

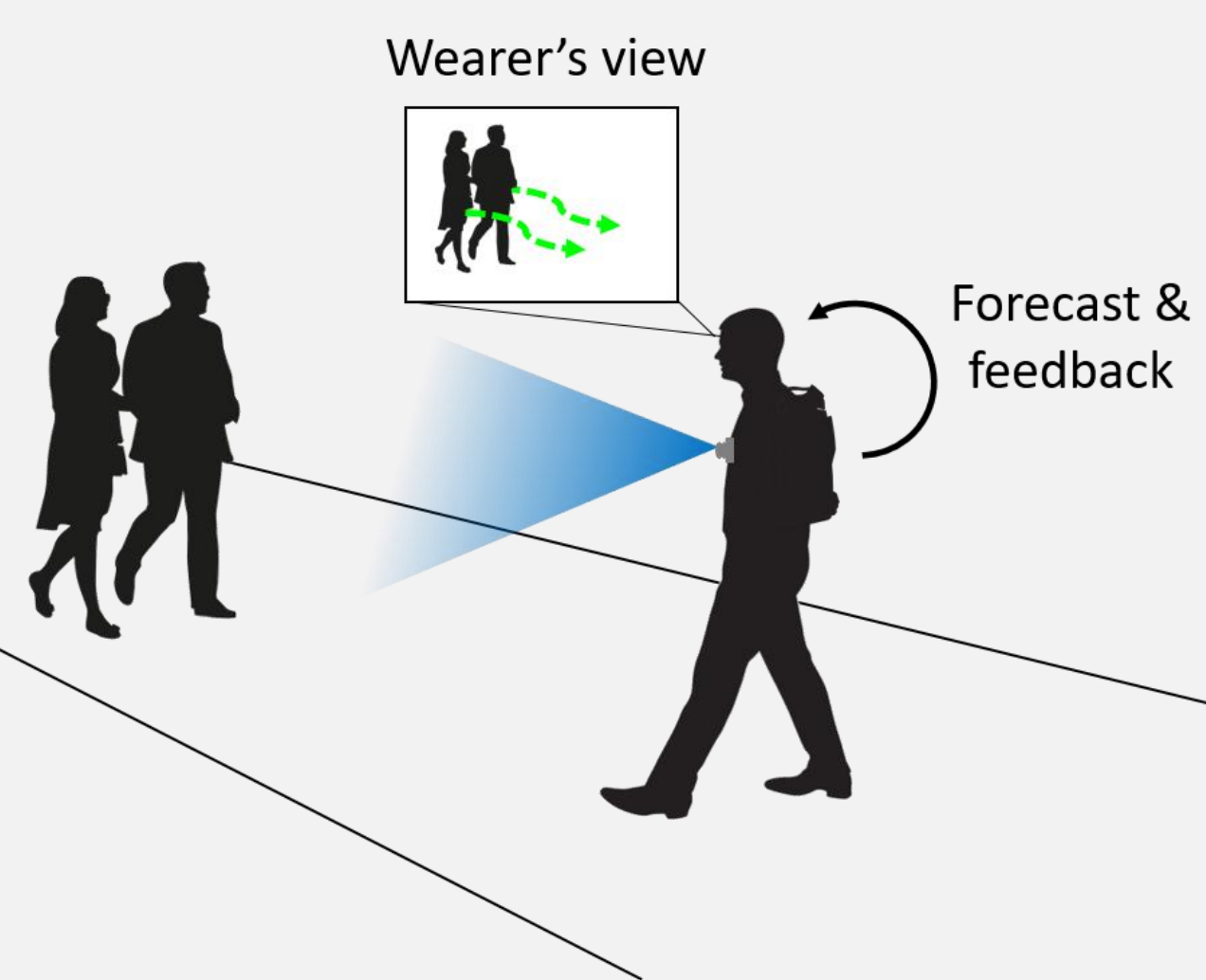
## 問題：人物位置予測

- 一人称視点映像中に映る人物の位置予測
- 「将来の装着者の視野」から見える2次元位置を予測



## 研究目的

- 胸部に装着したカメラからの人物位置のリアルタイム予測
- 予測結果の装着者への呈示
- 動的に変化する環境における装着者行動支援の実現



## アプローチと貢献

- 頭部装着のスマートグラスより予測結果を確認できるバックパック型ウェアラブルシステム
- 一人称視点映像用の人物位置予測手法 [Yagi+, CVPR'18] をノートPC上でリアルタイムで推論するための高速化
- 10フレーム/秒での予測結果呈示
- 実際的な状況に対応するためのデータバイアスの解消

## 関連研究

- 一人称視点映像における人物位置予測 [Yagi+, CVPR'18]
  - オフライン実装、リアルタイム性を追求せず
- 警告音による視覚障害者の衝突回避支援 [Kayukawa+, CHI'19]
  - 歩行者の接近を予測し、警告音で周辺歩行者に回避を要求

## 実験結果

### データバイアスの解消

- 先行研究のデータでは装着者および対象人物が共に歩行→装着者または対象人物が静止している場合に対応できず
- 装着者および対象人物が静止している映像を新たに収集

両者移動



装着者静止



両者静止



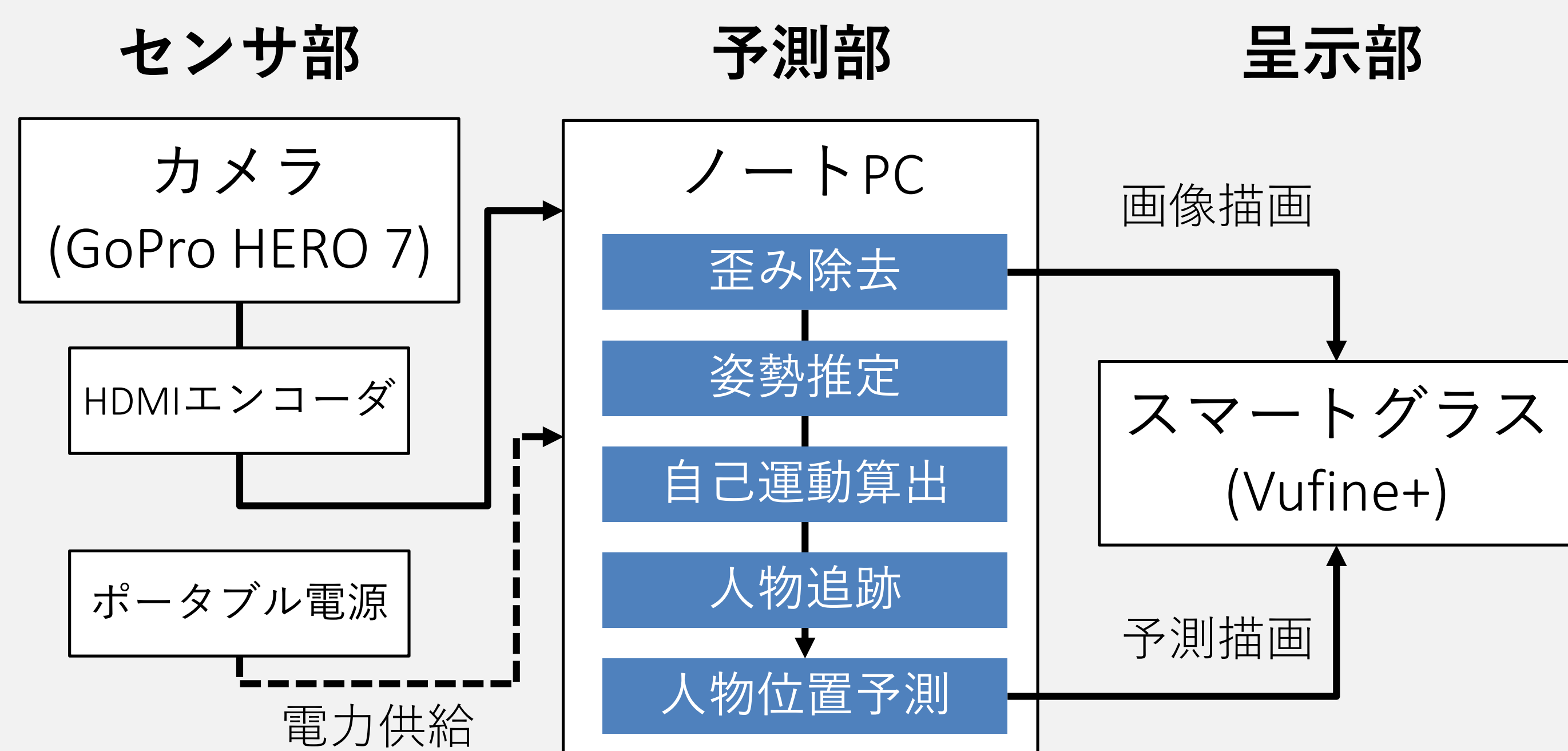
### 追加データの効果

- 装着者が静止している場合の予測誤差が大幅に改善
- 1秒後の最終予測誤差 (予測-正解間のL2距離)

手法 (単位: 画像幅*に 対する%)	評価 データ	追加 データ	歩行方向			
			Toward	Away	Across	Average
位置 + スケール	全て	なし	8.27	6.04	7.03	6.35
位置 + スケール	全て	あり	7.87	6.01	6.49	6.26
フルモデル	全て	なし	7.06	5.47	6.00	5.68
フルモデル	全て	あり	<b>6.61</b>	<b>5.39</b>	<b>5.60</b>	<b>5.51</b>
フルモデル	装着者静止	なし	5.72	3.04	5.31	4.09
フルモデル	装着者静止	あり	<b>3.92</b>	<b>2.32</b>	<b>3.47</b>	<b>3.00</b>
フルモデル	両者静止	なし	<b>2.45</b>	2.87	2.75	2.50
フルモデル	両者静止	あり	2.82	<b>2.18</b>	<b>1.75</b>	<b>2.18</b>

※水平視野角は約130°

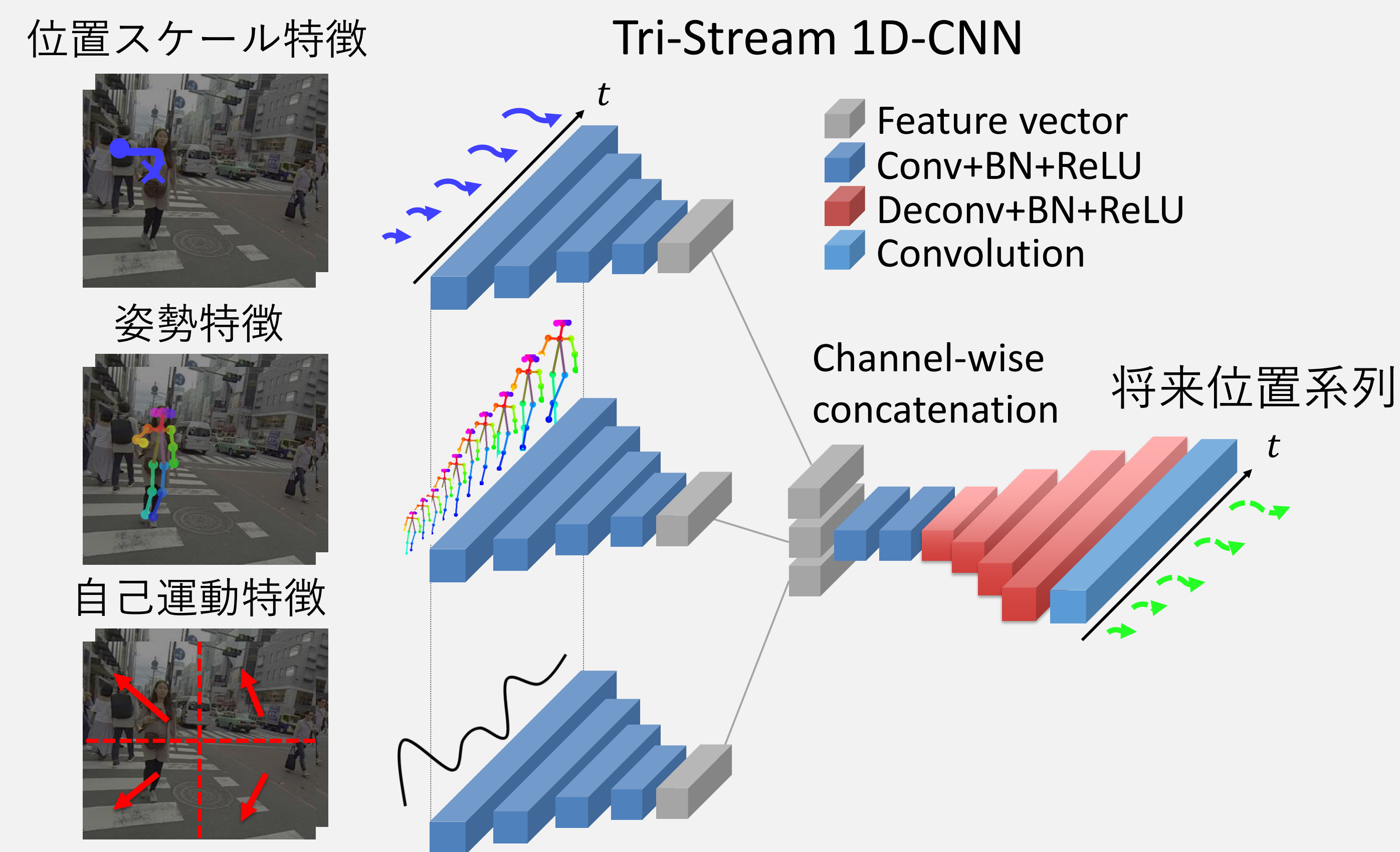
## 提案システム



- 画像取得/画像描画/予測描画を別々のスレッドに分割
- 映像取得遅延を最小化しつつオーバーヘッドを削減

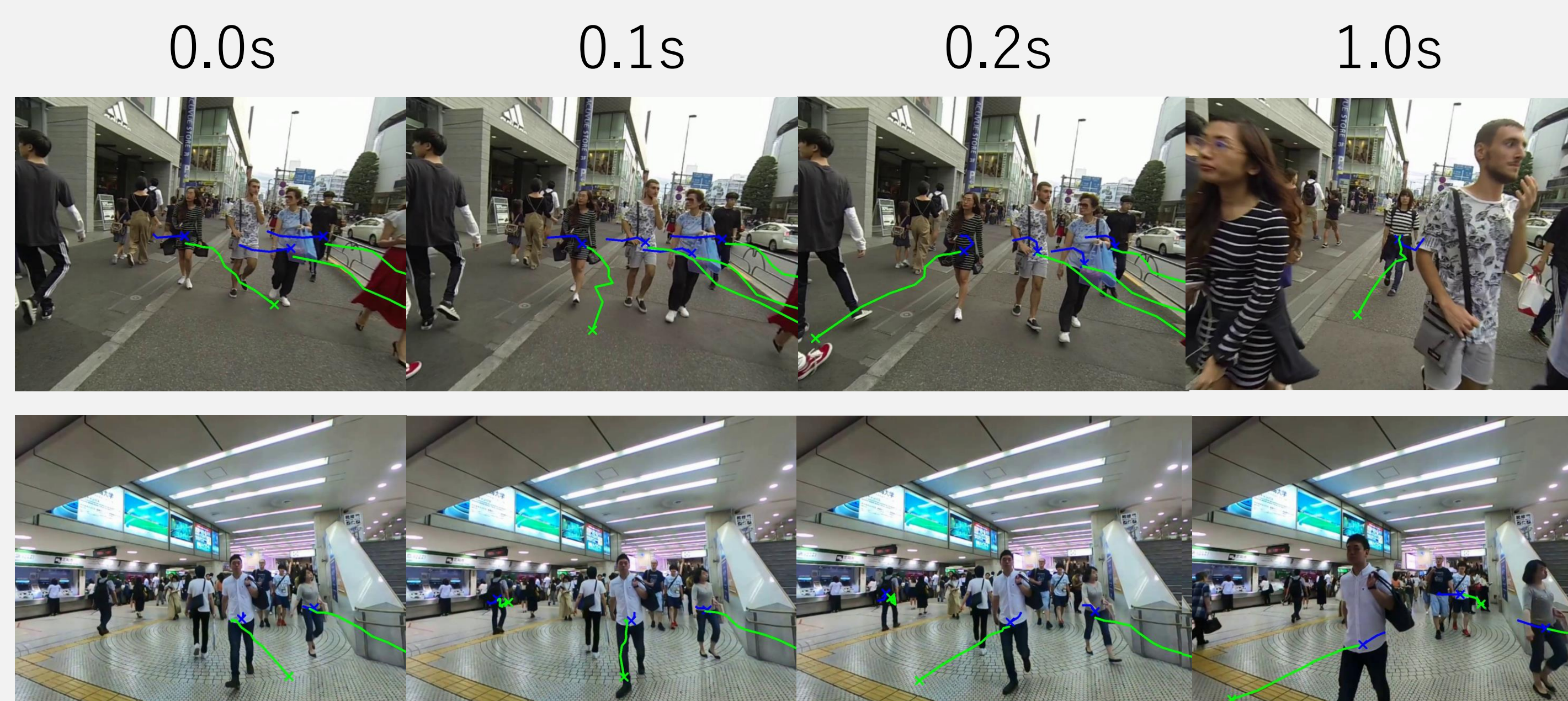
## 予測アルゴリズム

- 位置スケール、姿勢および自己運動 [Yagi+, CVPR'18] を使用
- 人物追跡にはカルマンフィルタ [Bewley+, ICIP'16] を使用
- 自己運動特徴を粗いグリッド状のフロー (LK法) に置換



## 予測結果

- 下記の映像が実際にスマートグラスに表示される:



## 推論速度

- Core i7 8750H 2.2GHz (6コア) + GeForce RTX2070
- 追跡人数が2人以下の場合に設計速度 (10FPS) を達成

予測人数毎の推論速度 (フレーム/秒)

予測人数	1	2	3	4	5	10
	12.7	11.2	9.6	9.0	8.6	6.5

## 今後の展望

- 電力効率の向上
- カメラの装着位置ずれを吸収するための個人適応手法
- 視覚障害者ナビゲーションのための非視覚的情報呈示