

姿勢変化に対応した Soft Decision Feature と Online Real Boosting による人物追跡

山下 隆義^{†,††} 藤吉 弘巨^{††} 勞 世竝[†] 川出 雅人[†]

[†] オムロン株式会社 技術本部 京都府木津川市木津川台 9-1

^{††} 中部大学 愛知県春日井市松本町 1200

あらまし 統計的学習手法の一つであるブースティング手法をオンラインでの物体追跡に拡張した Online Boosting が注目されている。Online Boosting は追跡対象に合わせて適した弱識別器を選択することが可能である。しかしながら、オフラインで得られた弱識別器は、人体のように姿勢や向き変化が生じる場合、最適な特徴量であるとは限らない。そこで我々は、弱識別器の形状をオンラインで最適化することが可能な Soft Decision Feature を提案する。Soft Decision Feature と Online Real Boosting を組み合わせることで、従来手法では追跡することができない体の向き変化や姿勢変動が生じる場合でも追跡することが可能である。

キーワード 人体トラッキング, ブースティング, Online Real Boosting, Edgelet, EHO

Human Tracking based on Soft Decision Feature and Online Real Boosting

Takayoshi YAMASHITA^{†,††}, Hironobu FUJIYOSHI^{††}, Shihong LAO[†], and Masato KAWADE[†]

[†] OMRON Corporation 9-1, Kizugawadai, Kizugawa-city KYOTO, JAPAN

^{††} Chubu University 1200, Matsumoto-cho, Kasugai-city, AICHI, JAPAN

Abstract Online Boosting is an effective incremental learning method which is applied offline boosting method to online object tracking. It can select best weak classifiers deal with object pose change. However, these weak classifiers do not select best ones, because the feature pool which is trained in offline has small number of weak classifiers. It needs the framework of select suitable classifiers deal with object pose and posture change. We propose the new type of feature which can adjust to object shape variation call Soft Decision Feature. This new adjustable feature and Online Real Boosting has better tracking performance in the scene of human pose and posture change.

Key words Human tracking, Boosting, Online Real Boosting, Edgelet, EHO

1. はじめに

近年、統計的学習手法の一つであるブースティングが注目されている。ブースティングは正解サンプルと非正解サンプルを入力とし、識別器候補の中から多数の弱識別器を選択することで強識別器を学習する学習手法である。ブースティングは、弱識別器が 0 または 1 の二値を出力する AdaBoost が一般的であるが、実数値を出力する Real AdaBoost [1] や弱識別器の選択方法に Logit 関数を導入した LogitBoost などがある。

Viola ら [2] が AdaBoost を利用した顔検出手法を提案して以降、ブースティングを用いた検出手法が多く提案されている [2] [3] [4]。Wu ら [2] は顔検出に Real AdaBoost を用いることで Viola の提案手法より少ない弱識別器数で高い検出率の手法を提案している。また、Zhu ら [3] はエッジの勾配方向に着

目した HOG 特徴量と AdaBoost を利用した人体検出手法を提案している。

これらの検出手法に用いられているブースティングはオフラインの学習手法であるが、Oza ら [5] はオフラインで学習した弱識別器を追跡対象に合わせてオンラインで更新する Online Boosting を提案している。Oza らが提案した Online Boosting は弱識別器の重みを更新するだけであるが、Grabner ら [6] は弱識別器の選択も同時に行うように拡張した手法を提案している。また、山下ら [7] はオフラインで Real AdaBoost により得られた弱識別器の確率密度関数をオンラインで用いることで、少ない弱識別器数で高い追跡性能を得る Online Real Boosting を提案している。

オフライン学習であるブースティングをオンライン学習に拡張した Online Real Boosting は、追跡対象の変化に合わせて

最適な弱識別器を得ることができる。しかしながら、弱識別器はオフライン学習で得られたセットの中から選択する必要がある。そのため、追跡対象の向き変化などによる形状変化が生じた場合、選ばれた弱識別器は追跡対象にとって最適であるとは限らない。このように、従来の Online Real Boosting はあらかじめオフラインで学習した限られた弱識別器セットから選択するため、Hard Decision Feature によるオンライン学習であるといえる。

そこで、我々は追跡対象の形状変化に合わせて適応的に弱識別器の形状を変化させることが可能な Soft Decision Feature によるオンライン学習手法を提案する。Soft Decision Feature は、オフライン学習時に弱識別器が複数の変形可能な形状を記憶しており、追跡対象の形状に合わせて弱識別器の形状を変化させることが可能な弱識別器である。我々が提案する Soft Decision Feature を用いた Online Real Boosting について、次章以降で説明する。

2. 従来の Online Boosting について

Oza らが提案した Online Boosting はオフラインで得られた弱識別器の重みを追跡対象に合わせて更新する。従来の Online Boosting はすべての弱識別器を利用することを前提としているが、Grabner らは弱識別器を複数のセクタというグループに分割し、追跡対象に合わせてセクタ内で弱識別器を選択する選択処理を導入している。さらに、山下らは Grabner らの Online Boosting に Real AdaBoost を用いることで少ないセクタ数で高い追跡性能を得る手法を提案している。ここでは、山下らが提案した Online Real Boosting と従来の Online Boosting の問題点について説明する。

2.1 Online Real Boosting

Online Real Boosting は Real AdaBoost により、オフラインで強識別器 $H_{strong} = \{h_1, \dots, h_M\}$ を学習する。弱識別器 $h_m(x)$ は、オフライン学習時に得られた正解サンプルおよび非正解サンプルの確率密度関数 $W_m^+(x)$, $W_m^-(x)$ をもっている。確率密度関数 $W_m^y(x)$ は弱識別器により算出される特徴量の値に対する正解および非正解サンプルらしさを出力する関数であり、一般的にはヒストグラムで表現される。確率密度関数 $W_m^y(x)$ は式 (1) のように更新される。

$$W_{m,j}^y(x) = W_{m,j}^y(x) + \lambda_L \quad (1)$$

ここで、 j はヒストグラムで表現されている確率密度関数 $W_m^y(x)$ の j 番目の要素であり、 λ_L はサンプル画像の重みである。また、 y はサンプルが正解または非正解をあらわすラベルである。

人物を追跡する際に、時刻 t における L 個の学習サンプル $S_t = \{(x_1, y_1), \dots, (x_L, y_L)\}$ を用いて、確率密度関数 $W_{m,j}^y(x)$ を更新する。オフラインで学習した弱識別器はセクタに分割されており、各セクタにおいて式 (2) の評価値 Z が最も小さい弱識別器を選択する。

$$Z = \sum_{j=1}^N \sqrt{W_j^+ W_j^-} \quad (2)$$

ここで、 N は確率密度関数 $W_{m,j}^y(x)$ のヒストグラムの要素数である。各セクタで選択された弱識別器 $h_m(x)$ を統合することで、式 (3) のように現時点 t における強識別器 $H_t(x)$ を学習する。

$$H_t(x) = \sum_{m=1}^M h_m(x) \quad (3)$$

2.2 問題点

Online Real Boosting による弱識別器の選択は、オフライン学習で得られた弱識別器に依存する。そのため、人体のように体の向きや姿勢の変化により形状が大きく異なる物体を追跡対象とする場合、あらかじめ用意された弱識別器セットに最適な弱識別器が含まれているとは限らない。従来のオンライン学習は限られた弱識別器セットである Hard Decision Feature を用いたオンライン学習であるといえる。

そこで我々は、オフラインで選ばれた弱識別器セットをもとに、追跡対象の形状変化に応じて、適応的に弱識別器の形状を変化させることができる Soft Decision Feature によるオンライン学習を提案する。

3. 提案手法

提案手法は図 1 に示すように、弱識別器と変形可能な形状を同時に学習するオフライン学習と、オフラインで得られた弱識別器を追跡対象に合わせて適応的に形状を変化させていくことが可能なオンライン学習から構成される。これらについて詳しく説明する。

3.1 Soft Decision Feature

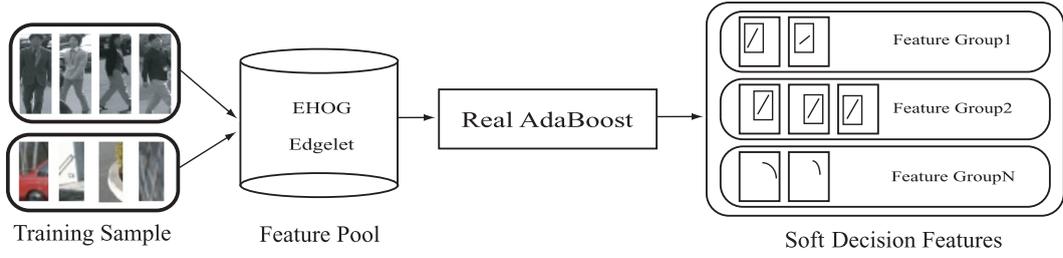
従来の特徴量はあらかじめ決められた位置やサイズ、形状に対する値を算出する Hard Decision Feature である。我々が提案する Soft Decision Feature は追跡対象に合わせて適応的に位置や大きさ、形状を変化させ、最も尤度の高い特徴量を算出することができる。Soft Decision Feature の特徴量のベースとして、Hou ら [8] が提案した EHOGE(Extended Histogram of Oriented Gradients) と Wu ら [9] の Edgelet を用いる。

EHOGE は勾配方向ごとのエッジ強度のヒストグラムを特徴量とする HOG を高速化した特徴量である。図 2 に HOG と EHOGE の違いを示す。図 2 の (a) のように、HOG はブロックに複数のセルを持ち、各セルの勾配方向をベクトル化したベクトル情報の特徴量としている。一方、EHOGE は図 2 の (b) のように、ブロック内にセルを持たず、ブロック内の複数の勾配方向の総和をもとに式 (4) のように特徴量を算出する。

$$F(D) = \sum_{i \in D} g_b(i) / Z_b \quad (4)$$

ここで、 $g_b(i)$ は勾配方向 $b = \{1, \dots, n, n = 9\}$ のヒストグラムであり、 D は b の集合である。

Offline Training



Online Training

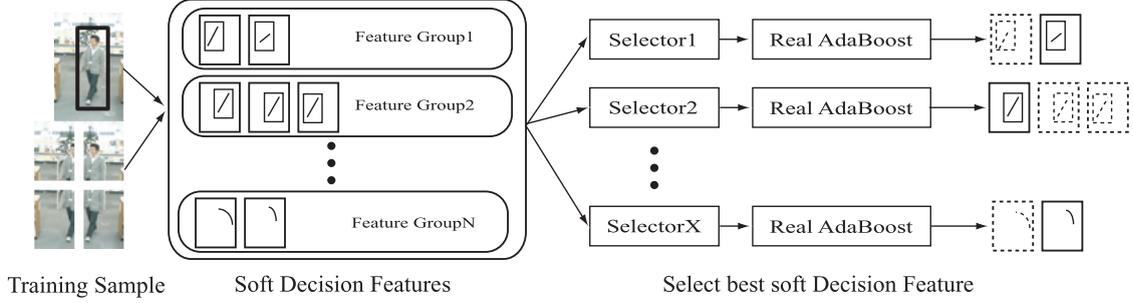


図 1 提案手法のフレームワーク

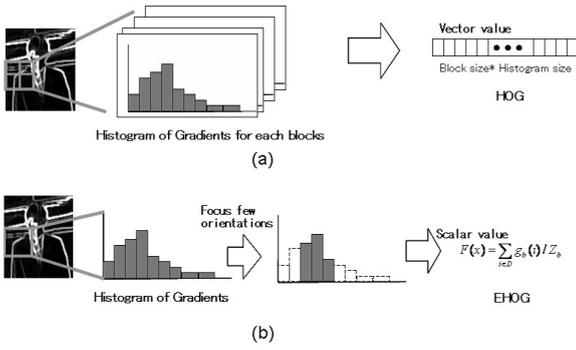


図 2 HOG と EHO の違い
(a)HOG の特徴量, (b) は EHO の特徴量

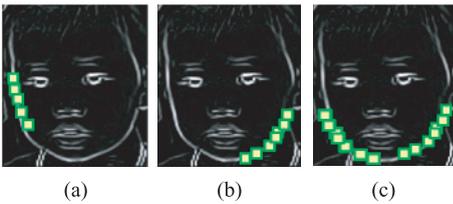


図 3 Edgelet の特徴量パターン
(a) 直線タイプ, (b) 円弧タイプ, (c) 対称性タイプ

また, Z_b は全勾配方向のヒストグラムの総和であり, 式 (5) のように示される.

$$Z_b = \sum_{b=1}^B g_b(i) \quad (5)$$

ここで, B は全勾配方向数であり, B は 9 としている.

一方, Edgelet は図 3 のようにエッジの直線または円弧, 対称性に着目した特徴量である. これらの形状における複数の位置 u_i でのエッジ強度と勾配方向の違いをもとに, 式 (6) のよ

うに算出する.

$$S(x) = \left(\frac{1}{K}\right) \sum_{i=1}^K E^I(u_i) \cdot l[V^I(u_i) - V_i^E] \quad (6)$$

ここで, K は参照する位置の個数であり, $E^I(u_i)$ は位置 u_i におけるエッジ強度である. また $l[V^I(u_i) - V_i^E]$ は位置 u_i における勾配方向 V^I と i 番目の要素の勾配 V^E との差をもとにした正規化関数であり, 式 (7) のようになる.

$$l[\cdot] = \begin{cases} 1 & x = 0 \\ 0.8 & x = \pm 1, \pm 5 \\ 0.5 & x = \pm 2, \pm 4 \\ 0 & x = \pm 3 \end{cases} \quad (7)$$

ここで, x は勾配方向の差である.

EHO と Edgelet はともにエッジに着目した特徴量であるが, EHO は局所的な領域の勾配変化を捉えているのに対して, Edgelet は部分的なエッジの類似度を捉えている. 本提案手法では, このような捉え方の異なる 2 つの特徴量を用いる.

Soft Decision Feature は式 (8) のように, 基準となる特徴量 F_{org} に対して, 変形関数 T により変形させることで新たな形状の特徴量 F_{trans} を得ることができる.

$$F_{trans} = F_{org} * T(\cdot) \quad (8)$$

変形関数 $T(\cdot)$ は特徴量の種類によりパラメータは異なる. EHO の場合, 変形パラメータとして, 位置およびサイズ, 勾配方向を用いる. EHO の Soft Decision Feature は式 (9) のように表すことができる.

$$F_{trans}^{EHO} = F_{org}^{EHO} * T(x, y, w, h, g) \quad (9)$$

ここで, 変形関数 T は移動量 x , y とサイズ変化量 w , h ,

勾配方向の変化量 g である。

また, Edgelet の場合, 特徴量は複数の要素 K から構成されており, 各要素 u_k は位置と勾配方向を持っている. Edgelet の Soft Decision Feature は式 (10) のように, 要素ごとに移動量 x, y と勾配の変化量 g を変形パラメータとしている.

$$F_{trans}^{Edgelet} = F_{org}^{Edgelet}(u_1 * T(x, y, g)), \dots, (u_k * T(x, y, g)) \quad (10)$$

これらのパラメータを制御することで, 全ての特徴量の形状の中から適応的に変形可能な形状をグルーピングすることができるが, 本提案手法では近隣の位置および大きさ, 勾配での形状変化に限定し, 学習を行う. 各パラメータは ± 1 の範囲とする.

3.2 Soft Decision Feature のオフライン学習

Soft Decision Feature のオフライン学習の手順を図 4 に示す. また, Soft Decision Feature の概要について図 5 に示す. オフライン学習はあらかじめ正解および非正解の学習サンプルを用意し, 1 のように均等に重みを初期化する. また, 特徴量候補もあらかじめ, 位置やサイズなどを変えて用意しておく.

図 4 の 2.1.1 の処理において, 特徴量候補 h_m に対して各サンプル画像から算出される特徴量をもとに確率密度関数 W_{org}^y を作成する. 確率密度関数 W_{org}^y は重みつきヒストグラムで表現されており, サンプル画像の特徴量をもとにヒストグラムの要素 bin_j を決定し, 重みを加算していく. Soft Decision Feature は, 図 4 の 2.1.2 および 2.1.3 において, 弱識別器候補の変形を行い, 各確率密度関数 W_{trans}^y を求める.

次に, 2.1.4 において, ベースとなる弱識別器候補の確率密度関数 W_{org}^y と変形した弱識別器候補 W_{trans}^y の類似度 S を算出する. 算出された類似度 S が閾値 TH より高ければ, 変形可能な形状とし, 2.1.5 のように確率密度関数の結合を行う. 同様に, 形状の変形を繰り返し行い, 1 つの弱識別器候補が変形可能な形状を図 5 のようにグルーピングしていく. これにより, Soft Decision Feature は 1 つの確率密度関数に対して, 複数の形状を持つことができる. 結合された Soft Decision Feature の弱識別器候補は, 2.1.5 で結合された確率密度関数 W_j^y をもとに弱識別器候補の評価値 Z を 2.1.6 で算出する. すべての弱識別器候補に対して, 評価値 Z を算出し, 評価値が最も小さい弱識別器候補を現試行回数 t で選択された弱識別器 h_t とする.

次に, 図 4 の 2.3 において, 選ばれた弱識別器 h_t の識別関数を算出する. 2.4 で算出された識別関数を用いて, 2.5 のように各学習サンプルを評価する. 2.5 で算出されたエラー率 ϵ_t をもとに, 2.6 のように各学習サンプルの重み λ_L を更新する.

Hard Decision Feature は図 4 の 2.1 において, 2.1.1 のように 1 つの弱識別器候補に対して, 確率密度関数の更新を行い, 2.1.6 のように評価値を求めることで弱識別器を学習する. 一方, Soft Decision Feature は 1 つの弱識別器候補の確率密度関数が類似する別の弱識別器候補を変形可能な形状として, グループ化する. これにより, オンラインでの学習時に変形可能な形状を得ることができる.

1. 学習サンプルの初期化 $\lambda_L = 1/N$
2. For $t = 1, \dots, T$ 学習の試行回数
 - 2.1 For $m = 1, \dots, M$ 弱識別器候補数
 - 2.1.1 確率密度関数 W の作成

if $h_m(x) \in bin_j$

$$W_{org,j}^y = W_{org,j}^y + \lambda_L \quad y = \{0, 1\} \quad (11)$$
 - 2.1.2 弱識別器の変形

$$h_{trans} = h_{org} * T(x, y, w, h, g)$$
 - 2.1.3 確率密度関数の更新

if $h_{trans}(x) \in bin_j$

$$W_{trans,j}^y = W_{trans,j}^y + \lambda_L \quad (12)$$
 - 2.1.4 類似度を算出

$$S = 2 \sum_{j=1}^N \sqrt{W_{org,j}^y W_{trans,j}^y} \quad (13)$$
 - 2.1.5 確率密度関数を結合

if $S > TH$

$$W_j^y = W_{org,j}^y + W_{trans,j}^y \quad (14)$$

2.1.2 から 2.1.5 を繰り返す
 - 2.1.6 評価値 Z の計算

$$Z_m = 2 \sum_{j=1}^N \sqrt{W_j^+ W_j^-} \quad (15)$$
 - 2.2 弱識別器候補の選択

$$m^* = \arg \min_m Z_m \quad (16)$$

$$h_t = h_{m^*} \quad (17)$$
 - 2.3 識別器の計算

$$h_t(x) = \frac{1}{2} \ln \left(\frac{W_t^+(j) + \beta}{W_t^-(j) + \beta} \right) \quad (18)$$
 - 2.4 エラー率の計算

$$\epsilon_t^y = \epsilon_t^y + \lambda_L |h_t(x_n)| \quad (19)$$

$$\epsilon_t = \frac{\epsilon_t^+}{\epsilon_t^+ + \epsilon_t^-} \quad (20)$$
 - 2.5 重みの更新

if $y h_t(x_m) > 0$ $\lambda_L = \frac{\lambda_L}{2(1 - \epsilon_t)}$ (21)

else $\lambda_L = \frac{\lambda_L}{2\epsilon_t}$ (22)
3. 強識別器

$$H(x) = \sum_{t=1}^T h_t(x) \quad (23)$$

図 4 Soft Decision Feature によるオフライン学習

3.3 Soft Decision Feature のオンライン学習と追跡処理

Online Real Boosting に Soft Decision Feature を用いたアルゴリズムの手順を図 6 に示す. オンライン学習は従来の Online Real Boosting と同様にセレクト数 N 個に分割し, 図

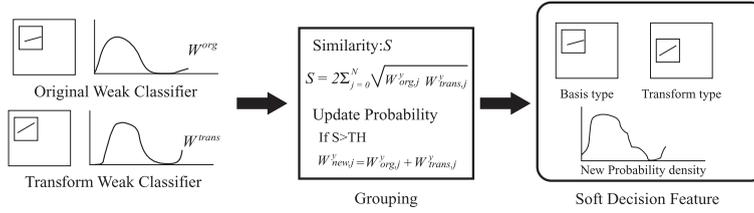


図5 Soft Decision Feature の概要

0. 弱識別器セット $h_{offline,m}(x)$ $m = 1, \dots, M$
1. 弱識別器 M 個をセクタ N 個に分割
2. 現時刻の学習サンプル $I(x, y), x = 1, 2, \dots, L, y \in \{+1, -1\}$
3. サンプルデータの重みを初期化 $\lambda_l = \frac{1}{L}$
4. For $n = 1, 2, \dots, N$ //セクタ数
 - 4.1 For $m = 1, 2, \dots, M/N$ //弱識別数
 - 4.1.1 弱識別器の確率密度関数 W を更新
if $h_{n,m}(x) \in bin_j$
$$W_{n,m,j}^y = W_{n,m,j}^y + \lambda_l \quad y = \{0, 1\} \quad (24)$$
 - 4.1.2 弱識別器の評価値 Z を算出
$$Z_{n,m} = 2 \sum_{j=1}^N \sqrt{W_{n,m,j}^+ W_{n,m,j}^-} \quad (25)$$
 - 4.2 Soft Decision Feature の選択
$$m^* = \arg \min_m Z_{n,m} \quad (26)$$

$$h_{n,org} = h_{m^*} \quad (27)$$
 - 4.3 Soft Decision Feature の弱識別器最適化
 - 4.3.1 弱識別器を変形
$$h_{n,trans} = h_{n,org} * T(x, y, w, h, g)$$
 - 4.3.2 確率密度関数を更新
$$W_{n,trans,j}^y = W_{n,trans,j}^y + \lambda_l \quad y = \{0, 1\} \quad (28)$$
 - 4.3.3 評価値 Z を算出
$$Z_{n,trans} = \sum_{j=1}^N \sqrt{W_{n,trans,j}^+ W_{n,trans,j}^-} \quad (29)$$
 - 4.3.4 弱識別器を選択
$$n^* = \arg \min_n Z_{n,trans} \quad (30)$$

$$h_n = h_{n^*} \quad (31)$$
5. 強識別器
$$H(x) = \sum_{n=1}^N h_n(x) \quad (32)$$

図6 Soft Decision Feature によるオンライン学習

6の4のような手順でセクタごとに最適な弱識別器を選択する。オンラインでの新しい学習サンプルは現時刻において、検出した人体の位置を正解サンプルとし、その周囲を非正解サンプルとする。非正解サンプルは人体領域を部分的に含むような位置である。

次に、4.1 および 4.2 において、オフライン学習と同様に、



図7 オフライン学習サンプル例

弱識別器 $h_{n,m}(x)$ の確率密度関数 $W_{n,m}^y$ を更新し、弱識別器の評価関数 $Z_{n,m}$ を求める。評価値 $Z_{n,m}$ が最も小さな弱識別器 h_{m^*} を Soft Decision Feature として選択する。4.3 で Soft Decision Feature の形状を変形させ、各変形した弱識別器 $h_{n,trans}$ の確率密度関数 $W_{n,trans}^y$ と評価値 $Z_{n,trans}$ を求める。4.2 と同様に、評価値が最も小さな Soft Decision Feature の形状をセクタの弱識別器 h_n の形状とする。これらの手順を全てのセクタに対して行い、強識別器 $H(x)$ とする。

追跡にはパーティクルフィルタを用いた。各サンプリング点の重みは式 (29) で算出した尤度をもとに算出している。追跡位置はサンプリング点のうち、強識別器 $H(x)$ が最も大きい位置を人体の位置として出力する。なお、本提案手法でのパーティクル数は 100 としている。

4. 実験

4.1 実験概要

提案手法の Soft Decision Feature を利用した Online Real Boosting と従来の Hard Decision Feature を利用した Online Real Boosting の追跡性能の比較実験を行った。比較実験は人体の向きが変化するようなシーンおよび姿勢が変化するシーン (約 800 フレーム) での追跡位置のばらつきを比較した。

はじめにオフラインで人体の弱識別器を学習した。学習データは 24×60 ピクセルに正規化し、約 5000 枚用いた。学習する弱識別器数は 500 個とした。オフライン学習に用いた学習サンプルは、図7に示すように人体の向きは正面だけでなく側面も含んでいるが、座り込みなどの姿勢変動が生じている画像は含んでおらず直立姿勢のみである。オフラインでの学習により、選ばれた弱識別器の上位 10 個と全弱識別器を重ね合わせた人体合成モデルを図8に示す。図8(a)の矩形で描画されている特徴量が EHOg であり、白線で描画されている特徴量が Edgelet である。図8(b)の人体合成モデルは Edgelet の形状

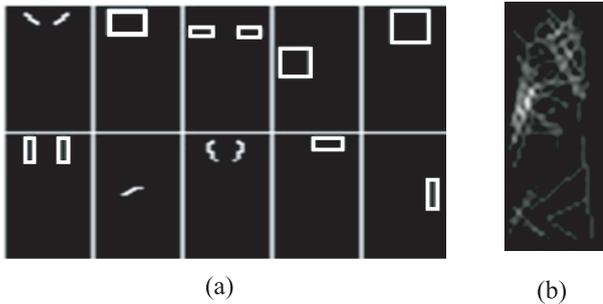


図 8 オフライン学習結果

(a) 上位 10 個の弱識別器, (b) 全弱識別器による合成人体モデル

表 1 追跡性能の比較

姿勢変化	提案手法		Hard Decision Feature	
	位置精度 S[%]	分散	位置精度 S[%]	分散
(a). 直立 (正面)	1.73	1.17	1.84	1.91
(b). 直立 (横)	1.83	4.26	3.95	9.72
(c). 座り込み	2.10	4.12	7.34	4.08
(d). 直立 (正面)	3.07	4.64	Miss track	-
(e). 直立 (横)	2.70	4.19	Miss track	-
全体	4.59	13.99	7.52	39.99

と EHOg の勾配方向を重ね合わせている。これより、特徴量は頭部付近や体の輪郭付近に多く分布していることがわかる。

次に、Online Real Boosting のセレクトタ数は 50 としており、各セレクトタにはオフラインで学習した弱識別器が 10 個ずつ含まれている。

追跡性能は追跡位置の精度を比較した。評価映像は、あらかじめ人体領域を囲むような矩形を正解矩形として入力している。正解矩形の中心座標を求め、中心座標と追跡結果の中心座標の距離をもとに位置精度を算出する。位置精度 S は式 (33) のように、人体領域の幅に対する正解中心座標からの距離の割合とする。

$$S = \text{Dist}(\text{Tracking}, \text{GroundTruth}) / \text{Width}[\%] \quad (33)$$

4.2 追跡精度の比較

体の向きおよび姿勢変化が生じる映像を用いて、追跡位置の精度を比較した。映像は正面向きの直立と横向きの直立、横向きの座り込みが含まれており、直立から座り込み後、直立するような映像である。

追跡結果例として、図 9 に提案手法、図 10 に Hard Decision Feature による Online Real Boosting の結果を示す。図 9 および図 10 の (a) は正面向きの直立時、(b) は横向きの直立時、(c) は座り込み時、(d) および (e) は座り込み後の正面向きの直立時、横向きの直立時の結果である。また、追跡位置の位置精度の比較結果を表 1 に示す。位置精度はシーケンス全体に対する精度と (a) から (e) に相当する姿勢ごとの精度を算出した。表 1 より、直立 (正面) 時の位置精度は大きな差がないが、直立 (横) 時は Hard Decision Feature に比べて、提案手法の位置精度がよいことがわかる。また、座り込み時も提案手法は直立時と同等の位置精度があり、図 9(c) のように人物の領域を正しく追跡することができていることがわかる。一方、Hard

Decision Feature の場合、位置精度が直立時より低下しており、図 10(c) のように人物の領域からずれていることがわかる。また、座り込み後の直立時は、図 10 の (d) および (e) のように追跡対象を正しく捉えることができていない。これは、追跡過程における位置ずれの影響により、弱識別器の確率密度関数の更新が正しく行われていないためである。提案手法の Soft Decision Feature は座り込みまでの位置精度が高いため、確率密度関数の更新が正しく行われており、座り込み後の直立時も位置精度が高く、図 9 の (d) および (e) のように追跡が成功している。これより、提案手法は Hard Decision Feature を用いた場合に比べて、追跡位置精度が高いため、姿勢変化が生じた場合でも正しく追跡対象を追跡することが可能である。

4.3 姿勢変化への汎化性

これまでは、姿勢変化として大きな変化を伴う座り込みについて実験を行い、提案手法が姿勢変化にロバストに追跡できることを示した。本節では、その他の姿勢変化に対する追跡性能について実験を行った。姿勢変化として、物を拾うシーンおよび走り出すシーンを対象とした。物を拾うシーンは前かがみにしゃがみこむため、座り込みと同様に大きな姿勢変化をとる。一方、走り出すシーンは座り込みやしゃがみこみと比較すると大きな姿勢変化を伴わないが、手や足が大きく動くため、時系列での変化が激しいシーンである。図 11 に提案手法による物を拾うシーン、図 12 に走り出すシーンでの追跡結果を示す。このように、提案手法はしゃがみこみや走り出すシーンでも正しく追跡することができている。これらの実験結果より、提案手法は座り込みやしゃがみこみ、走るシーンなどの様々な姿勢変化に対して正しく追跡することが可能である。

4.4 処理速度

提案手法の処理速度は、Pentium Core 2 Duo 2.4GHz において約 30ms とリアルタイムでの追跡を実現した。処理速度はセレクトタ数とセレクトタに含まれる弱識別器数に依存する。そのため、オフライン学習により選択される弱識別器の識別能力が高ければ、セレクトタ数を削減することができ、さらに高速化することができる。

5. Hard Decision V.S. Soft Decision

5.1 選択された特徴量の比較

図 9 および図 10 のうち、座り込み時のフレームにおける Soft Decision Feature と Hard Decision Feature を合成した人体形状を図 13 に示す。図 13(a) は Soft Decision Feature の追跡結果であり、図 13(b) はこの時に選択された弱識別器を合成することで得られる合成人体モデルである。また、図 13(c) は Hard Decision Feature による追跡結果、図 13(d) は弱識別器による合成人体モデルである。図 13(b) および (d) の線の色は弱識別器の重なりが大きいほど白色で表現されている。座り込み時の追跡結果である図 13(a) に対する Soft Decision Feature は頭部から背中付近に分布する弱識別器が選択されている。また、足元や腕付近にも均等に弱識別器が配置されており、体全体を捉えることができています。一方、座り込み時の追跡結果図 13(c) に対する Hard Decision Feature は図 13(d) のように頭



図 9 Soft Decision Feature による追跡結果例

(a) 直立 (正面) 時, (b) 直立 (横) 時, (c) 座り込み時, (d) 直立 (横) 時, (e) 直立 (正面) 時の追跡結果



図 10 Hard Decision Feature による追跡結果例

(a) 直立 (正面) 時, (b) 直立 (横) 時, (c) 座り込み時, (d) 直立 (横) 時, (e) 直立 (正面) 時の追跡結果



図 11 ものを拾うシーンにおける追跡結果例



図 12 走り出すシーンにおける追跡結果例

部を中心に弱識別器が選択されており、背中付近に分布する弱識別器が選択されていない。そのため、頭部を中心として人体の左側に追跡矩形がずれている。

これらの違いは弱識別器の確率密度関数の更新に大きく影響していると考えられる。提案手法の場合は人体全体を捉えることで弱識別器の確率密度が正しく更新されるため、座り込みによる姿勢変化が生じた後も正しく人体を追跡することができる。一方、Hard Decision Feature は、座り込み時に人体全体を正しく捉えることができていないため、弱識別器の確率密度関数が正しく更新されず、人体を追跡することができなくなっている。

このように、Soft Decision Feature は姿勢変化に対して、適応的に弱識別器の形状を変えるため、姿勢変化が生じるようなシーンでも正しく人体をトラッキングすることが可能となる。

5.2 計算コストの比較

Online Real Boosting による追跡はオンラインでの識別器の学習と学習した識別器による追跡の 2 つから構成される。本節では、それぞれの処理過程における計算コストについて、Hard Decision Feature を用いて姿勢変動に対応させた場合と比較する。Hard Decision Feature により姿勢変動へ対応させるためには、姿勢ごとの識別器をあらかじめ用意する必要がある。ここでは、姿勢パターンを直立と座り込みの 2 つとし、姿勢ごとの識別器は同数と仮定し N 個とする。また、オンラインで用いるセレクトラ数を S とする。全セレクトラにおいて、弱識別器候補の特徴量計算量は N であり、オンラインでの学習にかかる計算コストはセレクトラ数の数に関わらず $O(2 * N)$ となる。次に、識別器による追跡過程での計算コストは、パーティクル数を P とした場合、 $O(2 * P * S)$ となる。

一方、提案手法の場合、各 Soft Decision Feature が持つ形状変化パターン数を M とした場合、オンラインでの学習にか

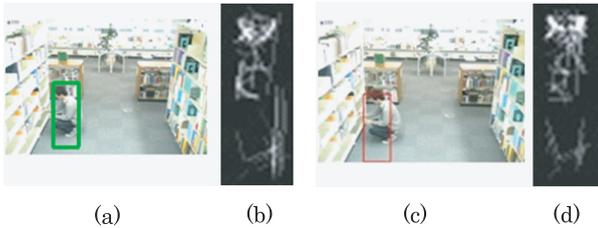


図 13 選択された弱識別器を合成した人体

(a)Soft Decision Feature による座り込み時の追跡結果, (b) 座り込み時に選択された Soft Decision Feature による人体合成モデル, (c)Hard Decision Feature による座り込み時の追跡結果, (d) 座り込み時に選択された Hard Decision Feature による人体合成モデル

かる計算コストは $O(N + M * S)$, 追跡過程での計算コストは $O(P * S)$ である. 提案手法は弱識別器候補数 N を 500, セレクタ数 S を 50, パーティクル数 P を 100 としている. これらのパラメータ設定の場合, Hard Decision Feature のオンライン学習の計算コストは $O(1000)$, 追跡過程の計算コストは $O(10000)$ である. 次に, 提案手法のオンライン学習の計算コストは $O(500 + 50M)$, 追跡過程の計算コストは $O(5000)$ である. これより, 追跡過程の処理速度は提案手法の方が 50% 高速である. また, オンライン学習は形状変化パターン数 M が 10 個以下であれば, 提案手法の方が高速である. 本稿での人体追跡の場合, 形状変化パターンは平均 6 個であり, オンライン学習の処理速度も Hard Decision Feature より 40% 高速である. これより, 提案手法は, 姿勢ごとの Hard Decision Feature を用いた場合より高速であることがわかる.

6. ま と め

我々は, 追跡対象の形状変化に対して, 適用的に弱識別器の形状を変えることができる Soft Decision Feature を用いたオフライン学習およびオンライン学習の枠組みについて提案した. 提案手法は Hard Decision Feature より追跡位置精度が向上することで, 追跡対象の姿勢変化に対してロバストに追跡することができた.

また, Soft Decision Feature は, 変形パラメータの設定範囲によって追跡対象の拡大・縮小や回転へ対応することが可能な弱識別器セットをグルーピングすることも可能である. 今後は, このような Soft Decision Feature に含まれる弱識別器の情報をもとに姿勢推定へ応用していく.

文 献

- [1] Schapire, R.E. and Singer, Y.: Improved Boosting Algorithms Using Confidence-rated Predictions, Machine Learning, No.37, pp.297-336, 1999.
- [2] Viola, P. and Jones, M.: Robust real-time face detection, International Journal of Computer Vision, Vol.57, no.2, pp. 137-154, 2004.
- [3] Wu, B., Ai, H., Huang, C., Lao, S: Fast Rotation Invari-

ant Multi-view Face Detection Based on Real AdaBoost, In Proc. IEEE Conf. on Automatic Face and Gesture Recognition, 2007.

- [4] Zhu, Q., Yeh, M.-C., Cheng, K.-T., and Avidan, S.: Fast Human Detection Using Cascade of Histogram of Oriented Gradients, In Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Vol.2, pp.1491-1498, 2006.
- [5] Oza, N. and Rusell, S.: Online bagging and boosting, In Proc. Artificial Intelligence and Statistics, pp. 105-112, 2001.
- [6] Grabner, H. and Bischof, H.: On-line Boosting and Vision, In Proc. IEEE Conf. on Computer Vision and Pattern Recognition, vol.1, pp. 260-267, 2006.
- [7] 山下, 芳, 川出: Real AdaBoost を利用した Online Boosting による人物追跡, CVIM 報告, vol.2007, No.31, pp.85-92, 2007.
- [8] Hou, C., Ai, H., Lao, S.: Multiview Pedestrian Detection Based on Vector Boosting, 8th Asian Conference on Computer Vision, 2007.
- [9] Wu, B. and Nevatia, R.: Detection of multiple, partially occluded humans in a single image by Bayesian combination of edgelet part detectors, In Proc. IEEE Conf. on International Conference of Computer Vision, vol.1 pp. 90-97, 2005.