

深度カメラを用いた近接者検知システムの構築に関する研究

A Study on the Construction of Nearby Person Detection System using Depth Camera

○高橋 遥平¹, 関 弘翔², 泉 隆², 香取 照臣²

*Yohei Takahashi¹, Hiroto Seki², Takashi Izumi², Teruomi Katori²

Abstract: Heavy machinery has many blind spot, and difficult to confirm around safety. Therefore, a contact accident with a heavy machinery may occur. In this paper, we studied the construction of nearby person detection system using the depth camera for the purpose of preventing heavy machinery accidents.

1. まえがき

工事現場などでは多数の重機が使用されているが、重機には後方などに死角があるため、人と重機の接触事故を未然に防ぐ意味で、周囲の安全確認は重要である。本研究では、重機事故防止の支援を目的として、重機後方に設置するカメラにより、1~5m の範囲にいる近接者を検知する、近接者検知システムを検討する。

本稿では、深度カメラを用いた近接者検知システム及び近接者検知手法について報告する。

2. 深度カメラ

重機後方に設置するカメラとして深度カメラの「Intel® RealSense™ Depth Camera D435」^[1]を用いる。この深度カメラからは、2つの赤外線画像からステレオマッチングにより距離情報を表す深度マップが取得できる。深度カメラの仕様を Table 1 に示す。

Table 1. Specifications of Depth Camera D435^[1]

Depth Technology		Active IR stereo
Depth Sensor	Depth Distance	0.2m ~ 10m
	Horizontal×Vertically	91.2°× 65.5°
	Frame Rate	Up to 90 fps
RGB Sensor	Horizontal×Vertically	69.4°× 42.5°
IR Projector	Horizontal×Vertically	100.4°× 69°

3. 近接者検知システムの構築

本研究の近接者検知システムは、人検出器を用いた人検知と、深度カメラを利用した距離推定からなる。

4. カメラ設置角度を考慮した距離補正

本システムは、カメラを重機の運転席付近など、人よりも高い位置に設置するため、設置角度の影響でカメラ座標系とワールド座標系が一致せず、人までの距離推定が正しく行えない。そのため、カメラ座標系における距離 Z [m]、高さ Y [m]を用いて、式(1)により設置角度 θ を考慮したワールド座標系における距離 Z' [m]を算出する。

$$Z' = Z \cos\theta + Y \sin\theta \quad (1)$$

5. 人検出器の構築

深度カメラからはカラー画像とグレースケールの赤外線画像、及び深度マップが取得できる。そこで本研究では、赤外線画像から画像特徴量を、深度マップから距離特徴量をそれぞれ多数抽出し、それらから人検出に有効な特徴量を Real AdaBoost^[5]により選択的に学習し、組み合わせることで高精度な人検出器の構築を検討する。

5.1 特徴量の抽出

深度カメラから取得する赤外線画像及び距離情報から、人と関連のある特徴量を抽出する。本研究では画像特徴量として HOG 特徴量^[2]、距離特徴量として距離ヒストグラム特徴量^[3]の抽出を行う。

5.1.1 HOG 特徴量^[2]

HOG 特徴量とは、画像の局所領域における輝度の勾配方向に着目した特徴量である。HOG 特徴量の抽出手順を以下に示す。

① 画像の正規化

画像を 64×128 pixel に正規化し、 8×8 pixel ごとに分割することで、 8×16 個のセルを生成する。

② 勾配方向と勾配強度の算出

各画素における輝度の勾配強度 m を式(2)により、輝度の勾配方向 θ を式(3)によりそれぞれ算出する。

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

$$\theta(x, y) = \tan^{-1}(f_x(x, y)/f_y(x, y)) \quad (3)$$

ただし、 $f_x(x, y) = L(x+1, y) - L(x-1, y)$

$$f_y(x, y) = L(x, y+1) - L(x, y-1)$$

③ 勾配方向ヒストグラムの作成

分割したセル単位で、勾配方向ごとに勾配強度を累積し、勾配方向ヒストグラムを作成する。

④ ブロック単位での正規化

③で作成した勾配方向ヒストグラムを、3×3 個のセルからなるブロック単位で正規化し、HOG 特徴量 v を式(4)により算出する。

$$v(n) = \frac{v(n)}{\sqrt{(\sum_{k=1}^{q \times q \times N} v(k)^2) + 1}} \quad (4)$$

n : 特徴量の番号 q : セル数 N : 勾配方向数

5.1.2 距離ヒストグラム特徴量^[3]

距離ヒストグラム特徴量とは、異なる 2 つの局所領域における、距離ヒストグラムの類似度から算出される特徴量である。

① 深度画像の正規化

深度マップを 64×128 pixel に正規化し、8×8 pixel ごとに分割することで、8×16 個のセルを生成する。

② 2つの矩形領域の選択

セルで構成される 2 つの矩形領域を選択し、それぞれの距離ヒストグラムを算出する。

③ Bhattacharyya 係数による類似度の算出

矩形領域内の距離ヒストグラムの総和が 1 となるように正規化し、正規化距離ヒストグラムとする。算出された、 m ビンからなる 2 つの正規化距離ヒストグラム p, q から、Bhattacharyya 係数^[4]の式(5)により類似度 S を算出し、距離ヒストグラム特徴量とする。

$$S = \sum_{n=1}^m \sqrt{p_n q_n} \quad (5)$$

5.2 Real AdaBoost による特徴量の学習^[5]

教師あり機械学習の一種である Real AdaBoost は、多数の特徴量から対象の識別に有効な特徴量を選択でき、識別結果として対象らしさを実数で出力できる。

① 学習サンプル重みの初期化

学習サンプルの重み D を式(6)で均一に初期化する。

$$D_1(i) = \frac{1}{N} \quad (6)$$

i : 任意の学習サンプル N : 学習サンプル数

② 確率密度分布の作成

全ての学習サンプルに対して、 m 番目の弱識別器候補により算出した特徴量を式(7)、式(8)により、 $|J|$ ビン毎にサンプル重みを足し合わせ、確率密度分布 w を作成する。ここで $W_{t,m+}^j$ は正解サンプルを、 $W_{t,m-}^j$ は不正解サンプルを用いて作成した確率密度分布である。

$$W_{t,m+}^j = \sum_{i: j \in J \cap y_i = +1} D_t(i) \quad (7)$$

$$W_{t,m-}^j = \sum_{i: j \in J \cap y_i = -1} D_t(i) \quad (8)$$

j : 特徴量(ビン) y_i : 属性 {1 : Positive, -1 : Negative}
 t : 学習ラウンド m : 弱識別器候補の番号

③ 弱識別器の類似度算出

弱識別器候補に対する 2 つの確率密度分布の類似度を Bhattacharyya 係数により算出する。ここで類似度が低いとき、正解サンプルと不正解サンプルの分布が大きく分離していることを意味する。 m 番目の弱識別器候補の確率密度分布の分離度 $z_{t,m}$ を式(9)に示す。

$$z_{t,m} = 1 - \sum_j \sqrt{W_{t,m+}^j W_{t,m-}^j} \quad (9)$$

④ 分離度が最大となる弱識別器の選択

式(10)より分離度が最大となる弱識別器候補を、現在の学習ラウンド t における弱識別器 $h_t(x)$ とする。

$$h_{t,m} = \arg \max z_{t,m} \quad (10)$$

⑤ 学習サンプルの重みの更新

弱識別器 $h_t(x)$ の出力をもとに、全ての学習サンプルの重みを式(11)より算出し更新する。

$$D_{t+1}(i) = D_t(i) \exp[-y_i h_t(x_i)] \quad (11)$$

⑥ 学習サンプルの重みの正規化

更新された学習サンプルの重みの総和が 1 になるように式(12)を用いて正規化する。

$$D_{t+1}(i) = \frac{D_{t+1}(i)}{\sum_{n=1}^N D_{t+1}(n)} \quad (12)$$

⑦ 強識別器の生成

学習した T 個の弱識別器の線形和が閾値 λ より高ければ人、低ければ非人であることを出力する強識別器 $H(x)$ を、式(13)により構築する。

$$H(x) = \text{sign} \left[\sum_{t=1}^T h_t(x) - \lambda \right] \quad (13)$$

6. まとめ

本稿では、深度カメラを用いた近接者検知システムの構築及び検知手法について検討した。今後は、検討した近接者検知手法の有効性の検証を行う。

参考文献

[1] Intel "Intel® RealSense™ D400 Series(DS5)Product Family Datasheet", https://software.intel.com/sites/default/files/Intel_RealSense_Depth_Cam_D400_Series_Datasheet.pdf (2018-09)
 [2] N.Dalal,B.Triggs:"Histograms of oriented gradients for human detection", IEEE CVPR05, pp. 886-893 (2005)
 [3] 池村 翔・藤吉 弘亘:「距離情報に基づく局所特徴量によるリアルタイム人検出」, 電子情報通信学会論文誌, 93(3), pp.355-364 (2010-03)
 [4]A. Bhattacharyya, "On a measure of divergence between two statistical populations defined by probability distributions," Bull. Calcutta Math. Soc., vol.35, pp.99-109 (1943)
 [5] R. E. Schapire and Y. Singer: "Improved Boosting Algorithms Using Confidence-rated Predictions", Machine Learning, Vol. 37, No. 3, pp. 297-336(1999-12)