

交通環境と人間行動を理解・再現する

西の都で実践する画像情報とXRの技術

小野晋太郎

福岡大学 工学部/大学院工学研究科 准教授

東京大学 生産技術研究所 リサーチフェロー

ハーモニック・モビリティ研究センター (ITSセンター)

福岡大学工学部 小野(晋)研究室



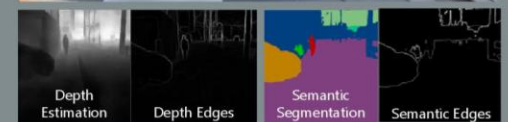
小野 准教授



廖 助教

- のりもの・交通、自動運転
- 画像情報処理、CG、VR、AR
- 地図・地理情報
- 交通環境の理解、人の行動理解
- 実映像VR・自由視点ドライビングシミュレータ
- デジタル文化遺産

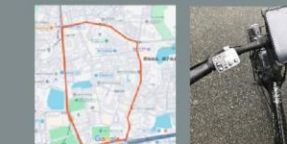
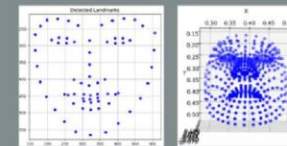
交通環境の理解



建設機械の 知能化



人の行動理解



実映像VR／自由視点ドライビングシミュレータ (DS)



実映像 と CG で運転操作は同じか？



デジタル文化遺産





小野晋太郎

- 1978.10 山口県生まれ、広島・松江育ち
- 1997.3 松江南高校 卒
- 2006.3 東京大学 博士(情報理工学)
恩師: 池内教授
- 2006.4 東京大学 生産技術研究所 特任助教
- 2013.4 同 特任准教授 (~2022.11)
- 2021.4 福岡大学工学部 准教授 (現職)
- 2017.1~2020.5
兼) ホンダ・リサーチ・インスティテュート・ジ ャパン (HRI-JP)

専門: 画像情報工学、ITS



福岡大学

- 中四国・九州で最大規模の私立総合大学
- 9学部31学科、学生数2万人
- 博多から20分、天神から15分 (直通)

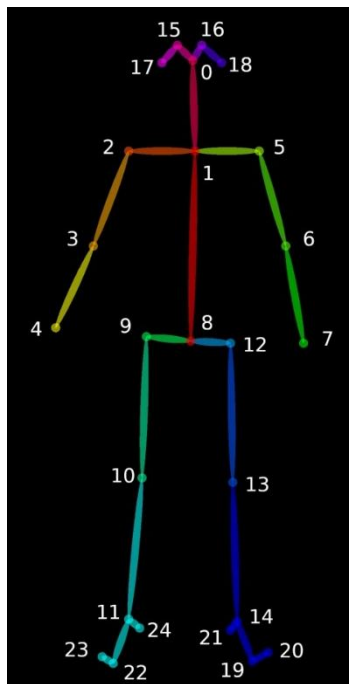
著名出身者

- 博多華丸 (工学部 → 吉本興業)
- 博多大吉 (経済学部 → 吉本興業)
- 梅野隆太郎 (スポーツ科学部 → 阪神)
- 榎田大樹 (スポーツ科学部 → 阪神・西武)
- 生野陽子 (法学部 → フジテレビ)
- 近藤奈央 (人文学部 → 気象予報士(NHKおはよう日本))
- 石川佳純ご両親 (体育学部・商学部)

交通環境の理解

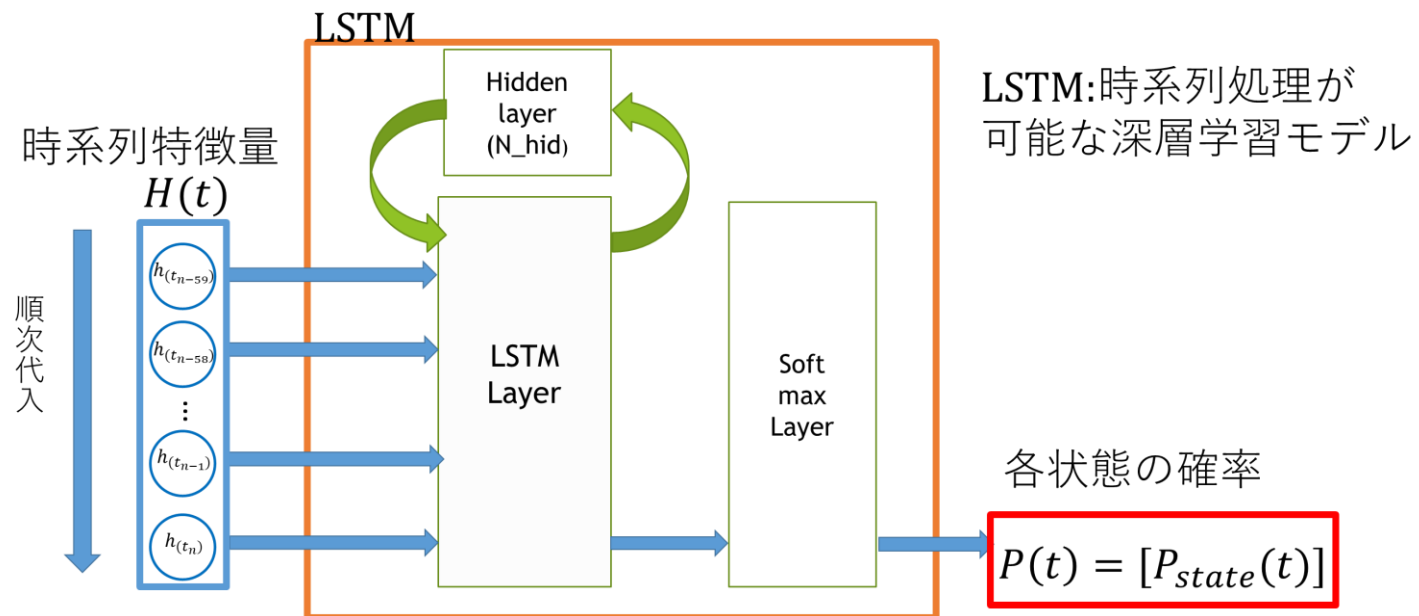
手信号の認識

骨格抽出 (OpenPose利用)



- 上記18箇所の正規化相対2D座標
- 各座標の推定信頼度
- これらの過去2秒分を学習

信号現示の推定（LSTMに基づく深層学習）



LSTM:時系列処理が可能な深層学習モデル

各状態の確率

$$P(t) = [P_{state}(t)]$$

2値分類

● Red: %

● Green: %

3値分類

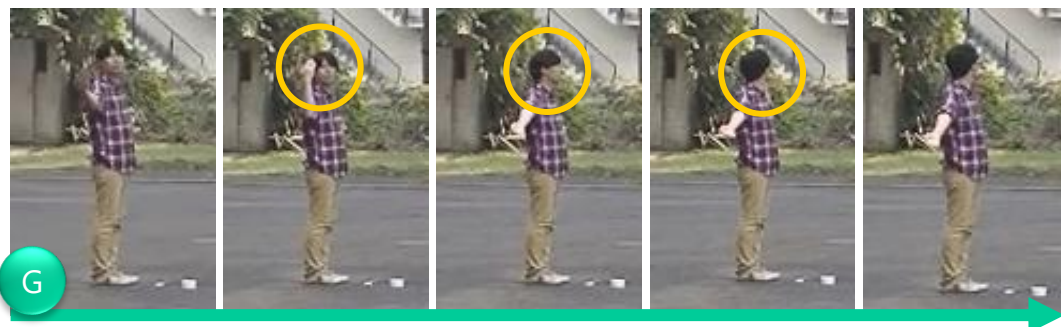
● Red: %

● Green: %

● Other: %

Yellowは
Redに含める

Red, Green に共通する動作がある (!)



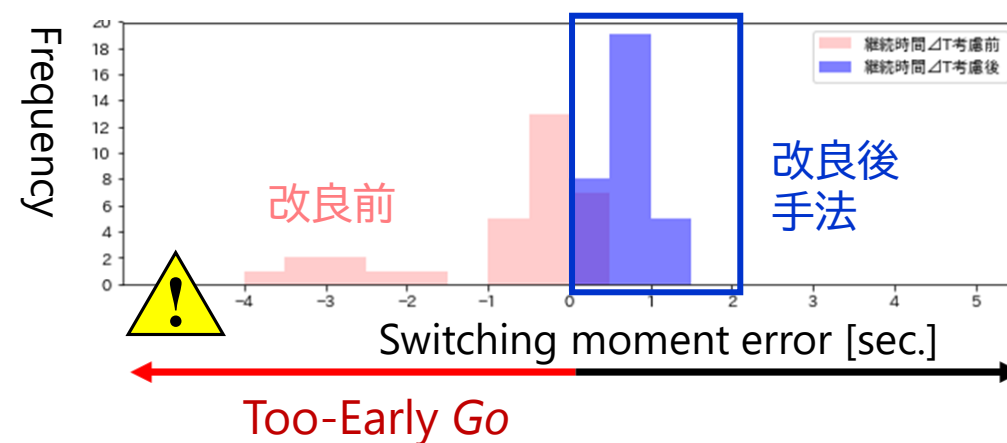
- 別扱いにして Red/Green/Other の3値に分類
- Other 判定時 → 一つ前の判定を引き継ぐ
- Green 判定が t 秒継続 → Go と判定

画像フレーム数による評価

- 判定の正確度: 98%
- StopをGoと判定する割合: 0.23%

判定の時間差による評価

Stop→Go判定の時間差の頻度分布



- 早すぎるGo判定: なし
- 遅すぎるStop判定: 最大0.43 s

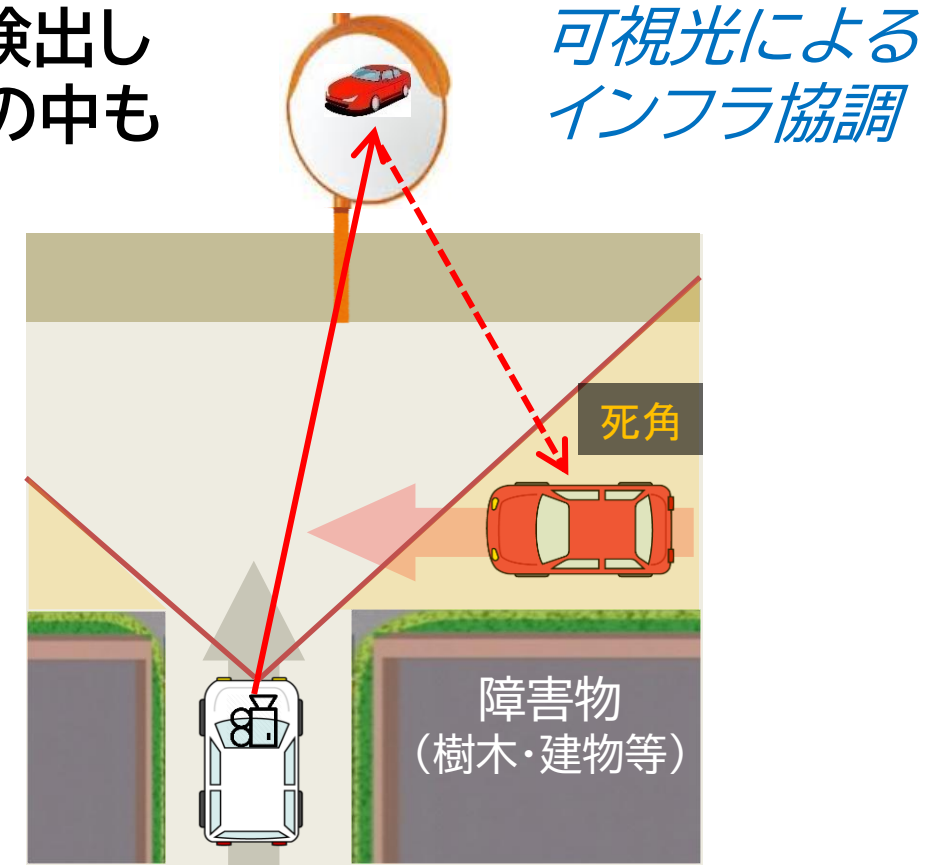
車載カメラとカーブミラーを利用した 死角の危険予知

25





ミラーを検出し
さらにその中も
見る



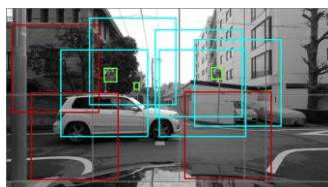
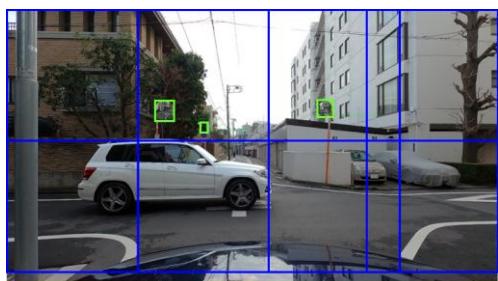
- 出会い頭の接触事故が多く発生
- 接近物体を自車から直接観測することが困難
- 高度な認識システムが必要

- カーブミラーは主要箇所に設置済み
- 車載カメラも一般化
- 運転支援(レベル2以下)にも活用可能

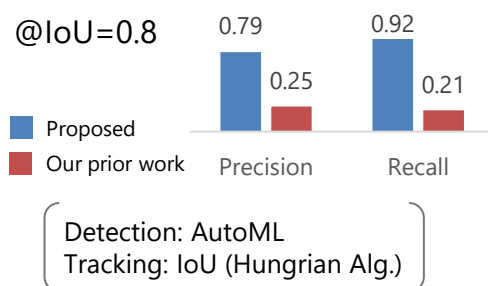
ミラー内の物体が接近中かどうかを**走行中の**車載カメラで認識

ミラー自体の認識

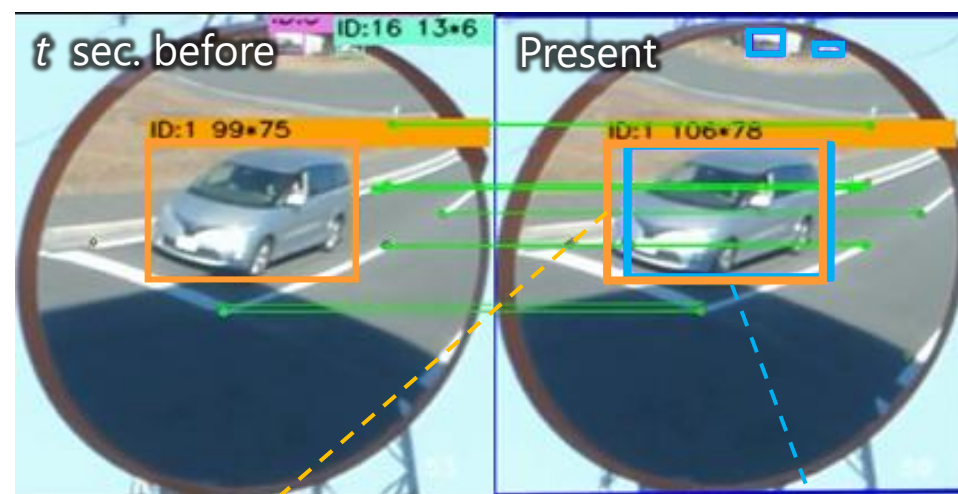
4Kカメラ画像を分割処理



無作為に
切り出して
学習に利用



ミラー内に映る路上物体の認識・接近判別



現在の車両検出箇所

過去の車両検出箇所
(現在視点に変換して表示)

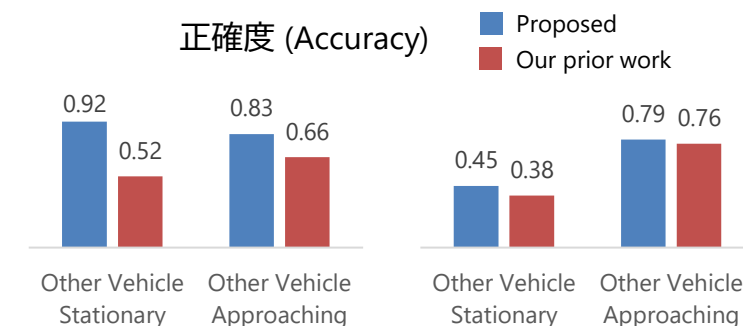
過去の視点を現在の視点に変換*してから
路上物体の変位を算出

*射影変換で近似

[Detection: efficientdet-d7x (COCO), Tracking: IoU (Hungarian Alg.)]

結果・考察

- 視点変換が可能な場合は、80% (フレーム数) 以上の正確度で判別
- 性能は特徴点の個数・分布に依存 (60点程度は特徴点ペアが必要)
- 今後: ミラーに関する地図情報の利用、あるいは構築・更新





元画像



二値画像



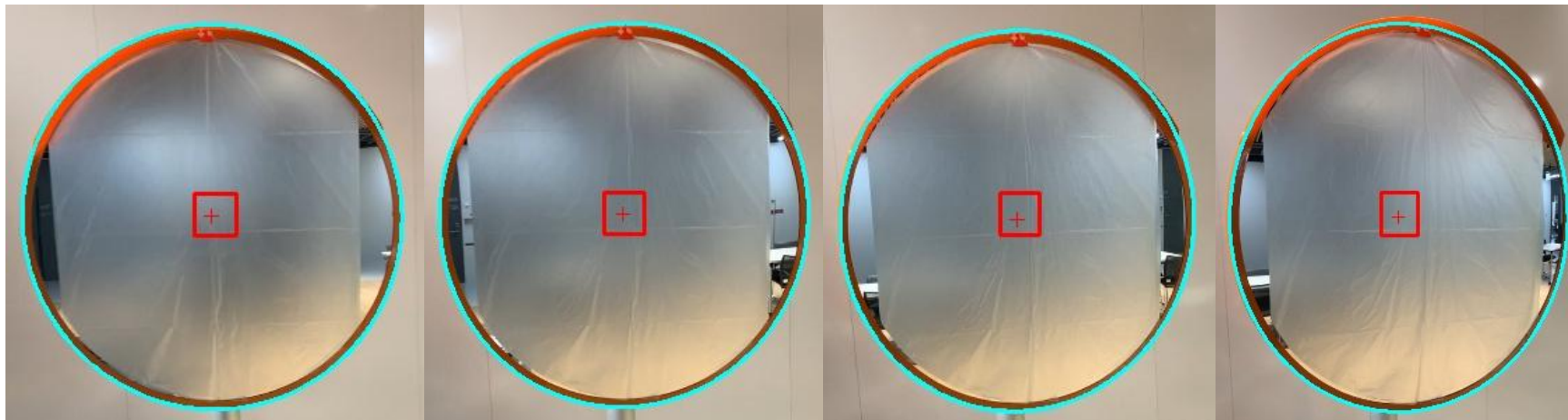
エッジ画像



輪郭画像



楕円検出



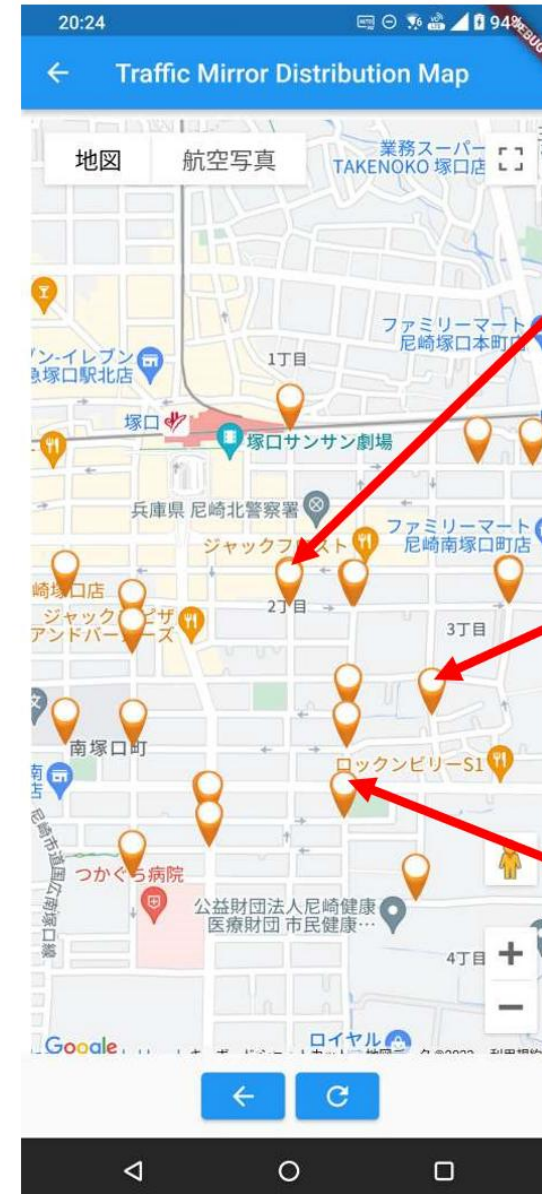
実測角度	0°	10°	20°	30°
推定角度	10.4°	16.0°	21.9°	29.2°

オープンデータを活用

- Open Street Map、Googleストリートビュー

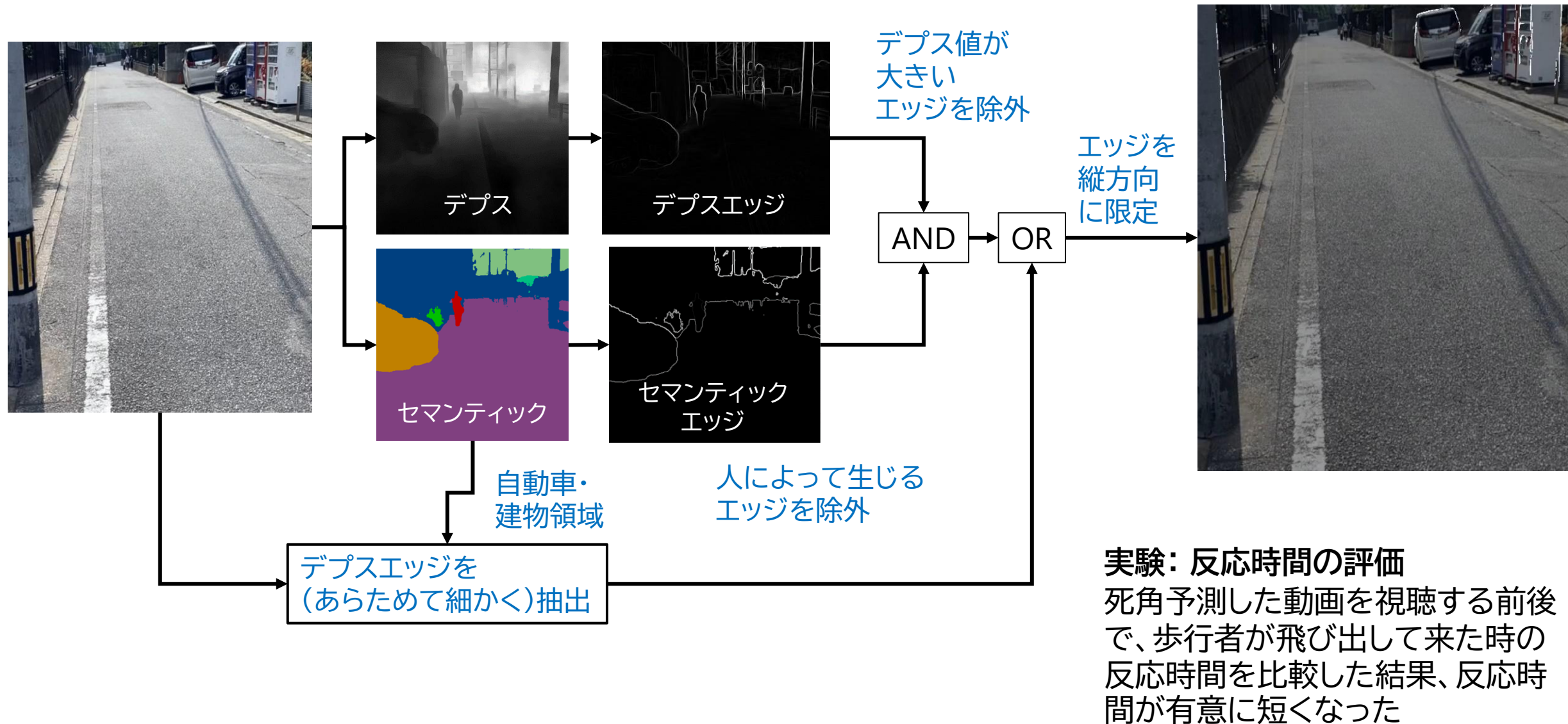
今後の展開

- カーブミラーマップを事前知識として活用したより効率的なカーブミラー検出
- 設置角度のズレなどの検出



目的：車載(自転車)カメラから死角を予測する

死角強調結果



奥行き(深度、デプス)推定

1台のカメラ画像からシーンの奥行きを推定



Boosting Monocular Depth Estimation

Miangoleh, S.M.H et al.: Boosting monocular depth estimation models to high-resolution via content-adaptive multi-resolution merging. CVPR 2021.

セマンティックセグメンテーション

画像内の各画素がどの物体種別に属するかを判別



STEGO

Hamilton, M et al. (2022): Unsupervised semantic segmentation by distilling feature correspondences. arXiv:2203.08414.

死角：強調なし
歩行者：飛び出す



死角：強調あり
歩行者：現れない



死角：強調なし
歩行者：飛び出す



反応時間を計測

歩行者に気付いたらキーボードを押す



T_B

動画を見る

反応時間を計測

歩行者に気付いたらキーボードを押す



T_A

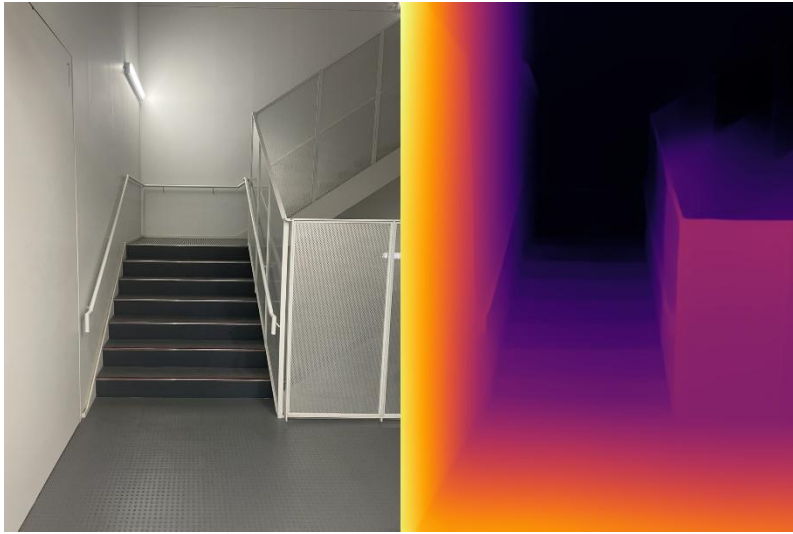
結果

$$\overline{T}_A = 0.744 \text{ s}$$

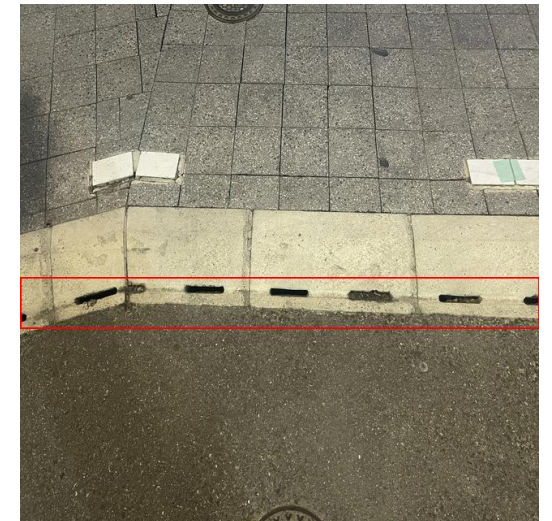
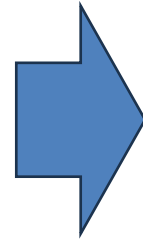
$$\overline{T}_B = 0.803 \text{ s}$$

有意差あり
($P = 0.047 < 0.05$)

N=13
全員異なるシーン



カラー画像 (深度画像)



-



- 検出サイズ・位置から
離隔距離に換算
- 30m先で誤差10cm程度

★ 公道上で方向転換が必要な駐車場の判別と分類



このような駐車場はほぼ2タイプに集約できる

タイプ1: 縦方向の不足
小規模な路面店など
タイプ2: 横方向の不足
小規模な空き地など

その他: (少数の例外)
駐車済み車両の台数・位置や
塀・フェンスの有無により
条件が異なる

(半)自動判別方法

空撮画像・デジタル地図
(道路が横方向になるように回転)

(A) 駐車車両の検出*
YOLOv5 (+COWC dataset)

* 自動化

$\theta \cong 90^\circ$

(B) 駐車区画端の検出*
マルチプレートマッチング
区画幅の n 倍を考慮

区画端から
公道までの
距離

$d < d_0$

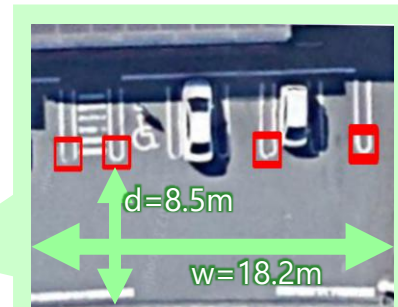
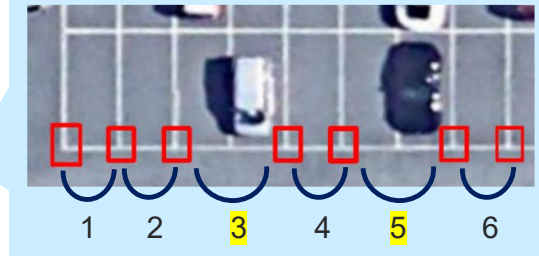
公道上で
切り返し
が必要
(タイプ1)

敷地の
幅

$w < w_0$

同
(タイプ2)

敷地内で
方向転換可能



交通環境における人の行動理解

自動運転に対する信頼度推定・自転車のドライバモニタリング

歩行者が自動運転車に対して抱く信頼度の推定

共同研究: 東京大学、日本自動車研究所(JARI)

Ryota Masuda, Shintaro Ono, Toshihiro Hiraoka, Yoshihiro Suda, "Trust Estimation for Autonomous Vehicles by Measuring Pedestrian Behavior in VR", ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2023.3

Ryota Masuda, Shintaro Ono, Toshihiro Hiraoka, Yoshihiro Suda, "Trust Estimation of Pedestrians Crossing Street Based on Measurement of Pedestrian Behavior in VR", Int'l Symp. on Future Active Safety Technology toward Zero Accidents (fast-zero), 2023.11



THE UNIVERSITY OF TOKYO



FUKUOKA
UNIVERSITY



- 自動運転化でドライバが不在になると、他の交通参加者とのインタラクションが不可能に
- 自動運転車に対する不信→歩行者は横断できない、過信→飛び出してしまう
- **信頼度を推定**して、適切な信頼度に調整

イメージ

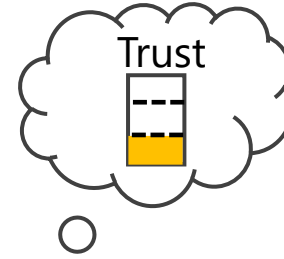


- 森山, 浜本, 栗, 平岡, 小野:「実環境における横断歩道接近車両に対する歩行者の信頼度推定(第1報)」, 自動車技術会秋季大会学術講演会, 2024.10
- R. Masuda, S. Ono, T. Hiraoka, Y. Suda: "Trust Estimation of Pedestrians Crossing Street Based on Measurement of Pedestrian Behavior in VR", Int'l Sym. on Future Active Safety Technology toward Zero Accidents (FAST-zero), 2023.11

■ 推定対象（目的変数）

- 歩行者の信頼度

- 横断前3段階、横断中3段階



■ 手掛かり（説明変数）

- 基本情報

- 歩行者から車両までの距離
- 歩行者の道路横断方向の位置・速度
- 車両の進行方向の位置・速度
- 横断開始までの経過時間

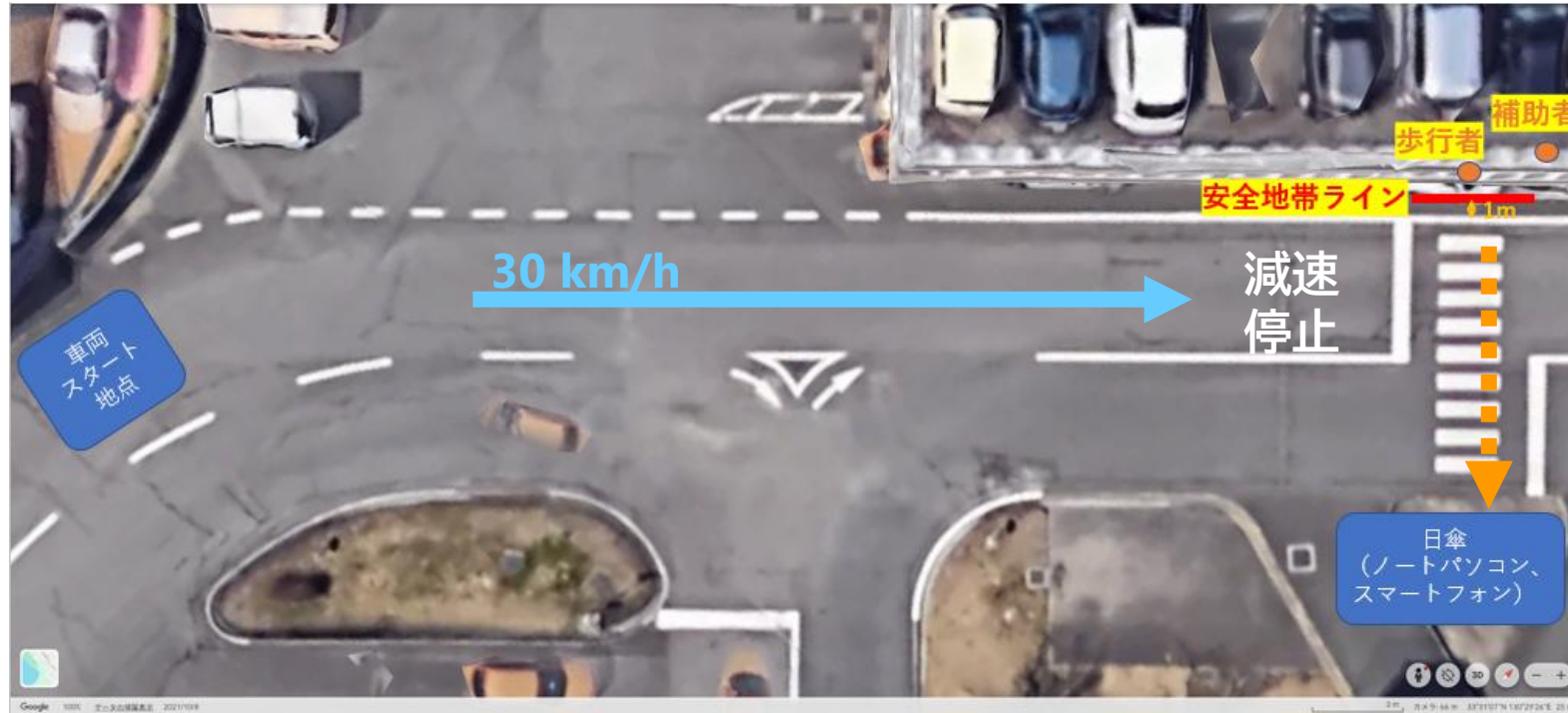
- さらに詳細な情報

- 骨格座標
- 歩行者の車両注視確率



- 信頼度の説明変数の選別(相関分析)
- 信頼度を推定できることの事前確認
- 注視確率に関する知見の獲得





協力: 南福岡自動車学校



データ収集

- GNSS, LiDAR, 車載カメラで走行データを記録
- フロントガラスにマジックミラーを貼付してアイコンタクトを防止
- 信頼度の正解値はボタンで入力

学習

- LSTM, AtLSTM, Transformer
- データ数のバランス
- 反応時間補正



車両注視確率

$$p(t) = f(x(t), R(t))$$

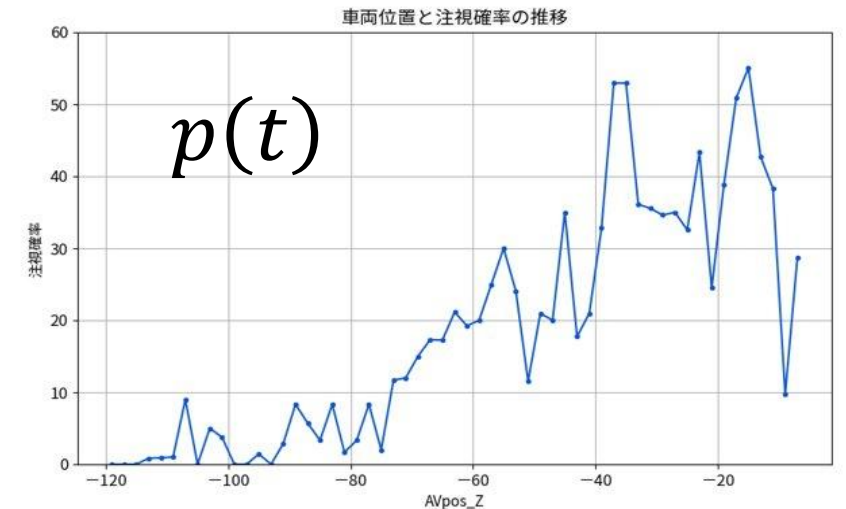
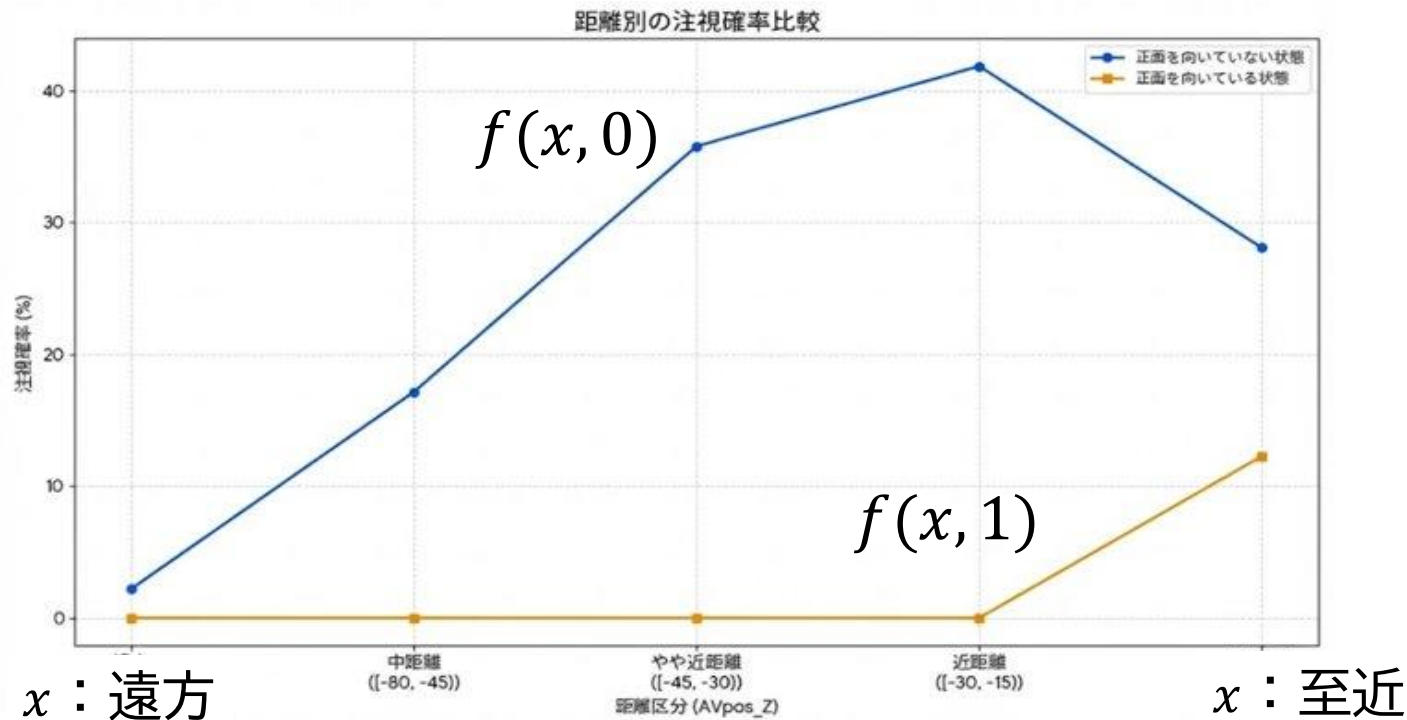
$x(t)$: 車両位置

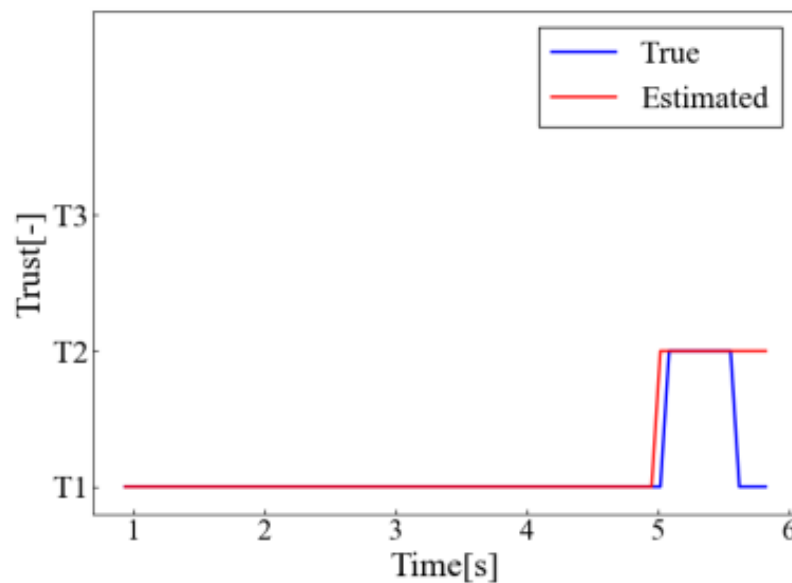
$R(t)$: 頭部姿勢

観測から頭部姿勢を
求めるのは難しい
横断方向(正面)を
向いているか否かの
2値とする



VR実験により $f(.,.)$ を決定

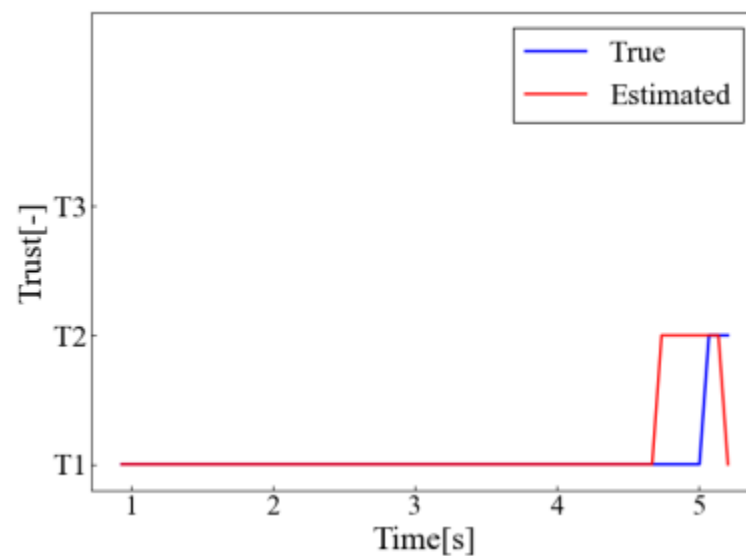


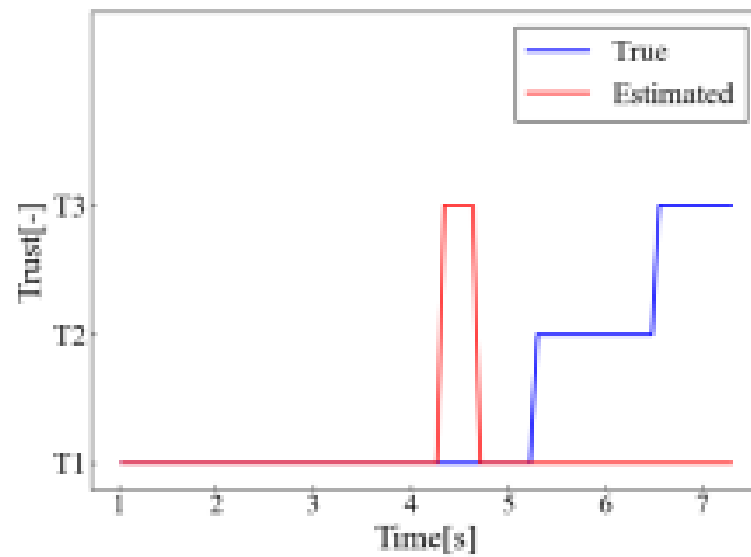


平均正解率
(時間一致率)

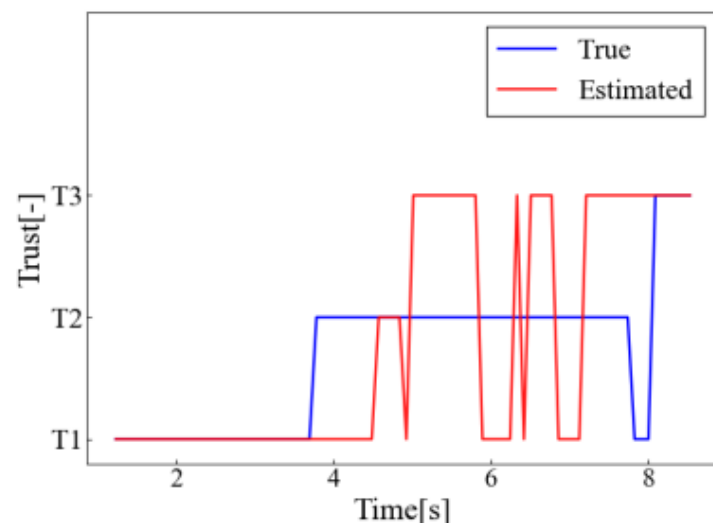
注視確率	正解率
考慮する場合	0.65
考慮しない場合	0.61

注視確率の効果





推定誤りが多い例
風の影響で服がはためき
骨格を誤って判定している



推定誤りが多い例
横断前の予備動作が小さい

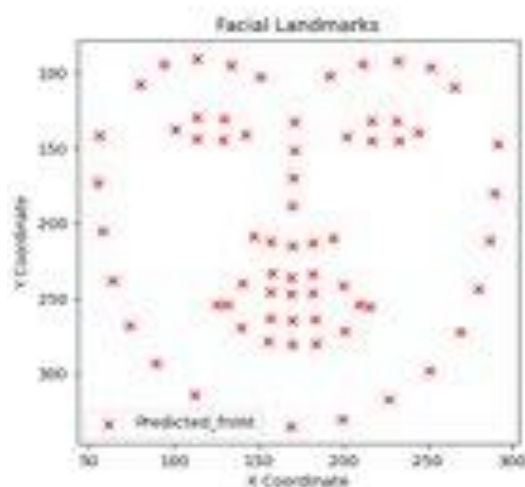
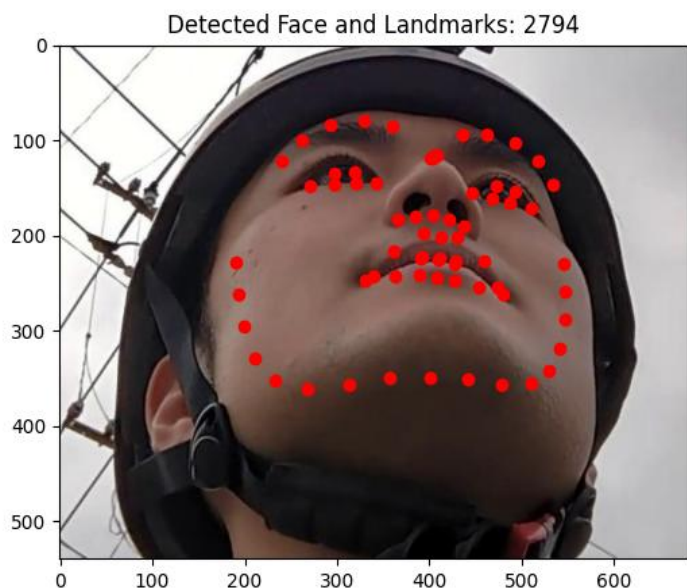
表情を反映した
「快適ルート」ナビなど

課題：ローアングル・時に逆光

※これは評価用カメラ



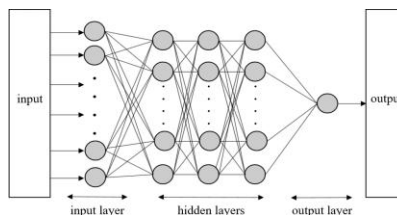
顔領域切り出し (RetinaFace) 特徴点抽出・顔向き推定 (PFLD)



輪郭、目、鼻、口など68点

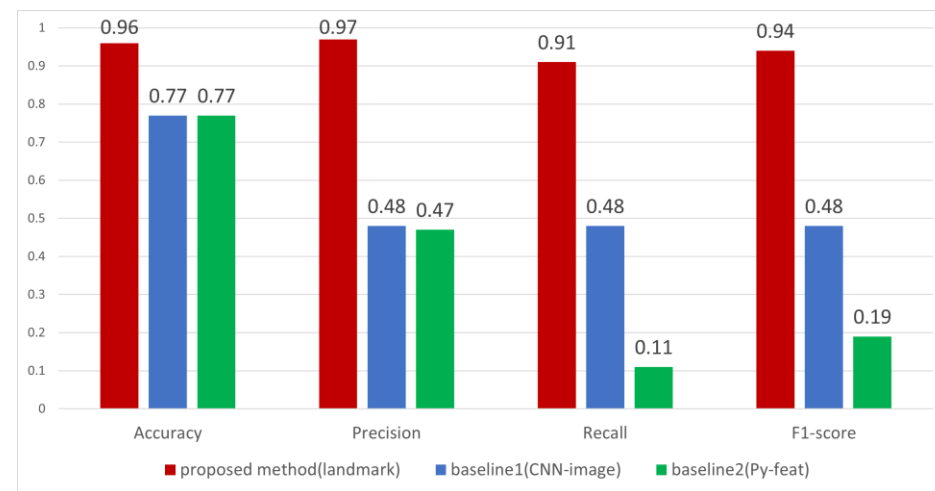
x_0	x_1	x_2	x_3	x_4	x_5	x_6	
183.4876	187.7374	194.033	207.7879	230.0054	263.6372	305.3344	y
173.5375	180.8946	190.1842	207.7653	234.0831	272.4012	319.409	

深層学習 (FFNN)



快適である
快適でない

結果



F値
 提案手法 0.94
 CNN* 0.48
 Py-Feat* 0.19

*CNN: 最も単純な画像分類
(特徴点を利用しない)

*Py-Feat: 既存の感情分類
(視点位置を考慮しない)

交通環境再現

実映像VR・ドライビングシミュレータ・街路AR

学習させたデータ



視点: 道路中央の直線上のみ

推定・生成したデータ

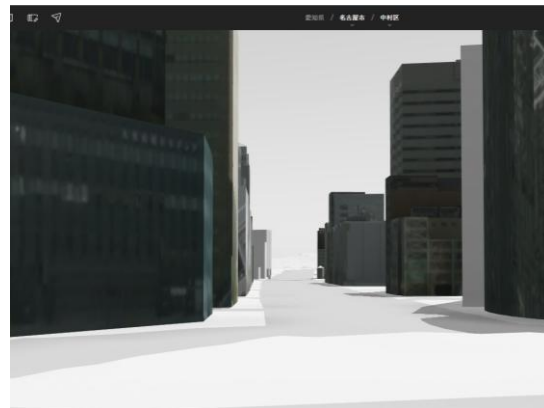


視点: 外れすぎなければ自由自在

従来型CG (3Dポリゴン+テクスチャ)



グランツーリスモ



PLATEAU

- 視点変更、物体の追加削除は自由自在
- 現実感が低い、工数が多い

実画像ベース



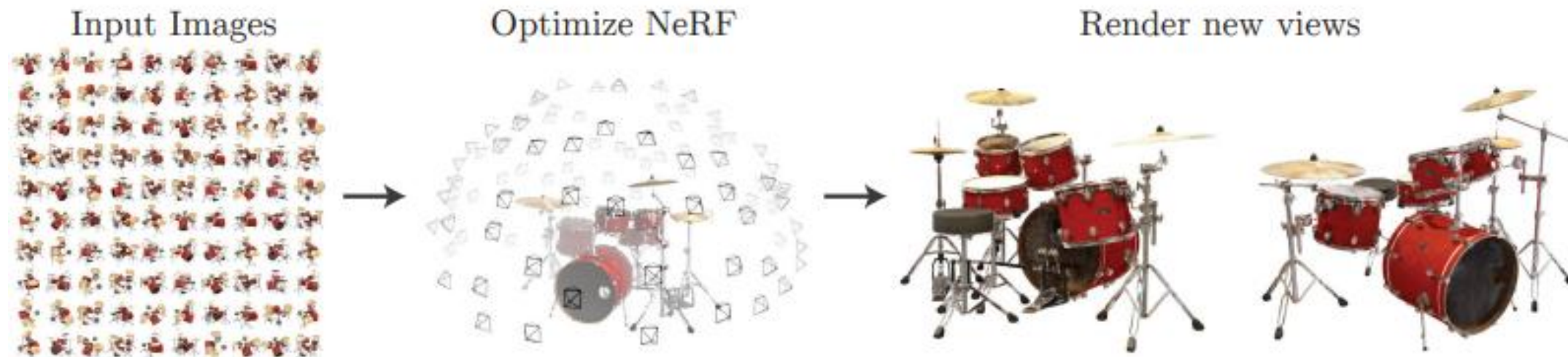
JR東日本トレインシミュレータ



ストリートビュー

- 現実感が高い、工数が少ない
- 視点変更、物体の追加削除には制限あり

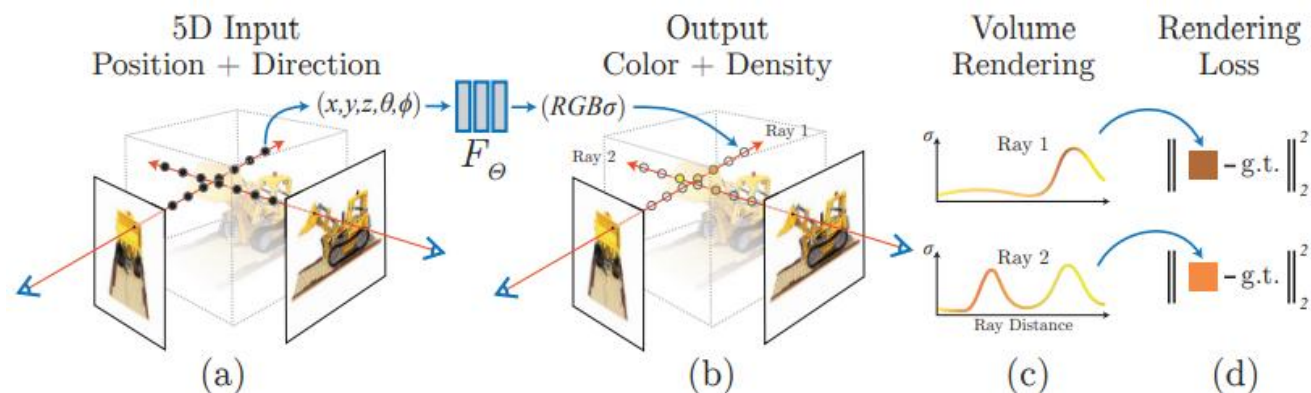
→ 機械学習で自由な
視点を再現 (NeRF)



撮影した画像

カメラの位置と姿勢を与える

撮影していない視点の
画像生成が可能

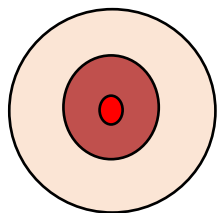


空間を格子状に分けて、色や密度を計算

弱点: 計算時間がかかる

- 学習も遅いが、学習済み結果を使った描画処理だけでも遅い
- 理由: 出力画素1ピクセルごとに「視線を奥へ向かって追いかけて、色や密度を積算するような処理」が必要

ドライバの
視野特性



中心視: $1 \sim 2^\circ$ 程度。詳細に認識

有効視野: $\pm 4 \sim 20^\circ$

中心視と同時に認識可能
速度や注視対象で変化

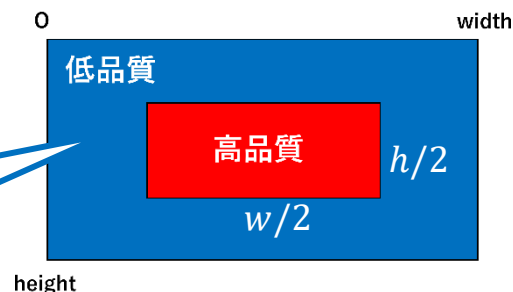
周辺視野: $180 \sim 210^\circ$

ぼんやりとしか見えない



1) 領域別 処理の導入

外周部において
空間探索処理の
回数を制限



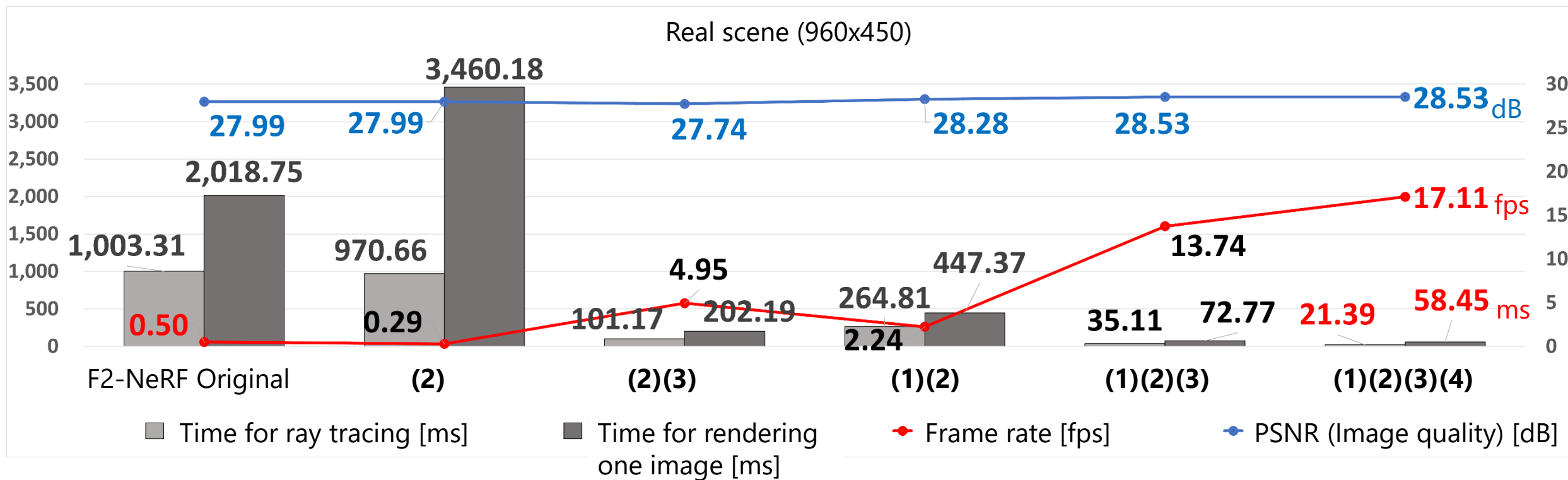
2) GPUによる 並列処理最適化

3) 分割処理によるメモリ負荷軽減



- 960×576 ピクセルのシーンにおいて
毎秒14コマでの描画速度を達成
- 元の手法は 毎秒 0.25 コマ
- 画質劣化も許容範囲 (PSNR=28.53)

画質を大きく劣化させることなく、オリジナル手法より 30倍高速化



(1) Region-Specific Process

(2) Parallel Processing by GPU

(3) Reducing GPU memory load

(4) Raising efficiency in ray tracing

撮影時とは異なる状況(他車両なし)で再現したい！



入力画像



出力画像
(開発手法)

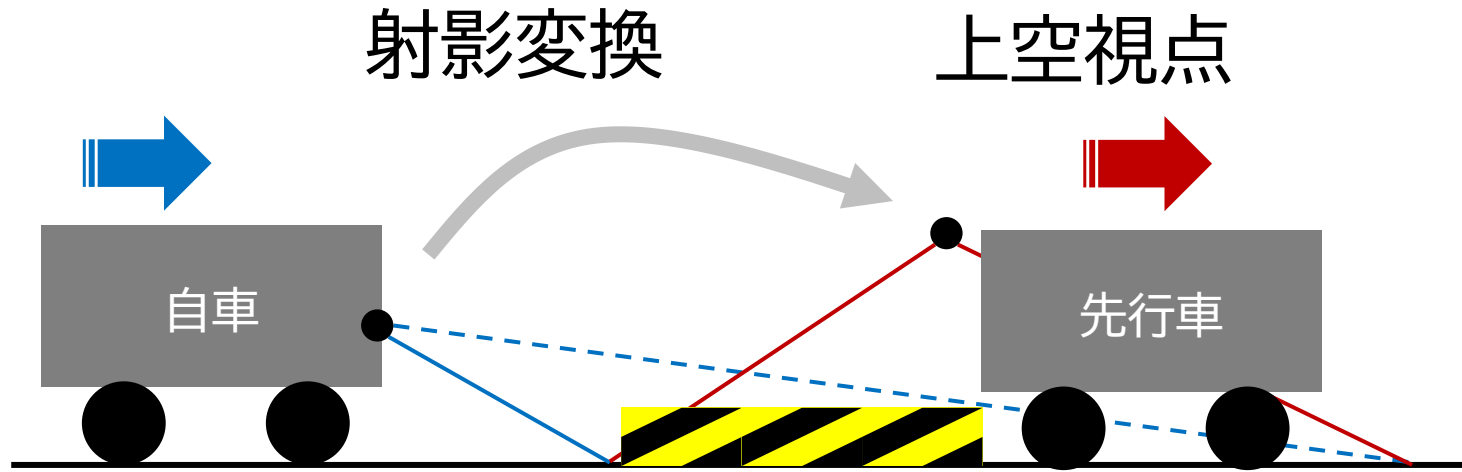


ProPainter [Zhou, ICCV2023]

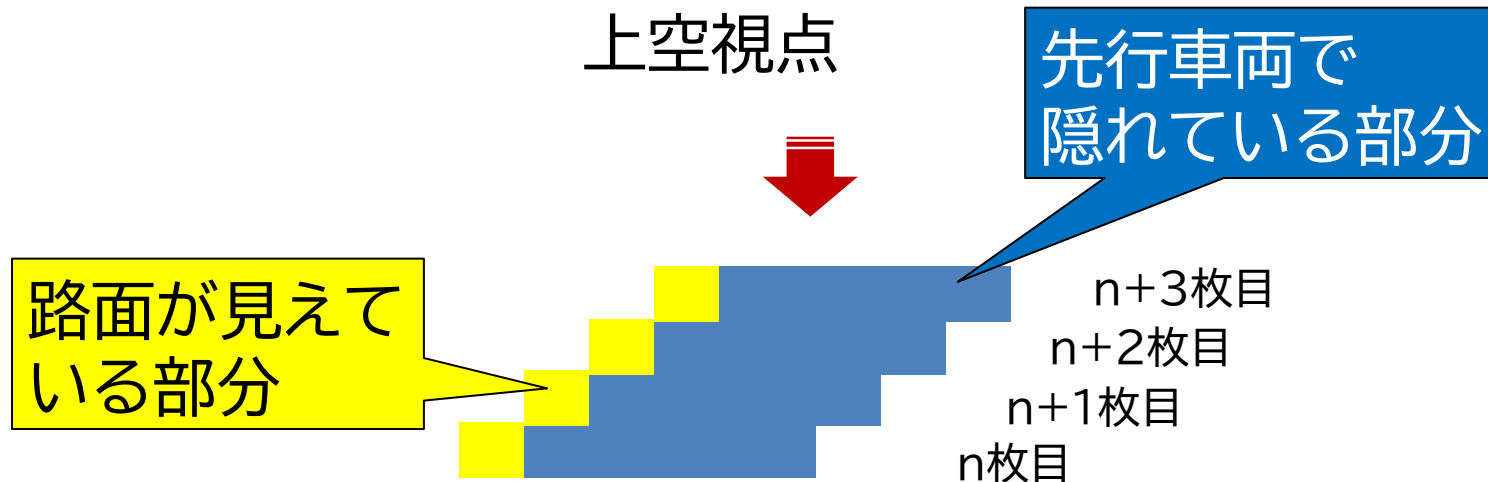


E2FGVI [Li, CVPR2022]

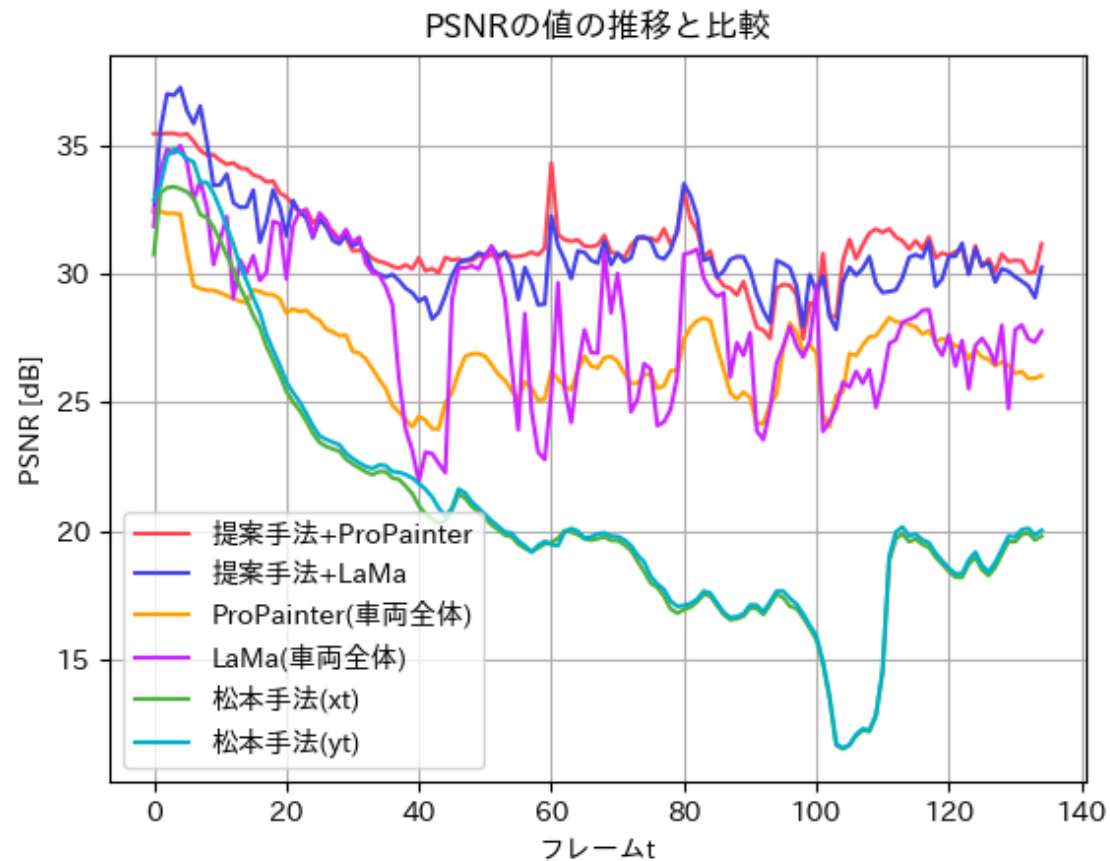
既存手法では路面が正しく
再現されない(後述)



車両が進行すれば
隠れていた部分は
現れる



車間の隙間に見える
情報を適切に集約・
再構成すれば良い



t=100付近



提案手法

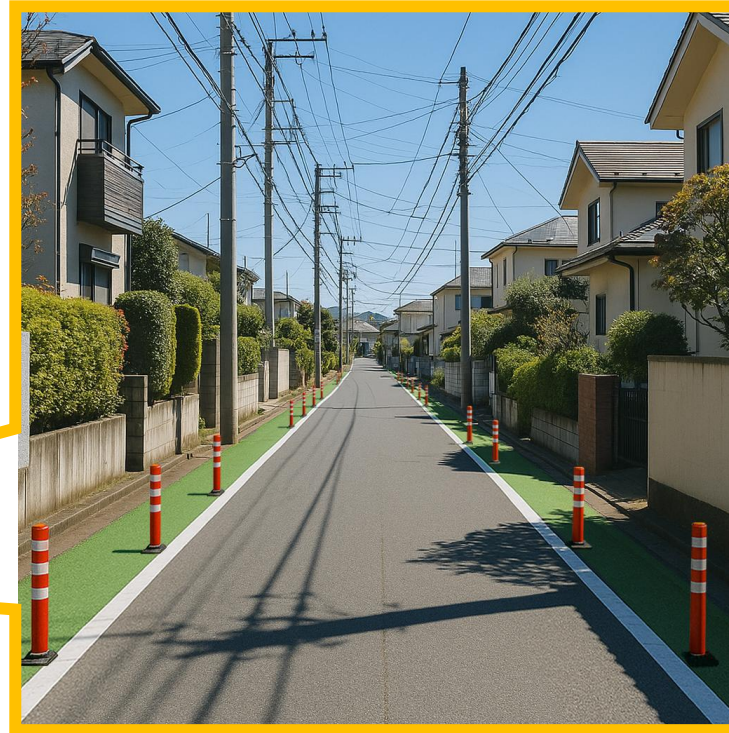


ProPainter

提案手法では
横断歩道も正しく再現



道路管理者



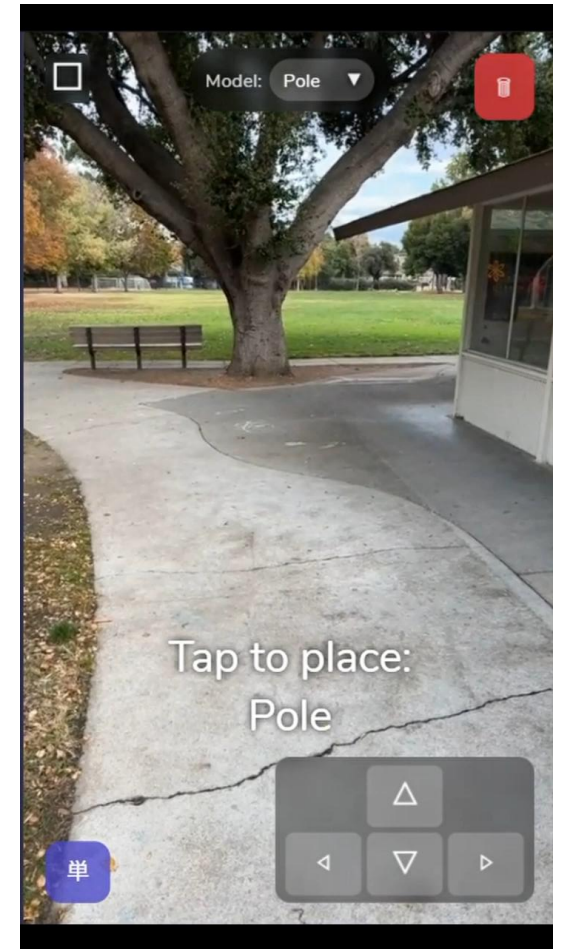
カラー舗装、ボラードなど

自治体職員からの主な意見:

- 直感的でわかりやすい
- 庁内説明や合意形成においてイメージを共有しやすい
- 設計・検討段階で可視化できる点が有用
- そのほか、機能追加要望など。否定的な意見なし

一般利用者による5段階評価

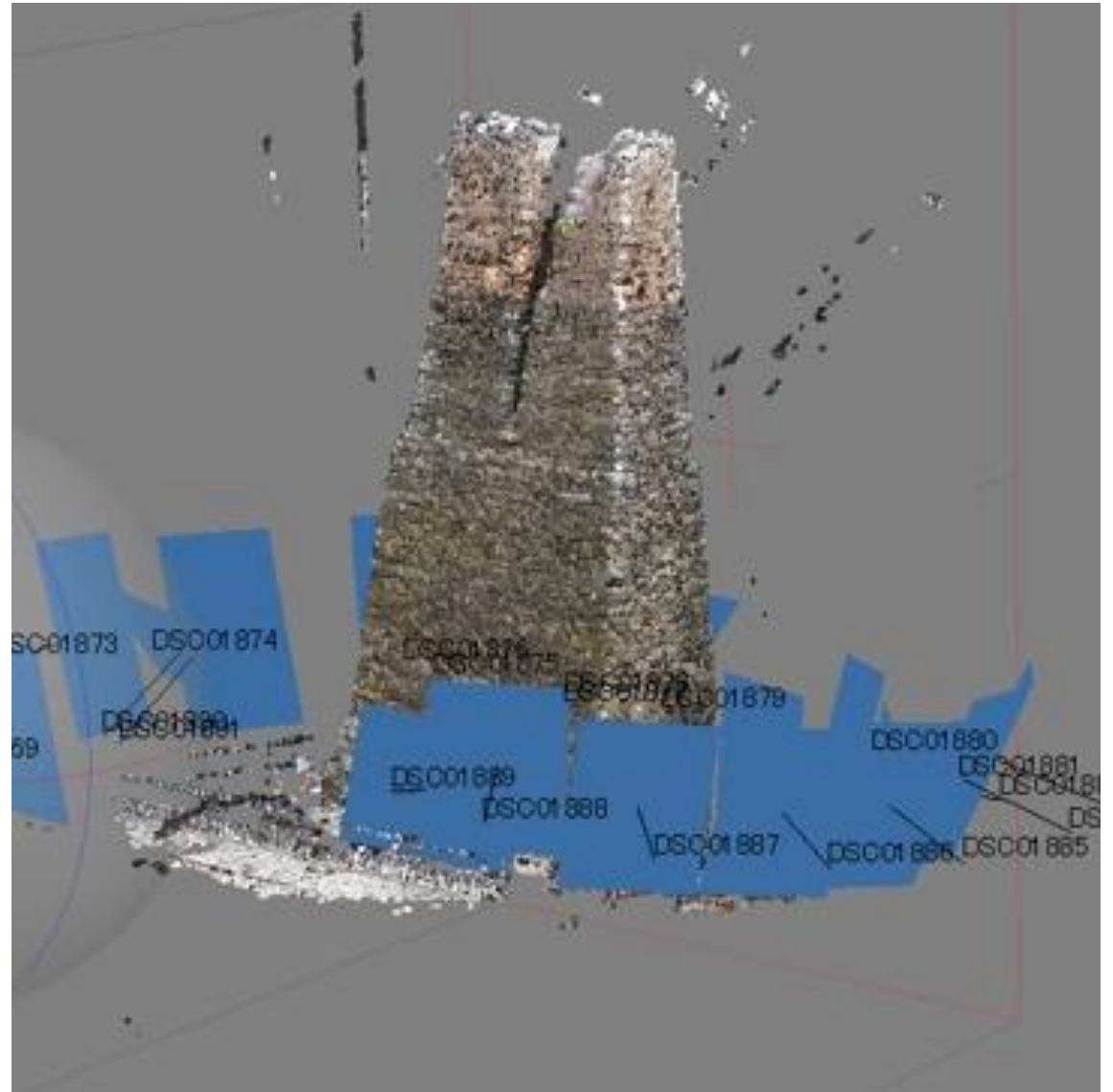
- イメージしやすさ 3.6
- 操作の直感性 5.0
- 安全対策への有用性 4.0

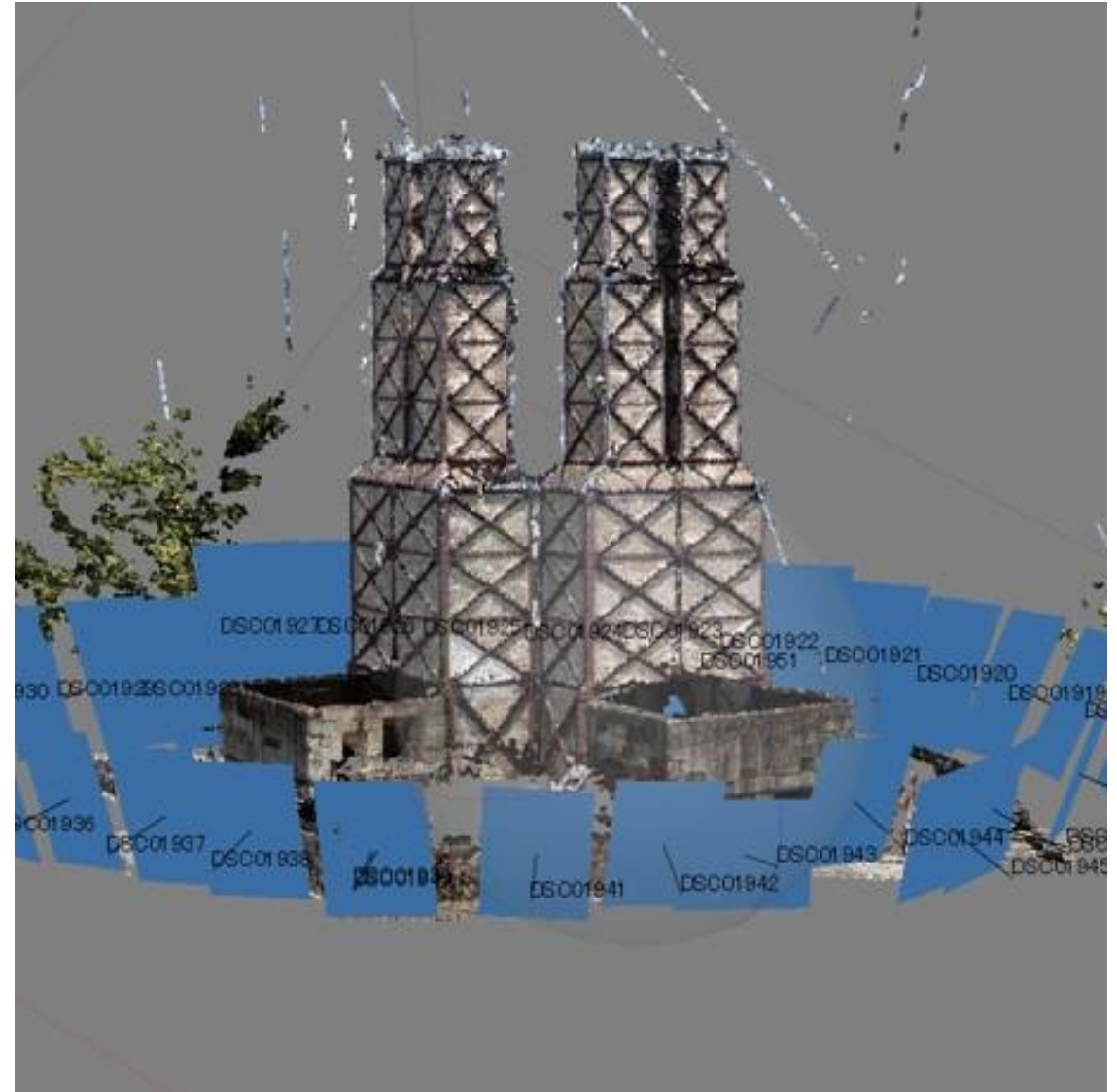


デジタル文化遺産

写真(約60枚)から復元

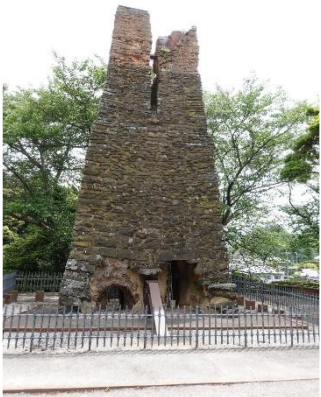






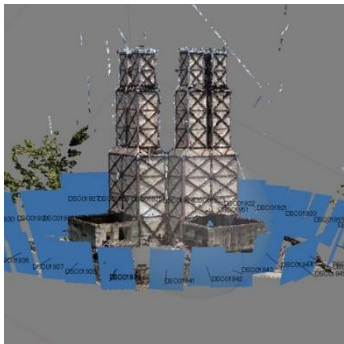
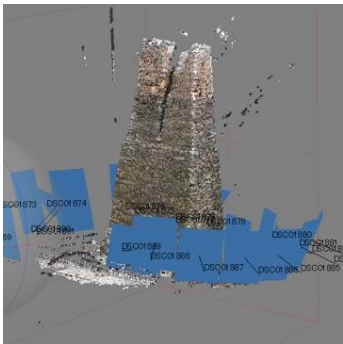
世界産業遺産

萩反射炉 (1855)



- ただ二つの現存する反射炉
- 劣化対応・コンテンツ化 → 簡便なデジタル保存

写真測量による三次元復元



- 特徴点マッチングによりカメラ位置を計算 (SfM)
- ステレオ視により密な三次元形状を復元 (MVS)

形状比較

どちらも佐賀藩の反射炉(現存せず)から技術許与を受けたとされ、類似している可能性がある

スケール・位置合わせ (ICP)



高さで正規化 底面で正規化

- 萩: 2つの台形状ブロックで構成
- 葦山: 3つの長方形ブロックで構成

断面解析



- 萩: 長辺に凹み
- 葦山: 鉄骨部分に膨らみ

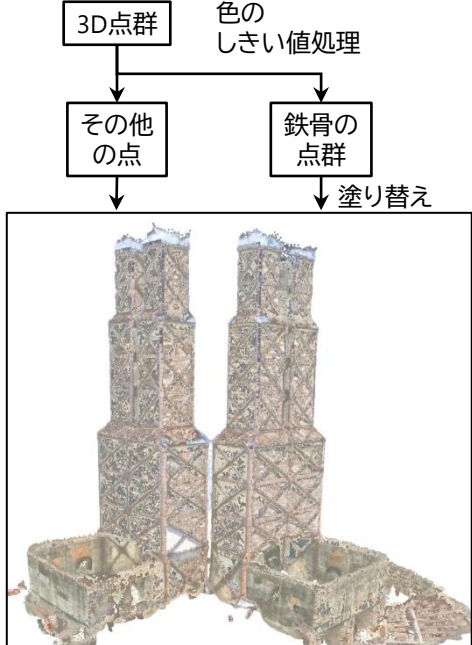
創建時の姿の仮想復元

- 葦山反射炉にあるX字型の鉄骨は、耐震補強のために追加された
- 表面も本来は漆喰
- 現在の見え方は、創建時と著しく異なる
- 遺産の正しい理解

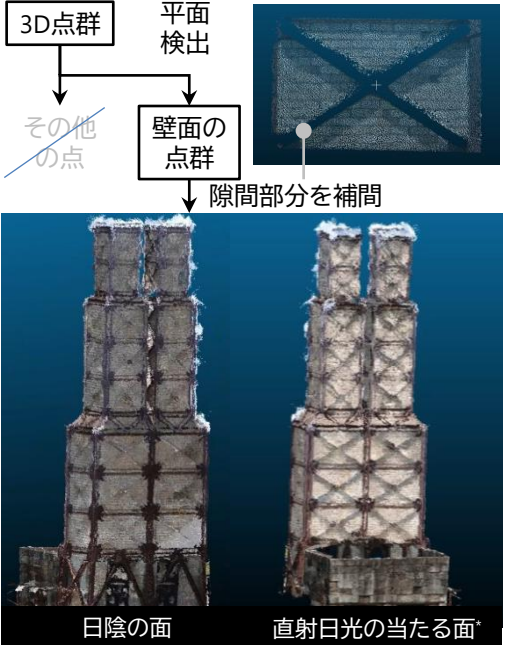
再現方法の検討

	壁面	鉄骨	現実の実現性
ケース (1)	そのまま	レンガ色に塗り替え	可能性あり
ケース (2)	そのまま	除去	困難
ケース (3)	漆喰復元	レンガ色に塗り替え	?
ケース (4)	漆喰復元	除去	困難

ケース (1)



ケース (2)



*既知の問題点: 壁面に落ちた鉄骨の影は修正できていない

今後

点群の補間精度向上、照明条件の考慮、漆喰部分の仮装復元、VRコンテンツ化 など