

視覚障害者歩行支援のための 単眼カメラを用いた歩行者検出システム

- A Monocular Pedestrian Detection System for Visually Impaired -

横浜市立大学大学院 知覚情報科学研究室
岸野 嵩久, Ruggero Micheletto

研究背景

視覚障害者の自立歩行を支援することは
福祉的な意義が大きい

自立歩行支援システム

- 歩行中の安全確保
- 経路ナビゲーション

- 歩行者との衝突は大きな
ストレスとなっている
- 研究開発例が少ない

普及の容易さから

モバイルデバイスを用いて携帯可能なシステムの実現を目指す

研究目的

単眼カメラの映像から

リアルタイムに前方の歩行者を検出
距離を推定して衝突の危険を通知するシステム



システム概要

映像から人を検出して距離を推定する

IR距離センサやマルチカメラを用いることなく
単眼カメラの映像のみから歩行者を検出して距離を推定する

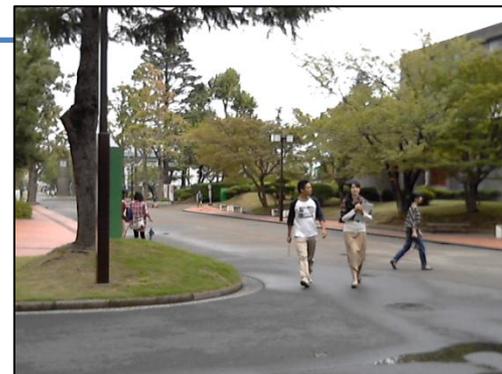
歩行者検出：

- 人物の姿勢や背景変化に強固な検出
- 背景差分によらない人検出

➡ HOG) 特徴量に基づくリアルタイム検出

距離推定：

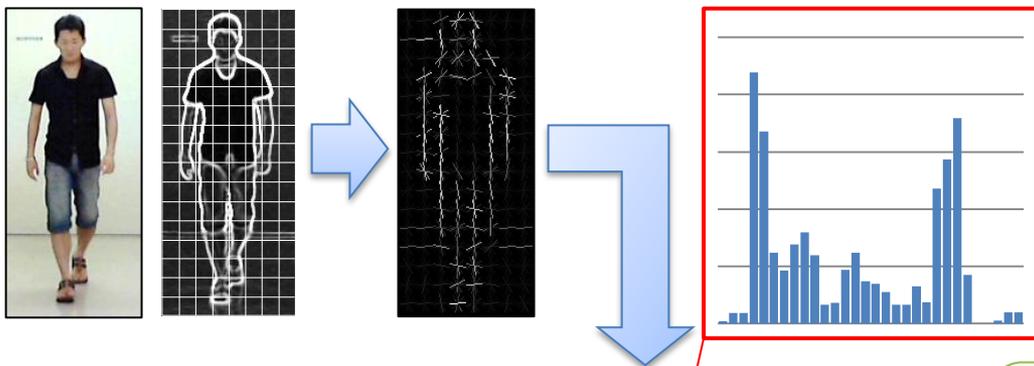
- 投影法を用いて単眼映像での距離推定を行う



歩行者検出法

Histograms of Oriented Gradients (HOG)

人物の輪郭線を方向をヒストグラム化した特徴量
局所的に勾配強度の正規化をすることで照明のバラつきに強固

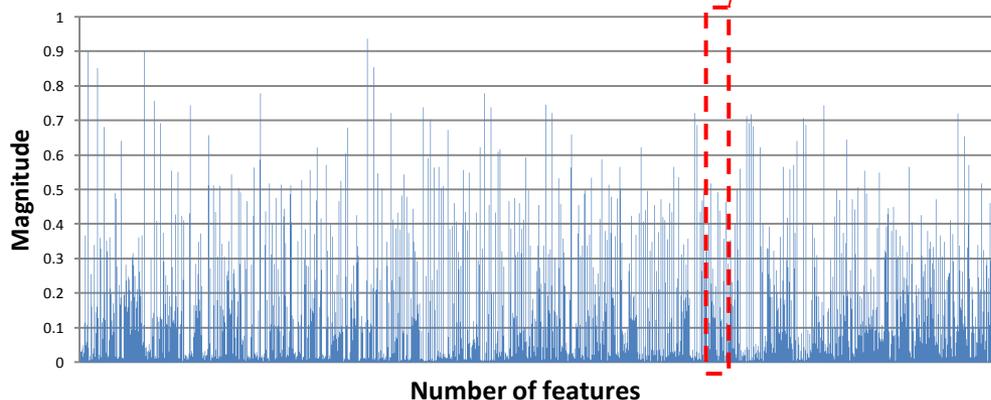


勾配強度：

$$m(x, y) = \sqrt{f_x(x, y)^2 + f_y(x, y)^2}$$

勾配方向：

$$\theta(x, y) = \tan^{-1} \frac{f_y(x, y)}{f_x(x, y)}$$



Real AdaBoostを用いた学習

人モデルと背景モデルの特徴量を比較
識別に有効な要素を学習・選出

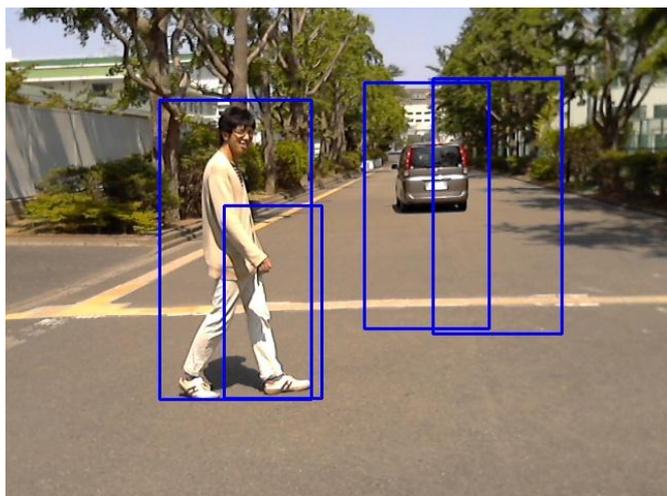
➡ 識別器の構築

識別器が画像上を走査して
歩行者を検出する

歩行者検出法

通常のHOG人物モデルの課題点

背景差分等の前処理がない場合には背景と人物のしい別が困難



➡ 複雑な背景環境では誤検出が多い

アプリケーション実装のためには背景識別をより高める必要がある

➡ 人の特徴をより詳細に捉えることが求められる

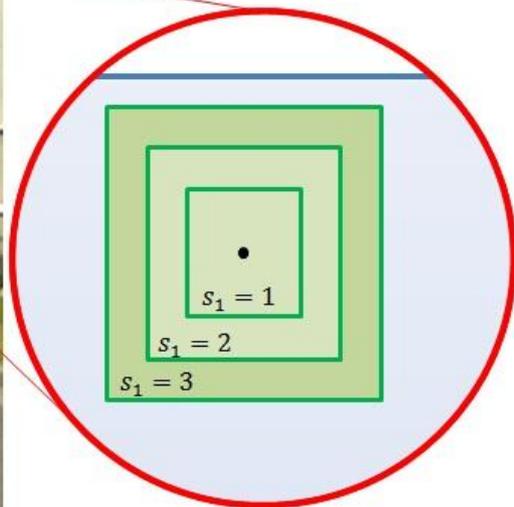
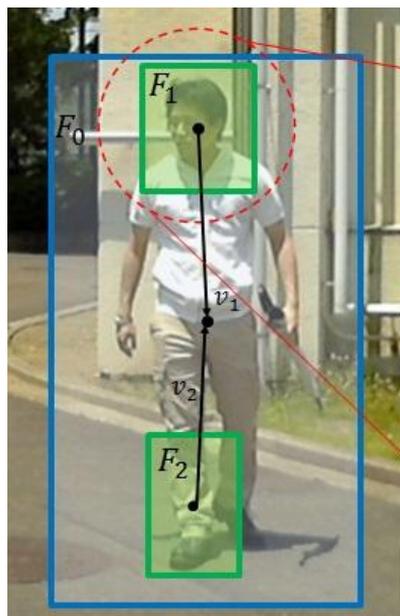
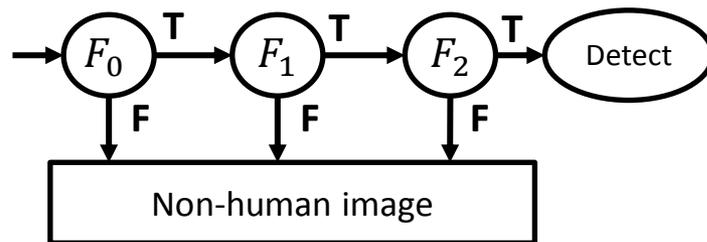
歩行者検出法

Multiple parts HOG detector

人物モデルを複数のパーツから構成するcascade型の検出器
複数のHOG特徴量識別器を組み合わせて人物をより詳細に表現

ベースフィルタ F_0 (全身)

パートモデル P_i (部位)



$$P_i = (F_i, v_i, s_i) \begin{cases} F_i : \text{パートフィルタ} \\ v_i : \text{位置} & s_i : \text{スケール} \end{cases}$$

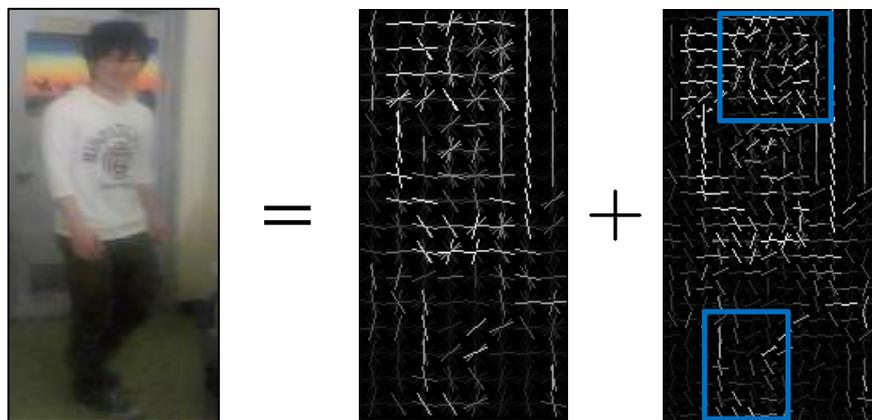
各パートモデルは可変であるため
姿勢変化による影響にも強固になる

歩行者検出法

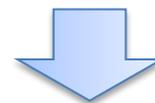
Multiple parts HOG detector

2種類の解像度のHOG特徴量を使用

→ 効率的により細かい特徴を表現できる



全身：姿勢による形の変化が大きい
部位：識別には詳細な特徴が必要



全身：低解像度HOG特徴量
部位：高解像度HOG特徴量

検出率自体を下げることなく

人物の詳細な特徴をより捉えることが出来る

歩行者検出法

検出結果

精度評価のために人物・非人物画像データへ検出器を適用

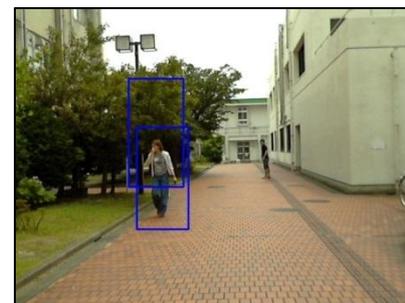
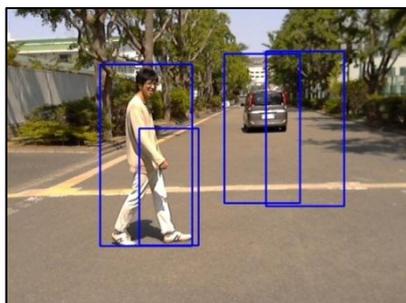
	precision	miss	false
HOG detector	91.90	8.10	25.21
Multiple parts HOG detector	88.70	11.30	6.29

実験・実環境共に誤検出を削減

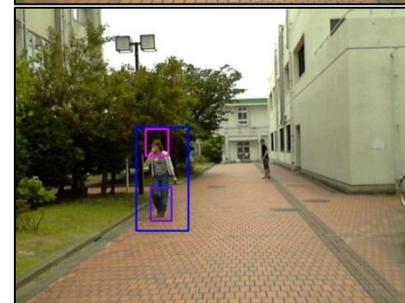
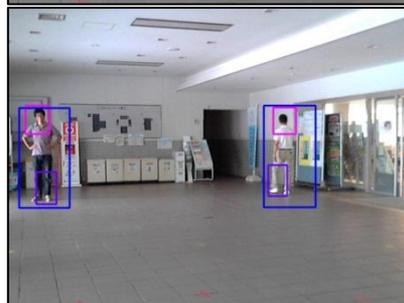
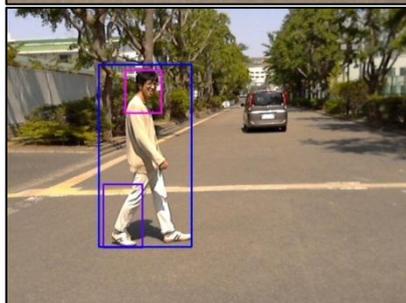
高精度の反面

1フレーム処理に平均1807.22[msec]

(a) Conventional HOG detector



(b) Multiple parts HOG detector



歩行者検出法

Particle Filterによる走査域特定

画像の走査を効率的に行い処理時間を削減

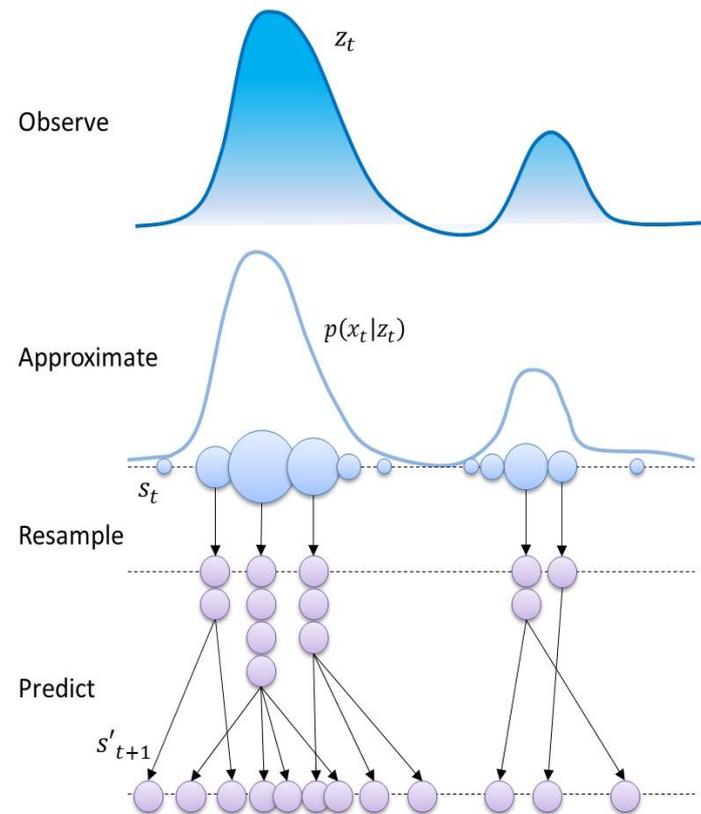
Particle Filter :

観測データを粒子サンプルで近似表現

尤度分布に比例した重みが各粒子に与えて
復元抽出することで対象を追跡

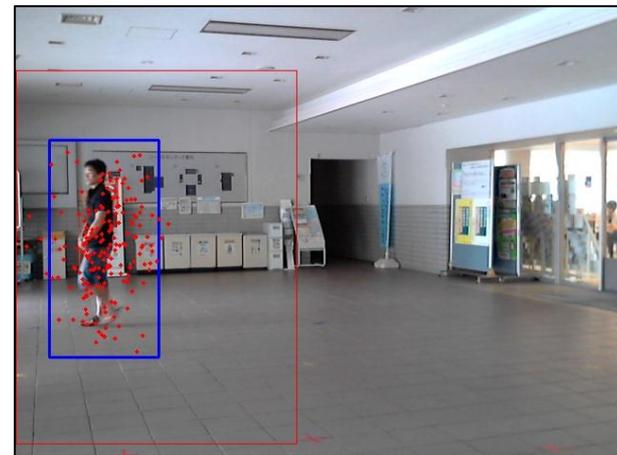
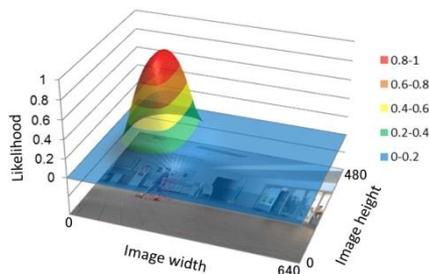
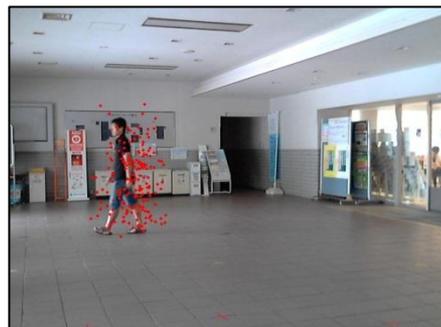
遷移モデルに基づいて各粒子が移動して
次フレームでの対象の位置を予測する

粒子群の重心と分散度から
次フレームの走査範囲と位置を決定



歩行者検出法

追跡結果(仮) :



検出器の出力が高い付近に粒子が散布

計測実験(仮) :

2つのシーンを対象に走査域特定を評価

➡ 処理時間を約67%短縮

リアルタイムの動作のために
走査域特定は有効であった

1フレームに対する平均処理時間(msec)

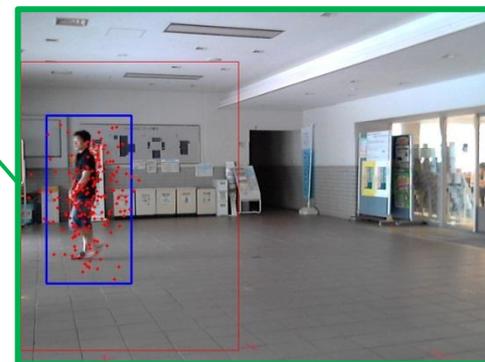
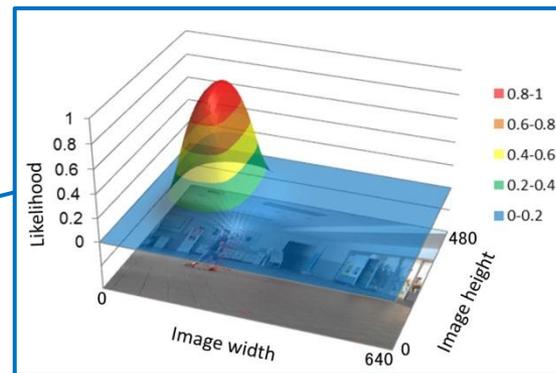
		Time(msec)
Multiple parts HOG	Scene1	1689.92
	Scene2	1775.37
	Average	1807.22
Multiple parts HOG + Particle Filter	Scene1	680.51
	Scene2	545.38
	Average	612.95

歩行者検出法

追跡結果(仮) :



検出器の出力が高い付近に粒子が散布
予測走査領域の特定も対象を補足



計測実験(仮) :

2つのシーンを対象に走査域特定を評価

➡ 処理時間を約67%短縮

リアルタイムの動作のために
走査域特定は有効であった

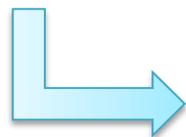
1フレームに対する平均処理時間(msec)

		Time(msec)
Multiple parts HOG	Scene1	1689.92
	Scene2	1775.37
	Average	1807.22
Multiple parts HOG + Particle Filter	Scene1	680.51
	Scene2	545.38
	Average	612.95

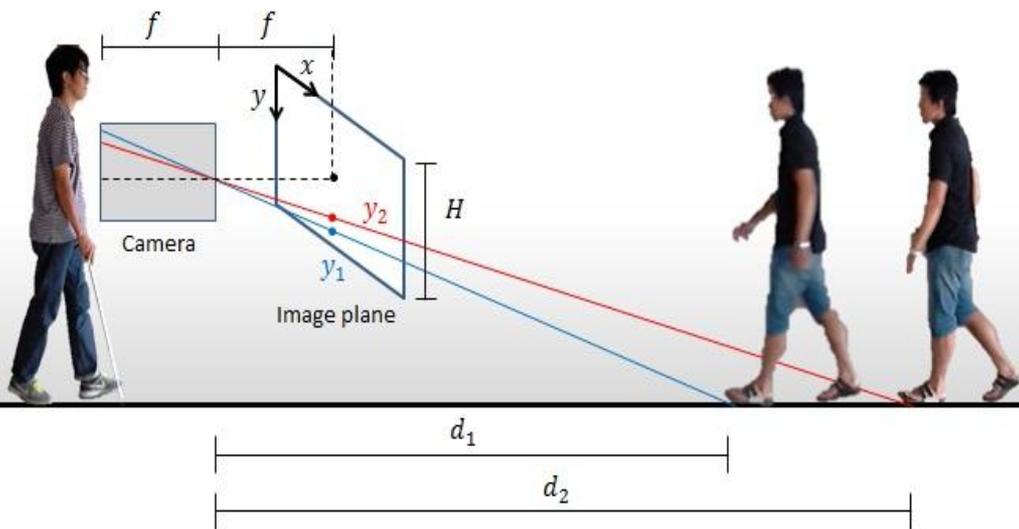
距離推定法

単眼カメラで距離を推定するには投影法を利用する

投影法 - 見た目の大きさの変化
- 位置の変化



身長による個人差がない
カメラ自体の傾き考慮により拡張性がある()



距離算手法：

地面との接地点(足)の座標に基づき歩行者の距離を算出する

$$d = fh \frac{2}{2y - H}$$

(h :カメラの高さ)

距離推定法

接近の程度を判断する目安として
距離を3段階のレベルに区分

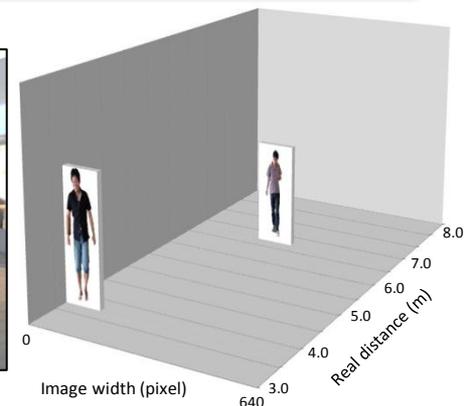
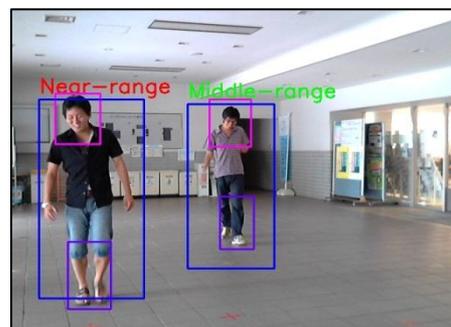
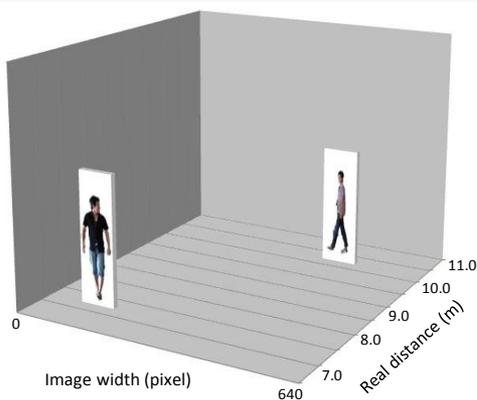
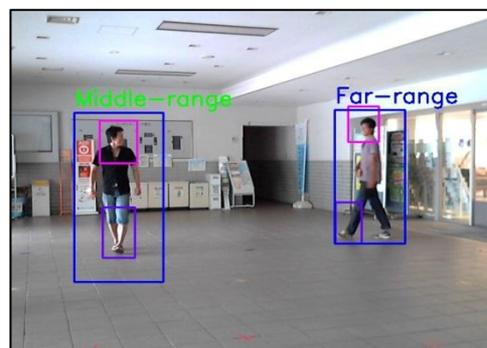
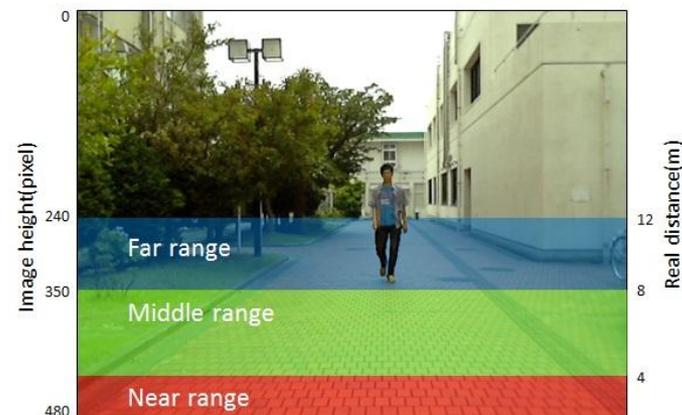
評価実験：

予備実験から $f = 540$ [pixel] と設定

カメラを $h = 1.40$ [m] の高さで地面と水平に固定

結果：

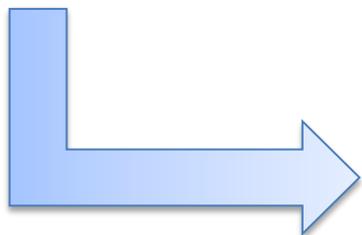
歩行者の存在する距離を推定しそのレベルを正確に表現



まとめ

視覚障害者の歩行を支援するシステムのために
カメラ映像からの歩行者の検出とその距離の推定を行うシステムを製作

- 複数の解像度と部位のHOG識別器を組み合わせ
従来のHOG検出器よりも誤検出を18.9%削減
- 時系列フィルタリングによる次フレームの
走査領域の特定によって処理時間を約67%短縮
- 投影法を利用した距離推定により
単眼カメラ映像でも接近の程度を正確に分別



今後の展望

- モバイルデバイスへの実装と傾きの考慮
- 処理時間短縮のためのメモリ量の削減

ご清聴ありがとうございました