

論文

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

池村 翔^{†a)} 藤吉 弘亘^{†b)}

Real-Time Human Detection Using Local Features Based on Depth Information

Sho IKEMURA^{†a)} and Hironobu FUJIYOSHI^{†b)}

あらまし 近年、統計的学習法と局所特徴量を用いた人検出に関する研究が多く取り組まれている。従来の人検出法では、HOG (Histograms of Oriented Gradients) 特徴量のようなこう配情報に基づく特徴量を用いる手法が多いが、人の重なりや複雑な背景に対して人検出が困難となる場合がある。更に、検出時には検出ウィンドウのスケールを変化させながら画像上を複数回にわたりラスラスキャンするため処理コストが高く、リアルタイム処理が困難である。そこで、本論文では TOF カメラから得られる距離情報を用いて人の重なりや複雑なシーンに頑健なリアルタイム人検出手法を提案する。提案手法では、距離画像から二つの局所領域の距離関係をとらえることができる距離ヒストグラム特徴量を抽出する。抽出された特徴量を用いて Real AdaBoost 識別器を構築し、人の識別を行う。検出時には三次元実空間における検出ウィンドウのラスラスキャンをすることで、人の大きさに合わない検出ウィンドウを削除して、高速化を実現する。三次元実空間における Mean-Shift クラスタリングにより人と識別された検出ウィンドウを統合することで、人がどこにいるかを自動的に検出することができる。評価実験の結果、誤検出率 5.0%において検出率 98.9%となり、HOG 特徴量を用いた従来法と比較して 4.9%検出率を向上させることができた。また、提案手法は約 10 fps でリアルタイムに人検出が可能であることを確認した。

キーワード 人検出, TOF カメラ, 距離ヒストグラム特徴量, Real AdaBoost

1. ま え が き

近年、オフィスや家、公共施設などの空間において、その空間内の人の意図を理解し行動を支援する技術の実現が期待されている。このような人を観る技術 [1] を実現するためには、空間内のどこに人がいるかを知る必要があり、高精度かつ高速な人検出の実現が課題となっている。人検出は従来から盛んに研究されており、様々な手法が提案されている [2]~[7]。従来の可視光カメラを用いた人検出手法では、局所特徴量と Boosting による統計的学習からなる手法が多い。局所特徴量にはこう配ベースの特徴量である EOH [2], HOG [3], Edgelet [4] 等が多く用いられており、高精度な人検出を実現できることが報告されている。しかし、可視光カメラ画像によるこう配ベースの特徴量は、複雑な背

景や人の重なりが多い場合、人の形状をとらえることが困難となり、検出精度が低下する場合がある。この問題に対して、Ess らはステレオカメラにより得られる距離情報を用いて混雑したシーンにおける高精度な人検出法を提案した [8]。しかし、ステレオカメラによる距離情報の取得は、カメラ校正やステレオマッチングによる画像間の対応計算が必要であるため、処理コストが高く、リアルタイム検出が困難である。また、従来の人検出手法においても、画像中における人のサイズが未知であるため、検出ウィンドウのスケールを変化させながら繰り返しラスラスキャンを行うことにより、計算コストが増大し、リアルタイム処理が困難となる。

そこで、本論文では TOF (Time of Flight) カメラから得られる距離情報を用いて人の重なりや複雑なシーンに頑健なリアルタイム人検出手法を提案する。提案手法は TOF カメラから得られる距離情報を用いて、局所領域間の距離関係をとらえる距離ヒストグラム特徴量を抽出し、Real AdaBoost により識別器を構築する。構築された識別器を用いて人検出を行い、

[†] 中部大学大学院工学研究科, 春日井市

Dept. of Computer Science, Chubu Univ., 1200 Matsumoto, Kasugai-shi, 487-8501 Japan

a) E-mail: si@vision.cs.chubu.ac.jp

b) E-mail: hf@cs.chubu.ac.jp

三次元実空間における検出ウィンドウのラスタスキャンによる高速化と、オクルージョン領域の考慮による検出精度の向上を実現する。

2. 距離情報に基づく局所特徴量

TOF カメラは、カメラの周囲に付いた LED より照射される赤外光が対象物に反射し、カメラで観測されるまでの時間を計測することにより、物体までの距離を計測するカメラである。本研究では、TOF カメラとして MESA 社の SR-3100 を用いる。SR-3100 は、0.3~7.5 m (分解能: 3 m において 22 mm) までの距離情報をリアルタイムで取得することができるが、屋外での撮影が不可能であり、屋内での利用に限られたカメラである。人を検出する際、TOF カメラより得られる距離情報を用いて人体と背景の距離関係をとらえることが有効であると考えられる。そこで提案手法では、二つの局所領域間の距離分布の類似度から得られる距離ヒストグラム特徴量を提案する。

2.1 距離ヒストグラム特徴量

距離ヒストグラム特徴量は、二つの局所領域から得られる距離ヒストグラムの類似度を特徴量としたものである。図 1 に示すように、距離画像を 8×8 pixel のセルに分割し、セルで構成される二つの方形領域を選択する。選択された二つの方形領域の距離情報から各距離ヒストグラムを算出し、それぞれの距離ヒストグラムの合計が 1 となるように正規化する。算出された m ビンからなる二つの正規化距離ヒストグラム p , q の各ビンを p_n , q_n とし、Bhattacharyya 距離 [9] による類似度 S を算出し、これを距離ヒストグラム特徴量とする。

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

これをすべての方形領域の組合せに対して行うことに

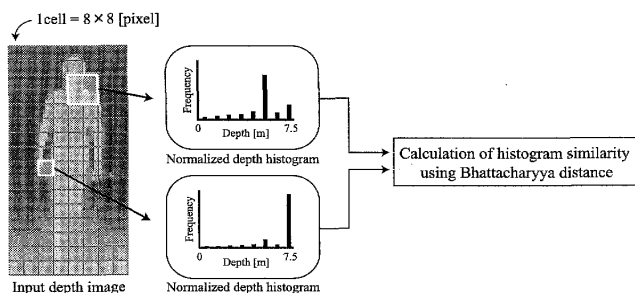


図 1 距離ヒストグラム特徴量の抽出
Fig. 1 Calculation of depth histogram features.

より、距離ヒストグラム特徴量ベクトルを算出する。距離ヒストグラム特徴量は、異なる領域の類似度から得られる特徴量であるため、カメラからの距離に依存しない二つの領域の相対的な距離関係を表現する特徴量となる。

2.2 方形領域サイズの可変

距離ヒストグラム特徴量では、図 2 に示すように距離ヒストグラムを算出する際、式 (2) を用いてヒストグラムの正規化を行うため、大きさの異なる方形領域同士における類似度を算出することが可能となる。ここで、 p_n を距離ヒストグラムの n 番目のビンとすると、正規化距離ヒストグラムの n 番目のビン p'_n は、次式により求めることができる。

$$p'_n = \frac{p_n}{\sum_{i=1}^m p_i} \quad (2)$$

局所特徴量の方形領域のサイズを変化させることで、人の肩や頭部のような一つのセルでは表現できない領域における距離関係をとらえることが可能となる。本手法では検出ウィンドウサイズを 64×128 pixel とするため、 8×16 のセルに分割できる。距離ヒストグラムの算出方形領域を 1×1 から 8×8 セルまでセル単位で変化させることにより、492 個の方形領域が存在する。492 個の方形領域の組合せから距離ヒストグラム特徴量を算出するため、一つの検出ウィンドウ内から $492C_2 = 120,786$ 個の特徴量候補が抽出される。

2.3 Integral Histogram による距離ヒストグラム計算の高速化

特徴量算出時の計算コストを抑えるために、本手法では Integral Histogram [10] を用いて高速に距離ヒストグラムを算出する。まず、距離画像の各画素にお

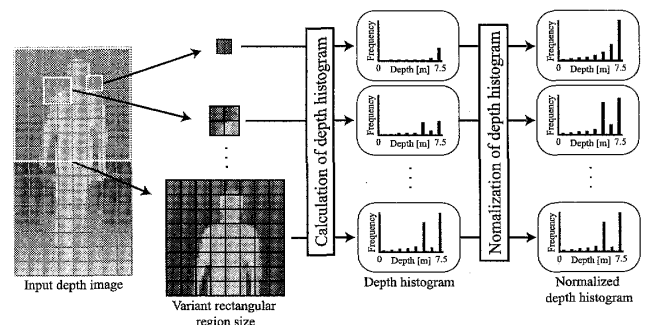


図 2 異なる方形領域サイズにおける距離ヒストグラムの正規化

Fig. 2 Normalization of depth histogram on variant rectangular region size.

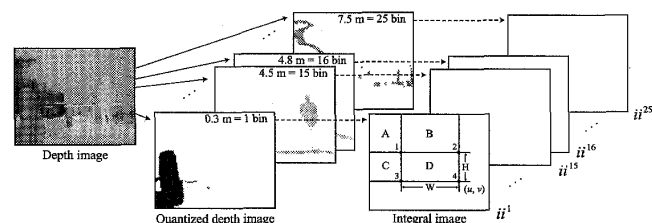


図 3 Integral Histogram の算出

Fig. 3 Calculation of integral histogram.

ける距離を 0.3 m 間隔で量子化する。本研究では 0~7.5 m を 0.3 m 間隔とするため、25 個のビンからなる距離ヒストグラムを算出することになる。次に、図 3 に示すように、ビン n に対応した 25 枚の量子化画像 i^n を作成し、各量子化画像 $i^n(u, v)$ から積分画像 $ii^n(u, v)$ を式 (3)(4) により算出する。

$$s^n(u, v) = s^n(u, v-1) + i^n(u, v) \quad (3)$$

$$ii^n(u, v) = ii^n(u-1, v) + s^n(u, v) \quad (4)$$

ここで、 $s^n(u, v)$ はビン n における行の画素の総和、 $ii^n(u, v)$ は列の s^n の総和を表す。ただし、 $s^n(u, -1) = 0$, $ii^n(-1, v) = 0$ とする。図 3 の D の領域から距離ヒストグラムの n 番目のビン D^n を算出するには、 n 番目の積分画像 ii^n から式 (5) により 4 点から和を求めればよいことになる。

$$D^n = (ii^n(u, v) + ii^n(u-W, v-H)) - (ii^n(u-W, v) + ii^n(u, v-H)) \quad (5)$$

これにより、領域の大きさに関係なく距離ヒストグラムの n 番目のビンの値を高速に求めることが可能となる。

3. 距離情報を用いた人検出

提案手法における人検出の流れを図 4 に示す。提案手法は、まず距離画像上において検出ウィンドウをラスタスキャンし、各検出ウィンドウから距離ヒストグラム特徴量を算出する。次に算出された特徴量のオクルージョン判定を行い、Real AdaBoost を用いて各ウィンドウが人か人以外かの識別を行う。ラスタスキャン後、三次元実空間における Mean-Shift クラスタリングにより人と識別された検出ウィンドウを統合し、人領域を決定する。

3.1 Real AdaBoost による識別器の構築

提案手法では、人の識別に Real AdaBoost [11] を用いる。図 5 に Real AdaBoost による識別器の構築

アルゴリズムを示す。Real AdaBoost は、ポジティブクラスの特徴量とネガティブクラスの特徴量の各次元の確率密度関数から分離度を求め、ポジティブクラスとネガティブクラスを最も分離できる特徴量を弱識別器として選択する。このとき分離度を評価値とするため、実数による識別結果の出力が可能である。学習により選択された弱識別器を $h_t(x)$ とすると、構築される最終識別器 $H(x)$ は式 (6) となる。

$$H(x) = \text{sign} \left(\sum_{t=1}^T h_t(x) \right) \quad (6)$$

3.2 三次元実空間におけるラスタスキャン

提案手法では、図 6 に示すように $y_w = 0$ とした地面上において、 60×180 [cm] の検出ウィンドウを z_w を変化させながら x_w 方向へのラスタスキャンを繰り返すことにより、三次元実空間の地面上をラスタスキャンする。三次元実空間中のラスタスキャンにより得られる検出ウィンドウの三次元座標を式 (7) を用いて画像座標 $[u, v]^T$ に投影し、投影された座標位置に対する領域内の距離情報から特徴量を計算し、識別を行う。

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = P \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (7)$$

$$P = A[R|T] \quad (8)$$

行列 P は透視投影行列であり、カメラキャリブレーションにより得られた内部パラメータ A と、外部パラメータである回転行列 R と平行移動ベクトル T から算出される。TOF カメラは画像座標 (u, v) と対応する $5 \times 4 \times 7.5$ [m] の空間内における世界座標 (x_w, y_w, z_w) を取得することができるため、容易にカメラキャリブレーションを行うことが可能である。ここで、内部パラメータ A は、TOF カメラが CCD センサを用いて赤外光を観測しているため、一般的なカメラと同様である。更に、Mure-Dubois らは公開されている内部パラメータと実際にキャリブレーションを行った結果の比較を行い、内部パラメータが近似していることを確認している [12]。

3.3 オクルージョンに対応した識別

複数の人が重なるような混雑したシーンにおいては、対象とする人領域にオクルージョンが発生する場合

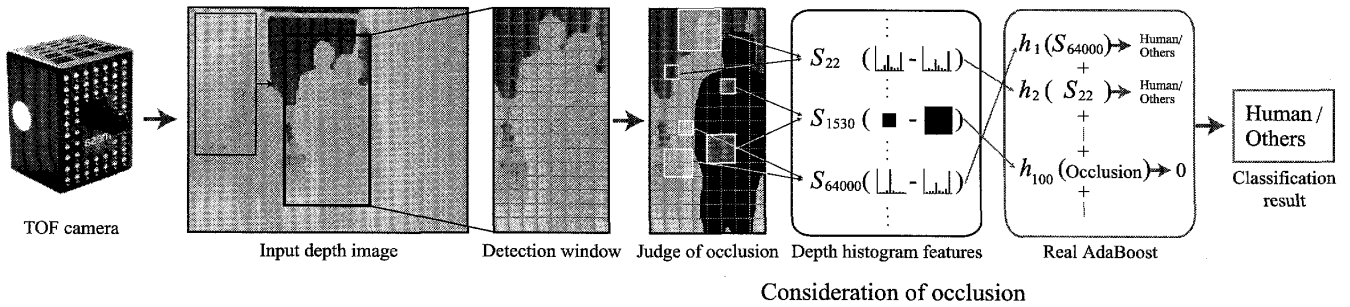


図 4 距離情報を用いた人検出の流れ

Fig. 4 Flow of human detection using depth information.

1. Training sample $I(x, y)$ $x = 1, 2, \dots, N$, $y \in \{+1, -1\}$
2. Initialize weight of training sample $D_1 = \frac{1}{N}$
For $t = 1, 2, \dots, T$ Number of training rounds
 3. For $m = 1, 2, \dots, M$ Number of weak classifiers
 - 3.1. Update probability density distribution

$$W_+^j = \sum_{i: j \in J_t^+} D_t(i)$$

$$W_-^j = \sum_{i: j \in J_t^-} D_t(i)$$
 - 3.2. Weak classifier candidate

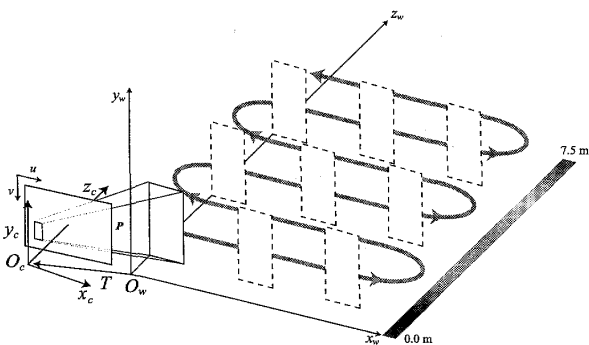
$$h(x) = \frac{1}{2} \ln \frac{W_+^j + \epsilon}{W_-^j + \epsilon}$$
 - 3.3. Calculate evaluated value

$$Z_m = 2 \sum_j \sqrt{W_+^j W_-^j}$$
 4. Select best weak classifier

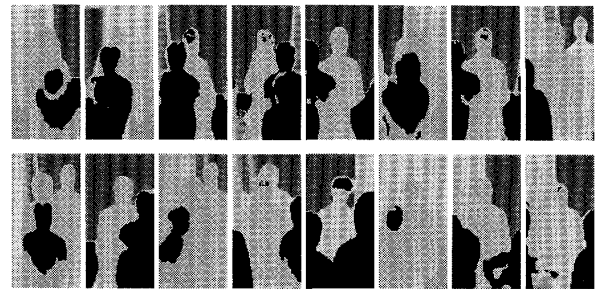
$$h_t = \arg \min Z_{t,m}$$
 5. Update weight of training sample

$$D_{t+1}(i) = D_t(i) \exp[-\gamma_i h_t(x_i)]$$
6. Final classifier

$$H(x) = \text{sign}\left(\sum_{t=1}^T h_t(x)\right)$$

図 5 Real AdaBoost のアルゴリズム
Fig. 5 Real AdaBoost algorithm.図 6 三次元実空間におけるラスタスキャン
Fig. 6 Raster scan in 3D space.

がある。オクルージョン領域から抽出される距離情報は、弱識別器の誤った応答を出力する原因となる。したがって、このようなオクルージョン領域をとらえる

図 7 抽出したオクルージョン領域例（黒色の領域）
Fig. 7 Example of occluded region.

弱識別器の出力は、そのまま最終識別器に統合しないようにする。提案手法では検出ウィンドウを三次元実空間においてラスタスキャンするため、ウィンドウの世界座標が既知である。そこで、TOF カメラから得られる距離情報と比較することで、カメラに対して検出ウィンドウより手前に存在する物体領域をオクルージョンとして判別し、識別に利用する。本研究では、人の自然な重なりが存在するとき、前面に存在する人と背後に存在する人の距離を 0.3 m 以上と仮定し、対象とする人の検出ウィンドウより 0.3 m 以上手前に存在する物体領域をオクルージョンとした。

3.3.1 オクルージョン領域の抽出

三次元実空間におけるラスタスキャン時の検出ウィンドウの距離 z_w を用いて、オクルージョン判定のしきい値を決定する。検出ウィンドウ内の各画素を (u, v) 、そのデプスマップを $d(u, v)$ とすると、各座標におけるオクルージョンラベル $O(u, v)$ は式 (9) で表される。

$$O(u, v) = \begin{cases} 1 & \text{if } d(u, v) < z_w - 0.3 \text{ m} \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

抽出されたオクルージョン領域（黒色の領域）を図 7 に示す。

3.3.2 オクルージョン領域を考慮した識別

t 番目の弱識別器 $h_t(x)$ の対象とする方形領域 B_t 内

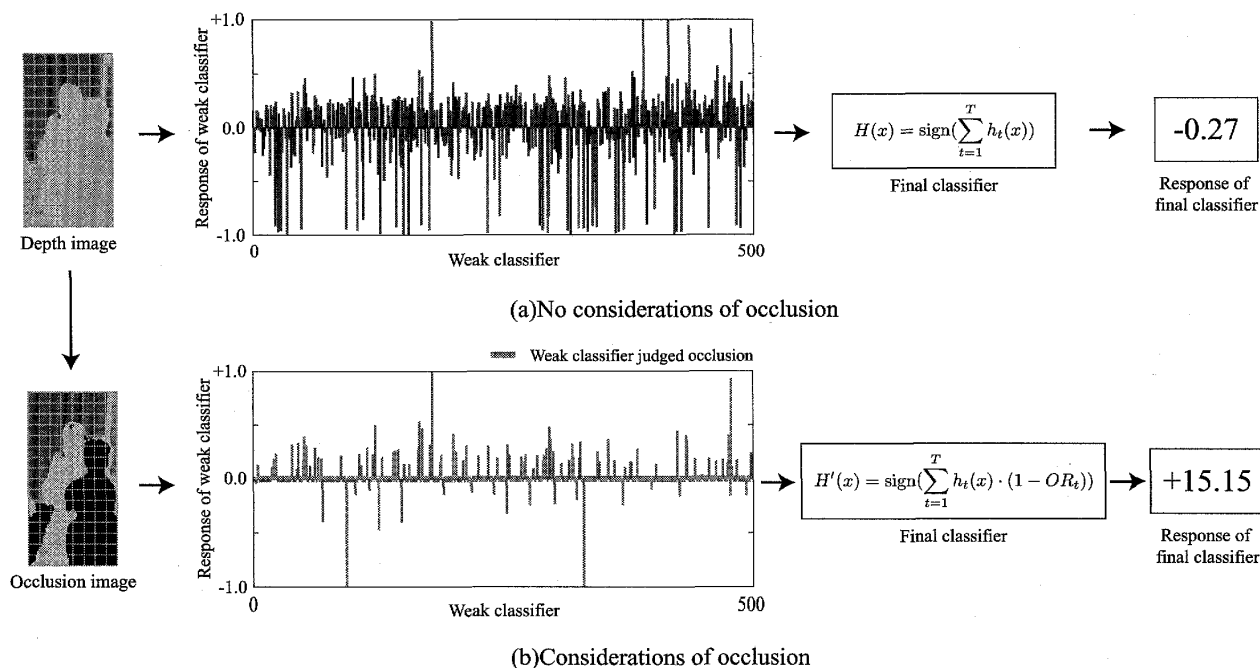


図 8 オクルージョン領域を考慮した識別例
Fig. 8 Example of classification with occlusion consideration.

に存在するオクルージョン領域の割合をオクルージョン率 OR_t とし、式 (10) より求める。

$$OR_t = \frac{1}{B_t} \sum_{(u,v) \in B_t} O(u,v) \quad (10)$$

算出されたオクルージョン率 OR_t を用いて、オクルージョン領域を考慮した最終識別器 $H'(x)$ は式 (11) で表される。

$$H'(x) = \text{sign} \left(\sum_{t=1}^T h_t(x) \cdot (1 - OR_t) \right) \quad (11)$$

図 8(a) に示すように、オクルージョン領域を考慮せず最終識別器により識別を行うと、多くの弱識別器の出力がマイナスとなり、その結果人以外と誤識別される。一方 (b) は、オクルージョン率を考慮して、最終識別器の出力を求めるため、人と正しい識別が可能となる例である。

3.4 三次元実空間における Mean-Shift クラスタリング

従来の可視光カメラによる人検出 [7] では、人と識別された検出ウィンドウを Mean-Shift クラスタリング [13] により統合し検出結果とする。しかし、画像空間における Mean-Shift クラスタリングのみでは、図 9(b)(d) に示すように人の重なりのある場合に検出ウィンドウが誤って統合されることがある。そこで

本手法では、図 9(c)(e) に示すように三次元実空間における Mean-Shift クラスタリングを行う。これにより、人の重なりがある場合でも距離情報によりクラスターを分離することが可能であるため、検出ウィンドウの誤った統合を抑制することができる。

三次元 Mean-Shift クラスタリングは式 (12) により Mean-Shift Vector $m(\mathbf{x})$ を算出する。ここで、 $\mathbf{x}$ は移動している検出ウィンドウの中心座標、 $\mathbf{x}_i$ は各データの三次元座標を表す。 k はカーネル関数、 h はバンド幅であり、本研究では $h = 0.3 \text{ m}$ とした。

$$m(\mathbf{x}) = \frac{\sum_{i=1}^n \mathbf{x}_i k \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right)}{\sum_{i=1}^n k \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right)} - \mathbf{x} \quad (12)$$

4. 識別実験による提案手法の評価

提案手法の有効性を確認するため、評価実験を行う。

4.1 データベース

データベースには TOF カメラで撮影したシーケンスを用いる。屋内で TOF カメラを約 2.5 m の高さに設置し、人の歩行シーンと複数の人が重なり合うシーンを対象とした。撮影した屋内のシーケンスから切り出した学習用ポジティブサンプル 1346 枚、学習用ネガティブサンプル 10000 枚を用いる。また、評価

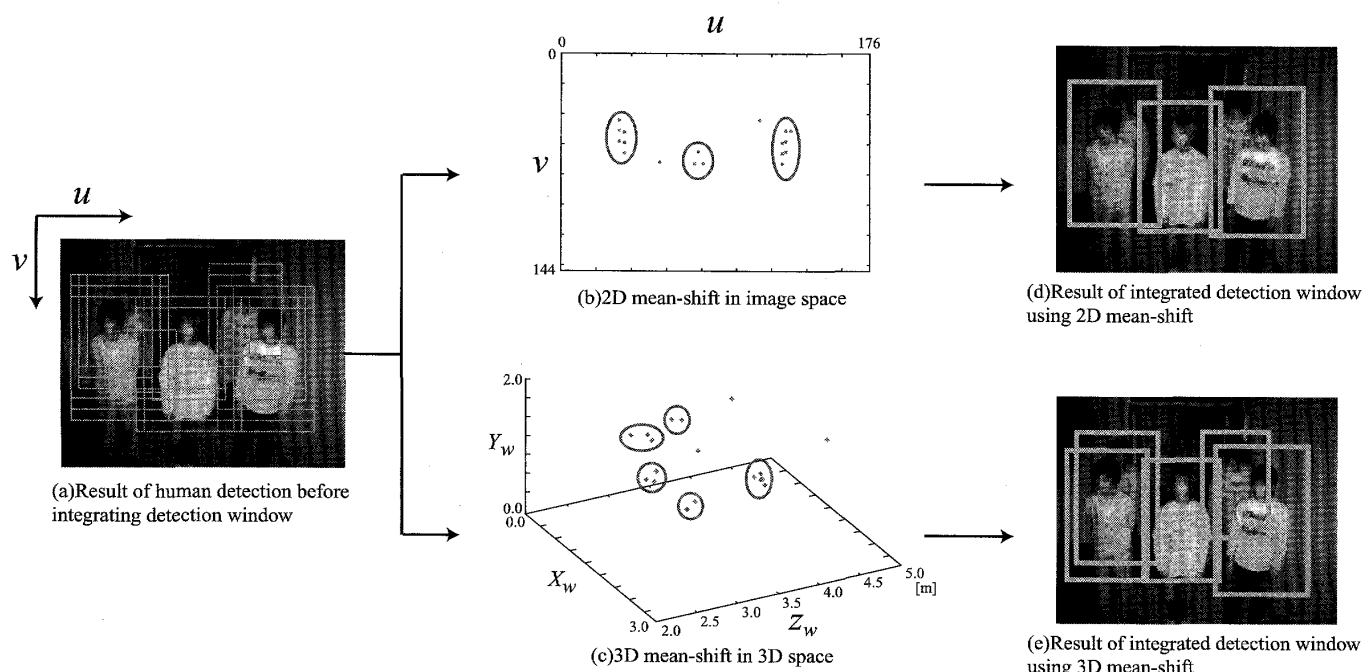
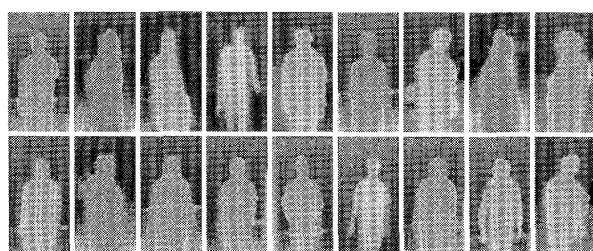
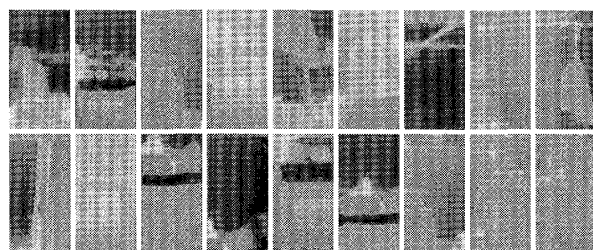


図 9 Mean-Shift クラスタリングによる検出ウィンドウの統合
Fig. 9 Integration of detection window by using mean-shift clustering.



(a) Positive sample (1,346 samples)

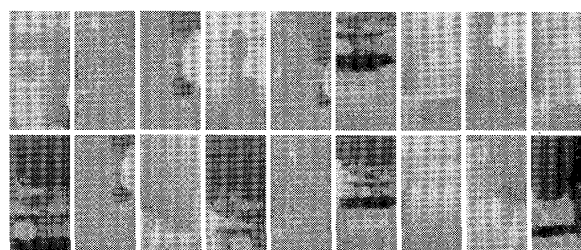


(b) Negative sample (10,000 samples)

図 10 学習用データベースの一部
Fig. 10 Example of training data.



(a) Positive sample (2,206 samples)



(b) Negative sample (8,100 samples)

図 11 評価用データベースの一部
Fig. 11 Example of test data.

には学習用とは別に作成した評価用ポジティブサンプル 2206 枚, 評価用ネガティブサンプル 8100 枚を用いる. TOF カメラは屋内において最長 7.5 m までの撮影となるため, 複数の人の全身を撮影することが困難である. そのため本実験では人の上半身 (全身の上部 60%) を検出対象とした. 図 10 に学習用, 図 11 に評価用のデータベースの一部を示す.

4.2 特徴量の評価実験

評価用データベースを用いて, 人の識別実験を行い, 特徴量の識別精度による比較を行う. 実験には, 距離画像から抽出される HOG 特徴量, 距離ヒストグラム特徴量 (方形サイズ可変なし, 方形サイズ可変あり), 更に, HOG 特徴量と距離ヒストグラム特徴量 (方形サイズ可変あり) の両者を用いた特徴量について比較を行う. 各特徴量ベクトルの次元数を表 1 に示す. 実験

表 1 特徴量ベクトルの次元数と計算時間

Table 1 Dimension number of feature vector and computation cost.

	Dimension number of feature vector	Computation cost per 500 features (1 detection window) [ms]
HOG	3780	0.021
Depth histogram features (Invariable of rectangular region size)	8128	0.067
Depth histogram features (Variable of rectangular region size)	120786	0.067
HOG + Depth histogram features (Variable of rectangular region size)	124566	0.085

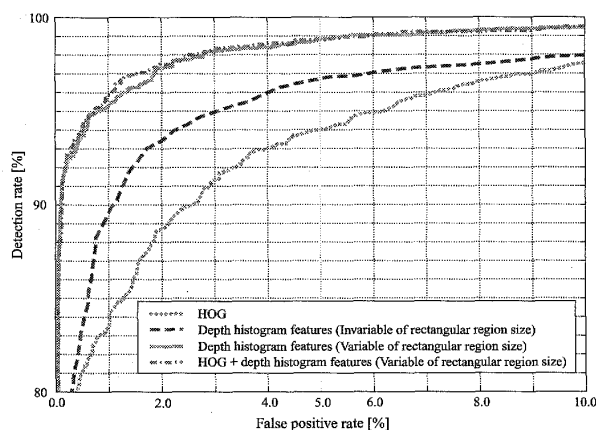
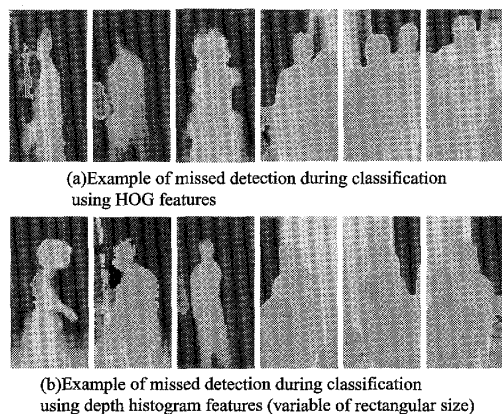
図 12 特徴量の評価実験結果
Fig. 12 Results of using features.

図 13 識別実験における未検出例

Fig. 13 Example of missed detection on classification.

結果の比較には、Receiver Operating Characteristic (ROC) カーブを用いる。ROC カーブとは、横軸に誤検出率、縦軸に検出率を表したものである。識別器のしきい値を変化させることによって、誤検出率に対する検出率の比較を行うことが可能である。グラフ左上に近いほど検出性能が良いことを表す。

図 12 に特徴量の評価実験結果を示す。距離ヒストグラム特徴量（方形領域サイズの可変あり）は、誤検

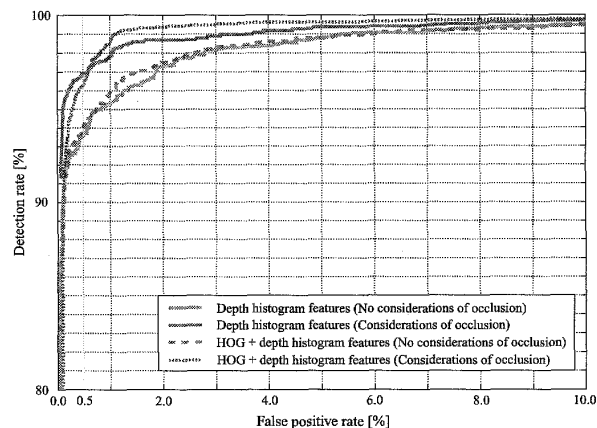
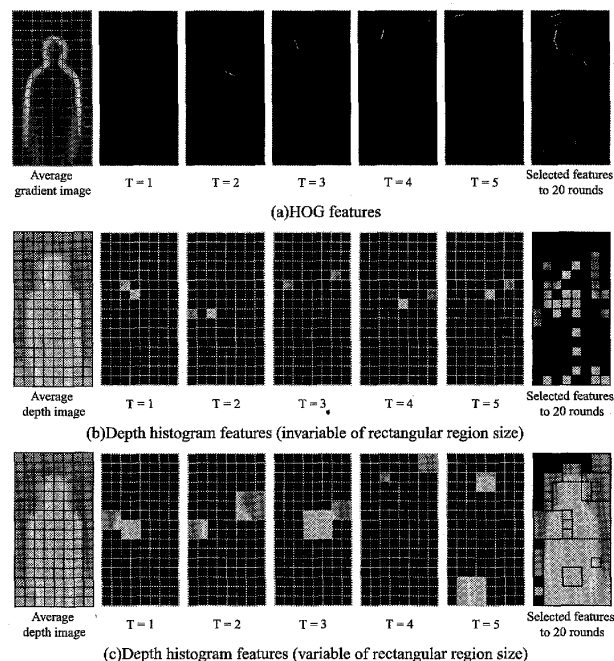
図 14 オクルージョン対応の評価実験結果
Fig. 14 Results of occlusion consideration.

図 15 学習により選択された特徴量

Fig. 15 Selected features by learning.

出率 5.0%において検出率 98.9%であり、距離画像の HOG 特徴量より 4.9%識別率を向上した。これは、方形領域サイズを可変にすることにより、領域の距離関係を有効にとらえることが可能となるためである。また、距離ヒストグラム特徴量（方形領域サイズの可変あり）のみと、HOG 特徴量と距離ヒストグラム特徴量（方形領域サイズの可変あり）の両者を用いた特徴量を比較すると、検出精度は同精度であった。これは、弱識別器の学習時に距離ヒストグラム特徴量（方形サイズ可変あり）が多く（99%）選択されており、HOG 特徴量が識別に貢献していないといえる。表 1 に 1 検出ウィンドウ当りの特徴量計算時間を示す。特徴量

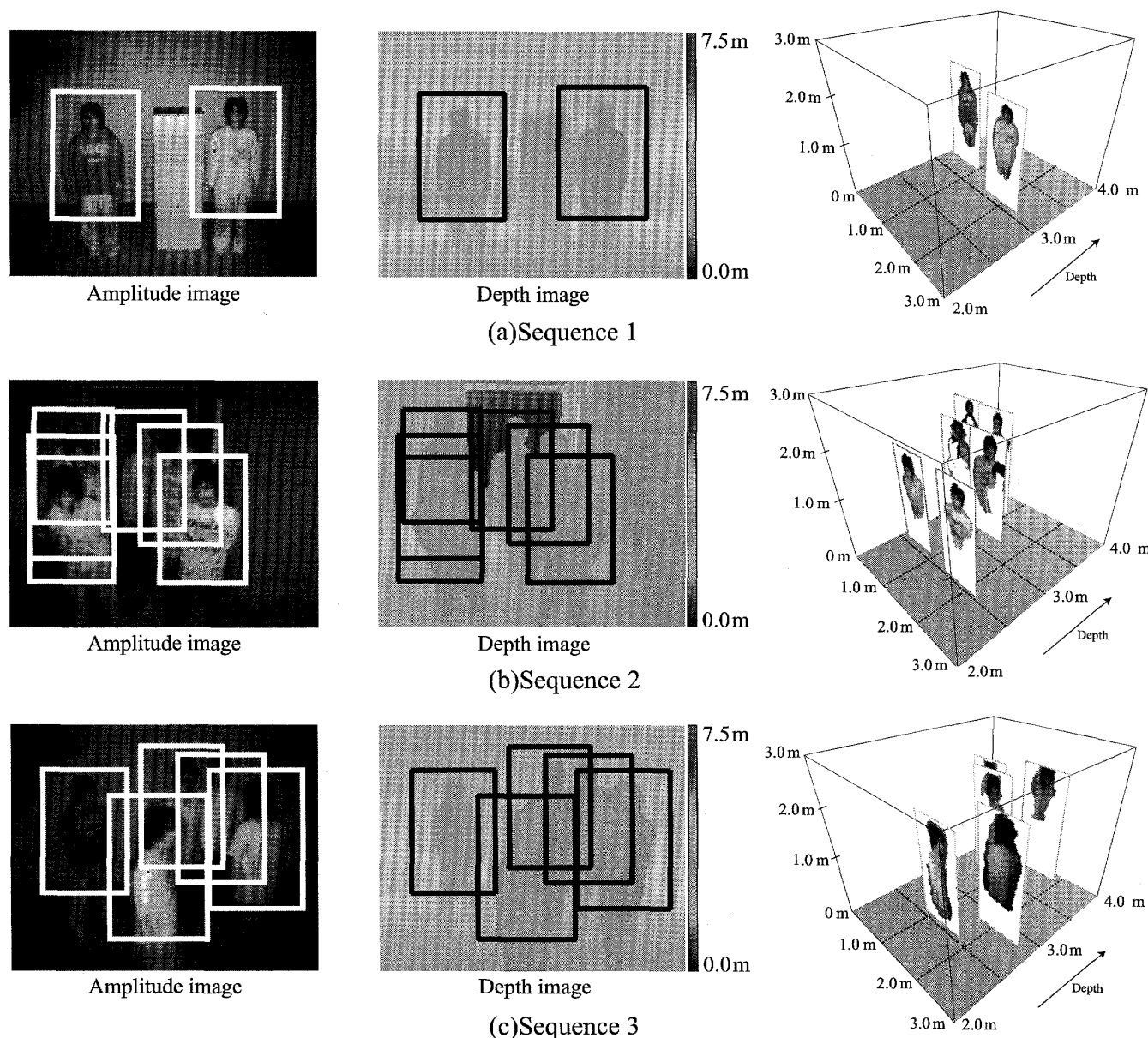


図 16 人検出例

Fig. 16 Example of human detection.

の計算時間は HOG 特徴量が最も高速であることが分かる。また、距離ヒストグラム特徴量（方形領域サイズの可変なし，方形領域サイズの可変あり）を比較すると，計算時間に差がないことが分かる。これは，Integral Histogram を用いて特徴量を計算しているため，計算時間が距離ヒストグラム算出領域の大きさに依存しないことを示している。識別精度と計算コストを考慮すると距離ヒストグラム特徴量（方形領域サイズの可変あり）のみを用いた場合に高速で高精度なリアルタイム検出が可能であるといえる。図 13 に誤識別した例を示す。誤識別サンプルの傾向として，形状変化の大きい場合や，オクルージョンが発生している

ことが分かる。

4.3 オクルージョン対応の評価実験

提案手法におけるオクルージョン対応の有効性を示すため人の識別実験により評価を行う。

図 14 にオクルージョン対応ありとなしの評価実験結果を示す。距離ヒストグラム特徴量（オクルージョン対応あり）は，誤検出率 0.5%において検出率 96.9%であり，オクルージョンが発生している場合でも高精度に識別が可能であることが分かる。更に，HOG+距離ヒストグラム特徴量（オクルージョン対応あり）においても検出率が向上している。これは，オクルージョン率を用いて識別に有効な弱識別器に重み付けし，最

表 2 1 フレームにおける処理時間 [ms]
Table 2 Computation cost per frame [ms].

	Processing cost of 1 detection window	Total (361 windows)
Feature calculation	0.067	24.31
Classification	0.125	45.34
Integration of windows	—	31.97
Total	—	101.62

終識別器の出力を求めることにより、オクルージョン領域の影響を抑えることができたといえる。

4.4 学習により選択された特徴量

図 15 に提案手法により構築した Real AdaBoost の初期の学習において選択された特徴量を示す。HOG 特徴量 (a) では、人の頭部や肩のエッジが選択されていることから、人と背景の境界線をとらえるように特徴量が選択されている。20 回までの学習の傾向としては、形状変化の少ない上半身の特徴量が多く選択されていることが分かる。次に、距離ヒストグラム特徴量 (方形領域サイズの可変なし) (b) では、近接した人領域と背景領域の組合せが多く選択されている。このような距離の異なる二つの領域の距離ヒストグラムの類似度は 0 に近くなり、二つの領域間に人と背景の境界を表現することが可能であることから、境界領域をとらえていることが分かる。20 回までの学習の傾向としては上半身の境界領域が多く選択されている。更に、距離ヒストグラム特徴量 (方形領域サイズの可変あり) (c) では、人と背景の境界領域を任意のサイズでとらえるように領域が選択されている。これは、人の境界領域の距離関係をとらえることで大まかな人の形状をとらえていると考えられる。また (c) においては、2 回目や 5 回目の学習で人の上下、左右の対称的な距離関係をとらえるように領域が選択されている。20 回までの学習の傾向としては境界領域が任意のサイズで選択されている。

4.5 リアルタイム人検出

三次元実空間における検出ウィンドウのラスタスキャンを用いた人検出結果を図 16 に示す。(a) では、人と同様の高さの物体を誤検出しないで、人のみを検出していることが分かる。更に (b)(c) では、向きの異なる人の重なりが存在しても、それぞれの人とその三次元位置を正確に検出できていることが分かる。また、表 2 に Intel Xeon CPU 3.00 GHz を用いた際の 1 フレーム (検出ウィンドウ 361 個) の処理時間を示す。提案手法は約 100 ms で処理を行うため、約 10 fps で

のリアルタイム検出が可能である。

5. む す び

本論文では TOF カメラから得られる距離情報を用いて人の重なりや複雑なシーンに頑健なリアルタイム人検出手法を提案した。評価実験の結果、従来法である HOG 特徴量より 4.9% 検出率を向上させることができた。更に提案手法は、検出ウィンドウ内におけるオクルージョンの割合であるオクルージョン率を算出し、オクルージョン領域を考慮した識別を行うことにより、オクルージョンが発生している場合でも高精度な識別が可能であることを確認した。提案手法は、特徴量算出、識別、検出ウィンドウの統合におけるすべての処理時間が約 100 ms であるため、約 10 fps でのリアルタイム検出が可能である。今後は、距離情報とその時間変化を用いた動作解析を行う予定である。

文 献

- [1] 藤吉弘亘, “人を観る技術: People Image Analysis—インビジブルロボットの視覚機能の実現,” 精密工学会サマーマセミナー 2008, pp.113–120, 2008.
- [2] K. Levi and Y. Weiss, “Learning object detection from a small number of examples: The importance of good features,” IEEE Comput. Vis. Pattern Recognit., vol.2, pp.53–60, 2004.
- [3] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” IEEE Comput. Vis. Pattern Recognit., vol.1, pp.886–893, 2005.
- [4] B. Wu and R. Nevatia, “Detection of multiple, partially occluded humans in a single image by bayesian combination of edgelet part detectors,” IEEE International Conference on Computer Vision, vol.1, pp.90–97, 2005.
- [5] P. Szabzmeydani and G. Mori, “Detecting pedestrians by learning shapelet feature,” IEEE Comput. Vis. Pattern Recognit., pp.511–518, 2007.
- [6] 三井相和, 山内悠嗣, 藤吉弘亘, “Joint HOG 特徴を用いた 2 段階 AdaBoost による人検出,” 第 14 回画像センシングシンポジウム (SSII08), IN1-06, 2008.
- [7] 山内悠嗣, 藤吉弘亘, 山下隆義, “Boosting に基づく共起表現による人検出,” 第 11 回画像の認識・理解シンポジウム (MIRU2008), pp.180–187, 2008.
- [8] A. Ess, B. Leibe, and L.V. Gool, “Depth and appearance for mobile Scene Analysis,” IEEE International Conference on Computer Vision, pp.1–8, 2007.
- [9] 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.
- [10] F. Porikli, “IntegralHistogram: A fast way to extract histograms in cartesian spaces,” IEEE Comput. Vis. Pattern Recognit., vol.1, pp.829–836, 2005.

- [11] R.E. Schapire and Y. Singer, "Improved boosting algorithms using confidence-rated predictions," Mach. Learn., no.37, pp.297-336, 1999.
- [12] J. Mure-Dubois and H. Hugli, "Fusion of time of flight camera point clouds," ECCV Workshop on M2SFA2, 2008.
- [13] D. Comaniciu and P. Meer, "Mean shift analysis and applications," IEEE International Conference on Computer Vision, pp.1197-1203, 1999.

(平成 21 年 6 月 23 日受付, 10 月 19 日再受付)



池村 翔

2009 中部大・工・情報卒。同年より同修士課程に在籍。画像処理, パターン認識に関する研究に従事。



藤吉 弘亘 (正員)

1997 中部大学大学院博士後期課程了。1997~2000 米カーネギーメロン大学ロボット工学研究所 Postdoctoral Fellow。2000 中部大学講師を経て 2004 より同大准教授。2006 米カーネギーメロン大学ロボット工学研究所客員研究員。工博。計算機視覚, 動画画像処理, パターン認識・理解の研究に従事。2005 年度ロボカップ研究賞。情報処理学会, 電気学会, IEEE 各会員。