動物体検出手法を提案した。実験では、提案手法による移動物体の20フレーム平均の検出率は最高で86.8%、識別精度は最高で97.8%となった。また、輝度情報または奥行情報のみを用いた場合に対する比較実験では、提案手法が輝度情報または奥行情報のみを用いた場合に比べ移動物体の検出率で優れていることを確認した。特に、輝度情報または奥行情報のみを用いた場合では検出することが難しかったカメラの奥行方向に低速で移動するような移動物体に対してもカメラが高速に移動している場合は正しく検出ができ、一般的な道路環境において提案手法が有効であることが確認された。

今後の展望として、アプリケーションなどへの応用が挙げられ、移動物体によって生成結果に歪みが生じる問題が解決されるのではないかと考えられる。

提案手法ではおおむね移動物体の検出が可能である一方、カメラが低速に移動する場合は、カメラの奥行方向に低速で移動する物体を検出することが困難であり、GPS・IMUにより得られる位置・姿勢情報の精度が十分でないことや、カメラの自動的な輝度補正が原因となり、遠方の地面点や細い柱状物体上の点などが誤検出されることがある。そこで今後の課題として、投影点の周辺の情報を用いてこれらに対応する必要がある。

## 謝辞

本研究を進めるにあたり，終始暖かくご指導，ご鞭撻頂いた視覚情報メディア研究室 横矢直和 教授に心より感謝申し上げます。また，本学での研究におけるあらゆる面での多大なご助言，ご協力を賜りましたことを深く感謝いたします。副指導教員として有益な御助言を頂きましたロボティクス研究室 小笠原司 教授ならびに，本研究を進めるに当たり終始細やかなご指導，ご助言頂いた視覚情報メディア研究室 佐藤智和 准教授に厚く御礼申し上げます。本研究へのご助言，ご協力を頂いた視覚情報メディア研究室 河合紀彦 助教，中島悠太 助教に深く感謝致します。研究について多大な御助言，ご指導，ご協力を賜った株式会社豊田中央研究所 町田貴史 氏に心より深く感謝致します。また，研究室での生活を支えていただいた視覚情報メディア研究室 石谷由美 女史に心より感謝いたします。本研究の遂行にご助言くださり，実装等にご協力していただきました，光メディアインタフェース研究室 青砥隆仁 氏，視覚情報メディア研究室 武原光 氏，大谷まゆ 女史，Fabian Lorenzo Baytion Dayrit 氏，Antonio Tejero de Pablos 氏をはじめ，研究室の皆様感謝の意を表します。

最後に，長い学生生活を温かく見守って下さいました両親に心より感謝の意を表します。

## 参考文献

- [1] Fumio Okura, Masayuki Kanbara, and Naokazu Yokoya. Mixed-reality world exploration using image-based rendering. *ACM Journal on Computing and Cultural Heritage*, Vol. 8, No. 2, pp. 9:1–9:26, 2015.
- [2] Akitaka Oko, Tomokazu Sato, Hideyuki Kume, Takashi Machida, and Naokazu Yokoya. Evaluation of image processing algorithms on vehicle safety system based on free-viewpoint image rendering. In *Proc. IEEE Intelligent Vehicles Symposium*, pp. 706–711, 2014.
- [3] 松元裕哉, 河合紀彦, 佐藤智和, 横矢直和. 実時間自由視点画像生成を用いた実画像に基づくドライブシミュレータシステム. 電子情報通信学会 技術研究報告, PRMU2014-174, 2015.
- [4] Toshihiro Asai, Masayuki Kanbara, and Naokazu Yokoya. Data acquiring support system using recommendation degree map for 3D outdoor modeling. In *Proc. Electronic Imaging*, pp.64910H-1-64910H-8, 2007.
- [5] Harlyn Baker. Three-dimensional modelling. In *Proc. International Joint Conference on Artificial Intelligence*, pp. 649–655, 1977.
- [6] Jiangjian Xiao and Mubarak Shah. From images to video: View morphing of three images. In *Proc. Vision, Modeling and Visualization Workshop*, pp. 495–502, 2003.
- [7] Leonard McMillan and Gary Bishop. Plenoptic modeling: An image-based rendering system. In *Proc. ACM SIGGRAPH*, pp. 39–46, 1995.
- [8] Gaurav Chaurasia, Sylvain Duchene, Olga Sorkine-Hornung, and George Drettakis. Depth synthesis and local warps for plausible image-based navigation. *ACM Transactions on Graphics*, Vol. 32, No. 3, pp. 30:1–30:12, 2013.

- [9] Tsunetake Kanatani, Hideyuki Kume, Takafumi Taketomi, Tomokazu Sato, and Naokazu Yokoya. Removal of moving objects from point cloud data for 3D modeling of outdoor environments. *IEEEJ Transactions on Image Electronics and Visual Computing*, Vol. 3, No. 1, pp. 54–62, 2015.
- [10] Arturo Flores and Serge Belongie. Removing pedestrians from google street view images. In *Proc. IEEE International Workshop on Mobile Vision*, 6 pages, 2010.
- [11] Norihiko Kawai, Kotaro Machikita, Tomokazu Sato, and Naokazu Yokoya. Video completion for generating omnidirectional video without invisible areas. *IPSJ Transactions on Computer Vision and Applications*, Vol. 2, pp. 200–213, 2010.
- [12] Norihiko Kawai, Naoya Inoue, Tomokazu Sato, Fumio Okura, Yuta Nakashima, and Naokazu Yokoya. Background estimation for a single omnidirectional image sequence captured with a moving camera. *IPSJ Transactions on Computer Vision and Applications*, Vol. 6, pp. 68–72, 2014.
- [13] 内山寛之, 出口大輔, 高橋友和, 井手一郎, 村瀬洋. 複数画像系列の部分画像選択に基づく移動物体を含まない車載カメラ映像の生成. 電子情報通信学会論文誌 D, Vol. 94, No. 12, pp. 2093–2104, 2011.
- [14] 高橋英之, 堀磨伊也, 神原誠之, 横矢直和. 全天球画像データベース作成のための動物体除去と色調統一. 画像の認識・理解シンポジウム (MIRU) 講演論文集, pp. 1933–1940, 2010.
- [15] Jizhou Yan, Dongdong Chen, Heesoo Myeong, Takaaki Shiratori, and Yi Ma. Automatic extraction of moving objects from image and LIDAR sequences. In *Proc. International Conference on 3D Vision*, pp. 673–680, 2014.
- [16] Carsten Rother, Vladimir Kolmogorov, and Andrew Blake. Grabcut: Interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics*, Vol. 23, No. 3, pp. 309–314, 2004.

- [17] Lili Huang and Matthew Barth. Tightly-coupled LIDAR and computer vision integration for vehicle detection. In *Proc. Intelligent Vehicles Symposium*, pp. 604–609, 2009.
- [18] Cristiano Premebida, Oswaldo Ludwig, and Urbano Nunes. LIDAR and vision-based pedestrian detection system. *Journal of Field Robotics*, Vol. 26, No. 9, pp. 696–711, 2009.
- [19] Alberto Elfes. Using occupancy grids for mobile robot perception and navigation. *IEEE Computer*, Vol. 22, No. 6, pp. 46–57, 1989.
- [20] Matthias Lindström and Jan-Olof Eklundh. Detecting and tracking moving objects from a mobile platform using a laser range scanner. In *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1364–1369, 2001.
- [21] Boris Kluge, Christian Köhler, and Erwin Prassler. Fast and robust tracking of multiple moving objects with a laser range finder. In *Proc. IEEE International Conference on Robotics and Automation*, pp. 1683–1688, 2001.
- [22] Erwin Prassler, Jens Scholz, and Paolo Fiorini. A robotics wheelchair for crowded public environment. *IEEE Robotics and Automation Magazine*, Vol. 8, No. 1, pp. 38–45, 2001.
- [23] Chris Urmson, Joshua Anhalt, Drew Bagnell, Christopher Baker, Robert Bitner, MN Clark, John Dolan, Dave Duggins, Tugrul Galatali, Chris Geyer, et al. Autonomous driving in urban environments: boss and the urban challenge. *Journal of Field Robotics*, Vol. 25, No. 8, pp. 425–466, 2008.
- [24] Sören Kammel, Julius Ziegler, Benjamin Pitzer, Moritz Werling, Tobias Gindele, Daniel Jagzent, Joachim Schröder, Michael Thuy, Matthias Goebl, Felix von Hundelshausen, et al. Team AnnieWAY’s autonomous system for the 2007 DARPA urban challenge. *Journal of Field Robotics*, Vol. 25, No. 9, pp. 615–639, 2008.

- [25] Michael Montemerlo, Jan Becker, Suhrid Bhat, Hendrik Dahlkamp, Dmitri Dolgov, Scott Ettinger, Dirk Haehnel, Tim Hilden, Gabe Hoffmann, Burkhard Huhnke, et al. Junior: The stanford entry in the urban challenge. *Journal of Field Robotics*, Vol. 25, No. 9, pp. 569–597, 2008.
- [26] 高橋英之, 堀磨伊也, 神原誠之, 横矢直和. 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-143, pp. 41–48, 2004.
- [27] Yuping Shen, Fei Lu, Xiaochun Cao, and Hassan Foroosh. Video completion for perspective camera under constrained motion. In *Proc. International Conference on Pattern Recognition*, pp. 63–66, 2006.
- [28] Dorin Comaniciu and Peter Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 5, pp. 603–619, 2002.
- [29] Yasuyuki Matsushita, Eyal Ofek, Weina Ge, Xiaoou Tang, and Heung-Yeung Shum. Full frame video stabilization with motion inpainting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 7, pp. 1150–1163, 2006.
- [30] 堀磨伊也, 河合紀彦, 神原誠之, 新井イスマイル, 西尾信彦, 横矢直和. パノラマビューシステムのための死角領域修復とプライバシー保護を行った全天球画像生成. 画像の認識・理解シンポジウム (MIRU) 講演論文集, pp. 923–929, 2011.
- [31] Bastian Leibe, Aleš Leonardis, and Bernt Schiele. Robust object detection with inter-leaved categorization and segmentation. *International Journal of Computer Vision*, Vol. 77, No. 1-3, pp. 259–289, 2008.
- [32] Changchang Wu. Visualsfm : A visual structure from motion system. ”<http://ccwu.me/vsfm/>”, 2013.

- [33] Michal Jancosek and Tomas Pajdla. Multi-view reconstruction preserving weakly supported surfaces. In *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 3121–3128, 2011.
- [34] Jaakko Astola, Petri Haavisto, and Yrjo Neuvo. Vector median filters. In *Proc. IEEE*, Vol. 78, No. 4, pp. 678–689, 1990.
- [35] Yuri Boykov, Olga Veksler, and Ramin Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 11, pp. 1222–1239, 2001.
- [36] Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. Vision meets robotics: The kitti dataset. *International Journal of Robotics Research*, Vol. 32, No. 11, pp. 1231–1237, 2013.