

第6章

実利用シーンに対する再学習の効率化

5章では、統計的学習法に基づく人検出において、学習サンプル収集に伴う人的コストを削減し、特定環境への適応を自動化する手法について述べた。しかしながら、再学習を行う際に学習に要する時間コストが大きな問題として残る。学習サンプルの追加と特徴空間を少数に抑えて省力化する手法として Boosting に基づく転移学習が提案されている。しかし、事前学習シーンと特定シーンの間に大きな変化がある場合、転移学習を用いて適応するのは困難である。そこで本章では、転移学習の高速性と再学習の高精度を両立するハイブリッド型転移学習に取り組む。6.1.1において転移学習について述べた後、6.1.3においてその問題点について述べる。6.2では学習効率に基づいて特徴空間を選択的に切り替えるハイブリッド型転移学習を提案する。これにより、追加する学習サンプル数の削減と学習の高速化による効率化が期待できる。

6.1 物体検出における転移学習

5章で述べた生成型学習では、CGを用いて学習サンプルの自動収集が可能である。これにより学習サンプル収集に要する人的コストの削減を実現したが、再学習を行う場合、識別器の学習に時間を要することが問題として残されている。識別器の学習時間は学習サンプル数と特徴次元の数に比例する。特に、生成型学習で様々な視点の学習サンプルを大量に生成した場合、学習時間を要することになる。設置環境などの特定シーンに対して短時間で学習を行う手法として、事前に類似した問題に対して構築された識別器を応用する転移学習が提案されている。本節では、転移学習の従来法とその問題点について述べる。

6.1.1 転移学習

転移学習とは機械学習の分野で利用される学習手法の一つである。転移学習という言葉には様々な解釈があるが、機械学習の分野においてはNIPS 2005 workshop - inductive transfer: 10 years laterの論文募集要項で、「新規タスクの効果的な仮説を効率的に見つけ出すために、一つ以上の別のタスクで学習された知識を得て、それを適用する問題」と定義されている。この定義に従う転移学習は、事前ドメインとよばれる事前に学習した識別器や学習サンプルを、目的ドメインと呼ばれる新たな環境に対する識別器の構築に利用することで効率化を目指している。転移学習による効率化として、人検出における俯角の変化への対応 [36] や新規サンプルの追加 [37] などが提案されている。転移学習を用いた研究は2010年代以降増加しており、SOINN[37] や TRAdaBoost[38], [39], CovBoost[36] などの手法が提案されている。以下では、共変量シフトに基づく転移学習を Boosting に導入した Covariate-shift Boost (CovBoost) を転移学習について述べる。

6.1.2 Covariate-shift Boost(CovBoost) による転移学習

Covariate-shift Boost(CovBoost) とは、事前学習によって学習した弱識別器の情報と事前学習に用いた学習サンプルを目的とする特定シーンの学習に用いることで、特定シーンにおける学習を効率化する Boosting 手法である。特定シーンでの学習サンプルが少量でも、事前シーンの大量の学習サンプルを用い、大量のデータを用いて全特徴量を探索した時と同程度の検出精度を維持することが可能となる。これにより、事前シーンの学習サンプルと検出対象である特定シーンの学習サンプルに見えの変化がある場合の再学習の省力化を実現した [36]。その後、半教師型オンライン学習による人検出に展開されている [40]。CovBoost の大きな特徴は特徴の転移と、共変量を用いた転移学習アルゴリズムの2点である。

特徴の転移

CovBoost は、学習の前準備として一度学習し、識別器として採用された有効な特徴量から、目標学習にも有効な特徴量を再生成する。これを特徴の転移という。特徴の転移は、事前学習によって選択

された特徴量を目的ドメインに転移させることで行われる。文献 [36] では、Haar-like 特徴量の転移を対象としている。学習により選択された Haar-like 特徴量の領域を示す座標 4 点を乱数を用いて変動し、特徴の転移候補として生成する。これを再度学習することで有効に転移した領域が選択される。有効である可能性が高い領域から生成するため、通常の学習と比べて高速となる。

共変量を用いた転移学習

Boosting 手法は、一般的に式 (6.1) に示す損失関数 L を最小化する強識別器 $H(\mathbf{x})$ を求めることが目標となる。ここで、 Ω は学習サンプルの総数、 \mathbf{x} と y はそれぞれ学習サンプルとクラスラベルを示す。

$$L = \sum_{\Omega} e^{-yH(\mathbf{x})} \quad (6.1)$$

CovBoost では入力する全学習サンプル集合 Ω に、事前学習で用いた学習サンプル (事前ドメイン) と再学習に用いる学習サンプル (目標ドメイン) の二種類が存在する。特徴量により観測される事前ドメインの確率密度分布 $p_a(\mathbf{x})$ は、目標ドメインの確率密度分布 $p_t(\mathbf{x})$ と異なる分布を持つ可能性がある。すなわち、 $p_a(y|\mathbf{x}) \neq p_t(y|\mathbf{x})$ であると考えられる。ここで、事前ドメインの学習サンプルのうち、目標ドメインのネガティブサンプルに類似したポジティブサンプルなど目標ドメインの学習を阻害するサンプルを、学習に利用しないことが必要となる。そこで、 $\frac{p_t(y|\mathbf{x})}{p_a(y|\mathbf{x})}$ として事前ドメインのサンプルに重みを与えることで目標ドメインの学習に有効なサンプルを取捨選択する。ここで $\frac{p_t(y|\mathbf{x})}{p_a(y|\mathbf{x})}$ は共変量と呼ばれ、共変量を用いて事前ドメインのサンプルを目標ドメインに挿入することを共変量シフトによる転移学習と呼ぶ。また、共変量 $\frac{p_t(y|\mathbf{x})}{p_a(y|\mathbf{x})}$ は記号 λ で表す。共変量シフトを用いた転移学習の目的関数を式 (6.2) に示す。ここで、 $(\mathbf{x}_i, y_i) / T_t$ と $(\mathbf{x}_j, y_j) / T_a$ はそれぞれ目標ドメイン t と事前ドメイン a の学習サンプルの特徴ベクトル \mathbf{x} とそのクラスラベル y を示す。

$$\tilde{L} = \sum_{(\mathbf{x}_i, y_i) \in T_t} e^{-y_i H_t(\mathbf{x}_i)} + \sum_{(\mathbf{x}_j, y_j) \in T_a} \lambda_j e^{-y_j H_t(\mathbf{x}_j)} \quad (6.2)$$

共変量 λ は、ベイズ推定を用いて次のように変形することができる。

$$\begin{aligned} \lambda &= \frac{p_t(\mathbf{x}, y)}{p_a(\mathbf{x}, y)} \\ &= \frac{p(\mathbf{x}, y, t)}{p(\mathbf{x}, y, a)} \\ &= \frac{p(t|\mathbf{x}, y)p(\mathbf{x}, y)}{p(t)} \frac{1}{\frac{p(a|\mathbf{x}, y)p(\mathbf{x}, y)}{p(a)}} \\ &= \frac{p(t|\mathbf{x}, y)p(a)}{p(a|\mathbf{x}, y)p(t)} \end{aligned} \quad (6.3)$$

$p(a) = p(t)$ と仮定すると、式 (6.3) は目標ドメインの条件付き確率と事前ドメインの条件付き確率の比となる。ここで、条件付き確率を次式のようなロジスティック関数に変形する。

$$p(t \mid \mathbf{x}, y) = \frac{1}{1 + e^{yH_t(\mathbf{x})}} \quad (6.4)$$

$$p(a \mid \mathbf{x}, y) = \frac{1}{1 + e^{yH_a(\mathbf{x})}} \quad (6.5)$$

ここで $H_a(\mathbf{x})$ とは、事前学習によって構築された強識別器を示し、 H_t は事前ドメインと目標ドメインの両方を使って学習された強識別器を示す。この式 (6.4) と式 (6.5) から λ は式 (6.6) のように変形できる。

$$\lambda = \frac{1 + e^{-yH_a(\mathbf{x})}}{1 + e^{-yH_t(\mathbf{x})}} \quad (6.6)$$

λ は事前ドメインのサンプルが目標ドメインにどれだけ適合しているかを識別器 H_a 、 H_t により表現し、目標ドメインに適合しているほど大きな値となる。

6.1.3 転移学習の問題点

Pang 等は、視点の異なる環境において CovBoost により識別器を転移することで、新規に収集する学習サンプルを 3 分の 1 に削減しても同等の性能を実現した。しかし、カメラの俯角が大きく変化する場合、図 6.1 に示すように画像上の見えが大きく変わるため、特徴の転移が困難となり、識別性能の低下を招く。図 6.2 は、カメラ俯角を $30^\circ \in 50^\circ$ に変化した際の転移学習による識別性能を DET により示したものである。事前学習には俯角 20° の学習サンプルを、特徴量には HOG を用いた。図 6.2 より、俯角を大きく変化すると性能が低下していることがわかる。これは図 6.1 の 50° のようにサンプルの見えが大きく変化すると、転移した特徴のみでは目標ドメインにうまく適合できないことを示している。

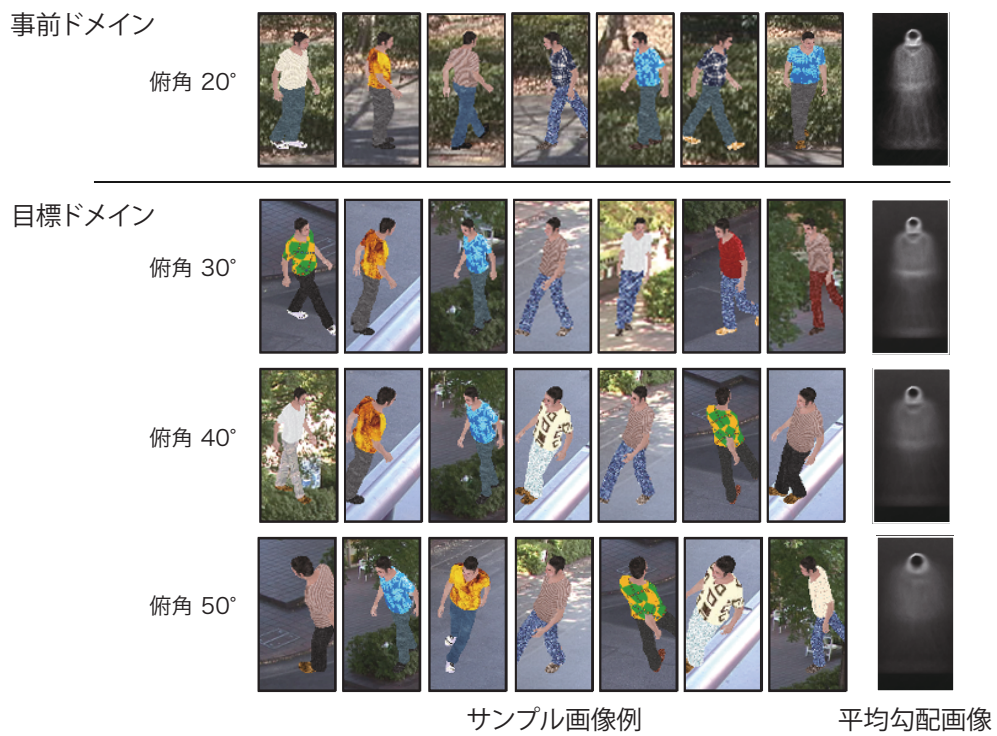


図 6.1: 俯角による学習サンプルの変化.

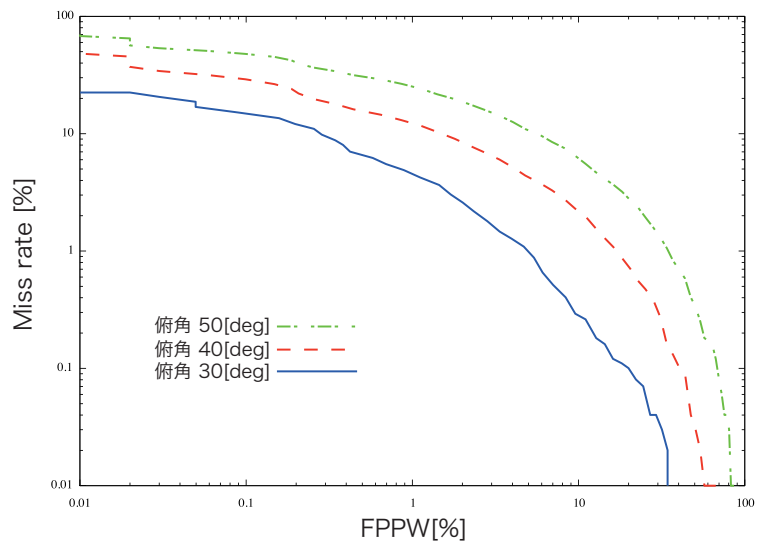


図 6.2: 俯角の変化による転移学習の性能.

6.2 ハイブリッド型転移学習による学習の効率化

転移学習を用いることで少量の目標学習サンプル収集で高精度な識別が可能となるが、事前学習と目標学習でのシーンが大きく異なる場合は適応できない。これは、シーンの大きな変化により特徴量の転移が困難となるためである。そこで本研究では、図 6.3 に示すように転移により得られる高速に探索可能な特徴量と、探索に時間を要する全ての特徴量をそれぞれ特徴空間として用意し、学習効率に基づいて転移特徴空間と全特徴空間を選択的に切り替えるハイブリッド型転移学習を提案する。事前に用意した大規模なデータセットを用いて事前学習を行い、事前識別器 H_a を得る。この H_a から特徴の転移によって得られる特徴空間を転移特徴空間 F_{Tr} とし、共変量を用いて目標ドメインと事前ドメインの両学習サンプルから学習を行い、目標識別器 H_t を構築する。これを転移学習ステップとする (図 6.3(a))。この際、学習サンプルに対するエラー率から学習効率を算出する。そして、その効率が低ければ、転移特徴のみで十分な成果が得られないとして、全探索特徴空間 F_{Re} より学習を行う再学習ステップ (図 6.3(b)) へ移行する。このようにハイブリッド型転移学習を定義し、再学習より高速かつ、従来の転移学習より高精度な識別器を構築することを目的とする。

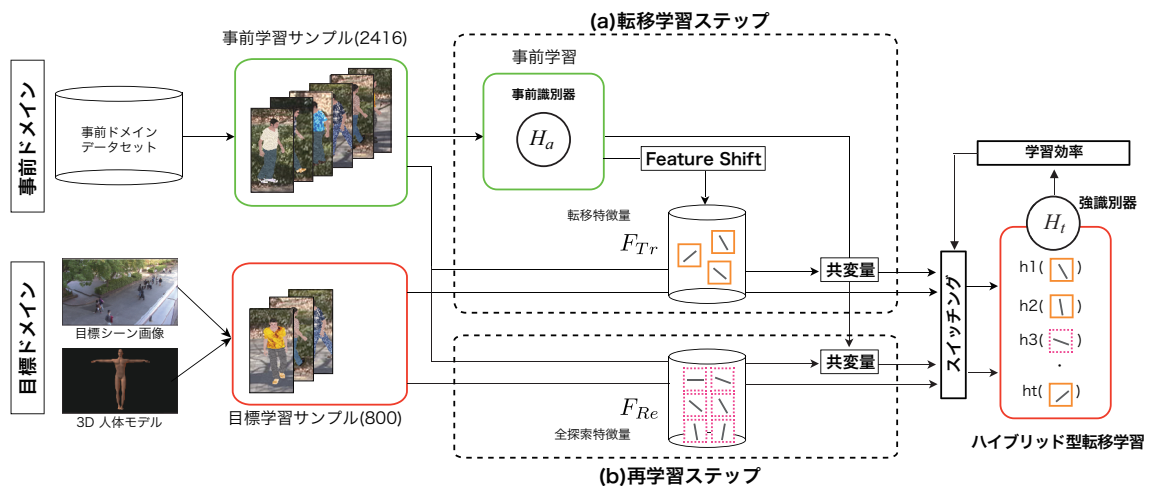


図 6.3: ハイブリッド型転移学習.

6.2.1 本研究の問題設定

本研究では、俯角の少ない人画像データの学習を事前ドメインとし、実際の設置環境である俯角の大きい特定シーンのデータを目標ドメインと定義する。

事前ドメイン

事前ドメインの学習はオフライン処理が可能であるため、大量のデータを使用する。学習サンプル

のポジティブサンプルは5章で述べた, CGで生成した人画像を用いるため, サンプル収集のコストはないものとする. 事前ドメインとしては俯角 20° の人画像 2,416 枚を使用し, HOG 特徴量による AdaBoost で学習した識別器を使用する.

～ポジティブサンプル数: 2,416 枚

～ネガティブサンプル数: 12,180 枚

～シーン: 俯角 20°

目標ドメイン

目標ドメインは, カメラの俯角の異なる特定シーンを使用する. 事前ドメインのデータと異なり, 新規に収集する必要がある点から, サンプル数が少ないことが望ましい. 目標ドメインとして, カメラ俯角を 30° , 40° , 50° として各俯角を 800 枚生成した 3 種類を用意する.

～ポジティブサンプル数: 800 枚

～ネガティブサンプル数: 12,180 枚

～シーン: 俯角 30° , 40° , 50°

本研究では, 上記に定義した問題に対して転移学習を行う. 再学習に比べ, 少数 (文献 [36] より事前ドメインの $1/3$) の目標ドメインサンプルを用いて, 精度を維持しながら高速に識別器を学習することを本研究の問題設定とする.

6.2.2 HOG 特徴量における Feature Shift

文献 [36] の CovBoost では Haar-like 特徴量の転移を行う. 本研究では, 人検出に用いられる HOG 特徴量に対して特徴の転移を行う. まず, 図 6.4(1) のように, 事前学習で選択された弱識別器の局所特徴量の中心座標を求める. この座標を中心に, 図 6.4(2) のように正規乱数に従い K 個の候補領域を発生させる. 文献 [36] では, $K = 50$ を妥当としている. ここで, 提案手法では, 候補領域から局所特徴量のヒストグラムを求め, 図 6.4(3) のように事前学習で選択された弱識別器の局所特徴量のヒストグラムと比較して両ヒストグラムの類似度を求める. ヒストグラムの類似度算出には次式に示すバタチャリヤ係数を用いる.

$$Bhattacharyya = \sum_{i=1}^n \sqrt{p(\mathbf{x})q(\mathbf{x})} \quad (6.7)$$

ここで $p(\mathbf{x})$ と $q(\mathbf{x})$ はそれぞれ異なるドメインの確率密度関数である. 最後に, 事前学習で選択された弱識別器と最も高い類似度を持つ転移候補を転移先の特徴量とし, その集合を転移特徴空間 F_{Tr} と定義する. これに対し, 再学習と同様に画像から全特徴量を抽出したものを全特徴空間 F_{Re} と定義する. 表 6.1 に各俯角ごとの類似度の平均値を示す. 6.1.3 で述べたように, 事前ドメインの学習データから離れるほど類似度は低くなる. それらの差が図 6.2 のように識別性能を低下させる一因であるといえる.

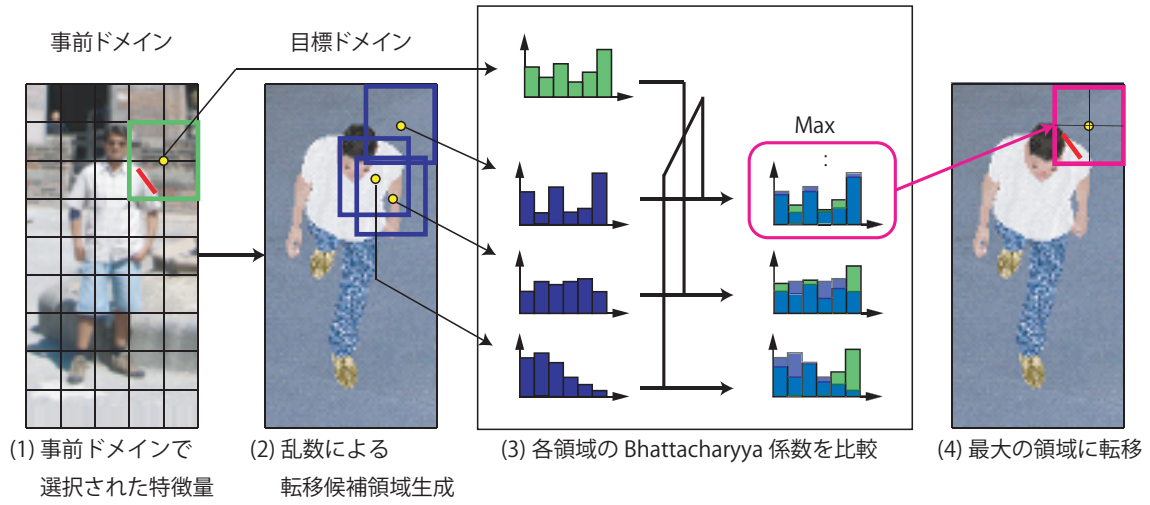


図 6.4: HOG 特徴量の転移.

表 6.1: 俯角による類似度の変化.

カメラ俯角 [°]	30	40	50
Bhattacharyya 係数	0.981	0.972	0.968

6.2.3 ハイブリッド型転移学習による学習

提案するハイブリッド型学習では、事前ドメイン T_a と目標ドメイン T_t の異なるシーンから切り出したサンプル群を用いる。目標ドメインの各サンプル \mathbf{x}_i と事前ドメインの各サンプル \mathbf{x}_j はクラスラベル y_i または y_j を持ち、ポジティブサンプルには +1, ネガティブサンプルには -1 を設定する。次に、サンプルの重みを初期化する。サンプルの重みは目標ドメインと事前ドメインのそれぞれで正規化したものを初期値とし、それぞれの重みを $D_t(\mathbf{x}_i)$ と $D_a(\mathbf{x}_j)$ と表現する。弱識別器の選択は、式 (6.2) を最小化するような $h(\mathbf{x})$ を求めることである。ここで、式 (6.2) を $h(\mathbf{x}) = 0$ の一次のテイラー展開により式 (6.8) の近似式を求める。

$$h_t(\mathbf{x}) = \arg \min_{h_t} \left(\sum_{(\mathbf{x}_i, y_i) \in T_t} e^{-2y_i D_t(\mathbf{x}_i)} y_i h_t(\mathbf{x}_i) + \sum_{(\mathbf{x}_j, y_j) \in T_a} \lambda_j e^{-2y_j D_a(\mathbf{x}_j)} y_j h_t(\mathbf{x}_j) \right) \quad (6.8)$$

この式を最小にする $h(\mathbf{x})$ を求めることで弱識別器 $h_t(\mathbf{x})$ を選択する。各弱識別器 $h(\mathbf{x})$ は、転移特徴空間 F_{Tr} と全特徴空間 F_{Re} を探索して求める。まず転移特徴空間 F_{Tr} を探索し、得られた弱識別器のエラー率から学習効率 ζ を算出する。その値がしきい値以下のとき、全特徴空間 F_{Re} において弱識別器の再選択が行われる。 ζ の算出法については次節で示す。次に、式 (6.9) でエラー率 ϵ_m を

算出する.

$$\epsilon_m = \frac{\sum_{h(\mathbf{x}_i) \neq y_i} e^{-2y_i D_t(\mathbf{x}_i)} + \sum_{h(\mathbf{x}_j) \neq y_j} \lambda_j e^{-2y_j D_a(\mathbf{x}_j)}}{\sum_i e^{-2y_i D_t(\mathbf{x}_i)} + \sum_j \lambda_j e^{-2y_j D_a(\mathbf{x}_j)}} \quad (6.9)$$

算出したエラー率から, 選択した弱識別器に対する重み α_m を次式により算出する.

$$\alpha_m = \frac{1}{4} \ln \frac{1}{\epsilon_m} \quad (6.10)$$

最後に, 目標ドメイン及び事前ドメインの学習サンプルの重み $D_t()$ と $D_a()$ を次式により更新する.

$$D_t(\mathbf{x}_i) = D_t(\mathbf{x}_i) e^{-2y_i \alpha_t h_m(\mathbf{x}_i)} \quad (6.11)$$

$$D_a(\mathbf{x}_j) = D_a(\mathbf{x}_j) e^{-2y_j \alpha_t h_m(\mathbf{x}_j)} \quad (6.12)$$

以上の処理を, 事前学習の学習回数と同数繰り返して, 新たな弱識別器 $h_t(\mathbf{x})$ を求める. 識別時は, 通常の AdaBoost と同様に, 次式に示すような弱識別器の重み付きの多数決を求めることにより識別を行う. 強識別器を式 (6.13) に示す.

$$H_t(\mathbf{x}) = \text{sign} \left[\sum_{t=1}^M \alpha_t h_t(\mathbf{x}) \quad th \right] \quad (6.13)$$

ここで th はしきい値であり, M は事前学習の学習回数である.

6.2.4 学習効率 ζ に基づく特徴空間選択

前節で述べた特徴転移では, 転移尤度の高い特徴量の転移特徴空間 F_{Tr} において転移学習を行う. これにより学習時間の削減 (探索コストの低下) が可能となるが, 事前学習データと目標学習データに大きな相違がある場合, 共通な特徴が観測されないことが考えられる. そこで, 提案手法では転移特徴空間と, 再学習同様の全特徴空間を選択的に切り替える. 転移が有効な場合には高速な転移学習を行い, 転移が困難な場合には全特徴空間を用いて学習を行う. 本研究では, それぞれの特徴空間を以下のように定義する.

転移特徴空間

～特徴次元: 100 (事前学習により選択)

～弱識別器のしきい値探索: 必要なし

～計算コスト: 低

～ドメイン間の差が大きいと性能低下

全特徴空間

～特徴次元：3,780

～弱識別器のしきい値探索：各次元ごとに 100 段階

～計算コスト：高

～目標ドメインに最適化

ここで必要となる切り替え指標は、転移学習で補えるか否かを判定する値である。図 6.5(a) に転移学習時と再学習時の Boosting における弱識別器のエラー率 ϵ (式 (6.9)) の推移を示す。

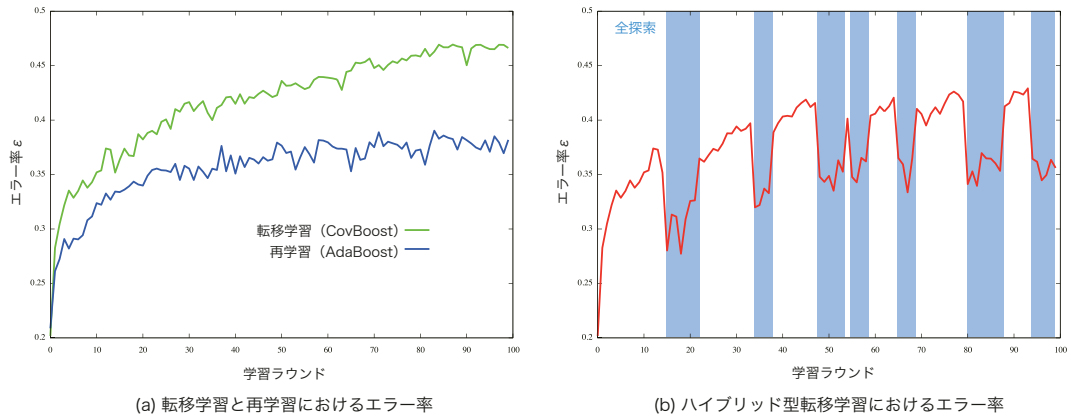


図 6.5: 弱識別器のエラー率の推移 ϵ .

弱識別器のエラー率は転移学習，再学習共に学習が進むにつれ高くなる。これは，学習が進行するにつれ適応困難なサンプルの重みの値が高くなるためである。この値の変化量が小さくなると，ある程度学習は収束し，大きな改善はないと考える。そこで，本研究ではエラー率 ϵ の勾配を学習効率 ζ として定義し，切り替えの指標として利用する。このエラー率の傾きを観測し，転移学習が進行して傾きが緩やかになり，その絶対値がしきい値を下回る際に全探索を適用する。なお，勾配の算出には直近 5 点の最小二乗近似を用いる。全探索により有効な特徴を発見できれば大幅に ϵ が下がるため勾配が拡大し，再度転移学習へと移行する。学習効率を用いた切り替えを行った際のエラー率の推移を 6.5(b) に示す。図より，全探索へ移行する度に大幅にエラー率が改善され，十分に ϵ が下がると転移学習に再度移行することがわかる。また，学習の終了条件は学習サンプルに対する誤識別率の収束を用いる。一般には誤識別率に加えエラー率の収束も用いられるが，本手法ではエラー率が大きく振動的な振る舞いをするため終了条件としては用いない。

6.3 評価実験

提案手法の有効性を示すために、識別精度と学習速度の観点から検証を行う。

6.3.1 実験概要

HOG 特徴量と AdaBoost による識別器を事前学習とする。HOG 特徴量のセルサイズは 8、ブロックサイズは 2 とすると計 3,780 次元となる。転移時に生成する特徴候補数は 1 特徴量あたり 50 とする。事前学習サンプルのポジティブサンプルとしては、図 6.1 で示した俯角 20° の CG を用いて生成した人画像を 2,416 枚使用し、ネガティブサンプルとしては、ポジティブサンプルを生成したシーンの背景画像から 12,180 枚切り出して使用する。目標学習のポジティブサンプルとしては、図 6.1 でそれぞれ示した俯角 30~50° の CG を用いて生成した人画像を 800 枚用い、ネガティブサンプルとして、事前ドメイン同様に切り出した背景画像 12,180 枚を用いる。ただし、比較手法である再学習については CG を用いて生成した人画像を 2,416 枚使用し、事前学習を行わない。評価用サンプルはそれぞれ、ポジティブサンプルには特定シーンの俯角に設定した 3D 人体モデルによる CG 生成画像 10,000 枚を、ネガティブサンプルには生成に用いた背景画像から 10,000 枚切り出したものを使用する。性能評価は EER (Equal Error Rate) により行う。

6.3.2 俯角変化に対するハイブリッド型転移学習の精度

表 6.2 に、カメラ俯角を変化させた 3 種の目標シーンに対する各手法の識別性能を示す。

表 6.2: EER による性能評価。

カメラ俯角 [°]	提案手法 [%]	転移学習 [%]	再学習 [%]
30	2.24	2.61	2.10
40	2.52	4.78	2.39
50	2.36	4.23	2.35

表より、従来の転移学習では、シーンの変化とともに性能が大きく低下することがわかる。一方、提案手法は大きなシーン変化が生じてても、再学習と同等であることがわかる。これは、転移特徴量のみでは対応できなかった点を、再学習と同等の全探索特徴により改善できたといえる。

6.3.3 考察

図 6.6 は提案手法で選択された特徴量のうち、転移学習ステップで選ばれたものを (a) 転移特徴量、再学習ステップで選ばれたものを (b) 全探索として可視化したものである。図 6.6(a) より、転移特

微では標準的な肩のエッジや脚部の縦方向のエッジなどが転移できていることがわかる。一方、図 6.6(b) の全探索では横エッジが目立ち、俯角の変化により発生した上半身の見えの変化に適応した特徴が選択されている。図 6.6(c) に提案手法全体として (a) と (b) を重ねたものを、(d) に再学習で選択された特徴量を表示する。両者を比較すると、特徴の位置関係や勾配方向が類似していることから、(c) の提案手法は転移特徴と全探索の組み合わせにより (d) の再学習に近い特徴の構成を獲得することで高精度な転移学習を実現したといえる。

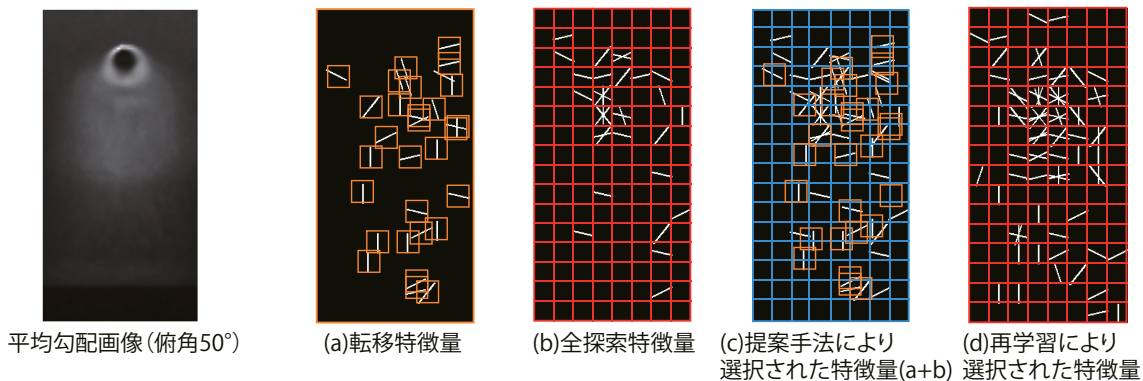


図 6.6: 選択された HOG 特徴量.

6.3.4 ハイブリッド型転移学習の高速性

提案手法は転移学習と同様の少サンプルで、再学習に近い精度を発揮する。しかし、目標シーンへの適応のための再学習にはサンプルの収集コストの他に学習に要する時間コストの問題も存在する。そこで、提案手法と従来の転移学習、再学習に要する時間の比較を表 6.3 に示す。

表 6.3: 目標識別器学習に必要な計算時間の比較.

カメラ俯角 [°]	提案手法 [min]	転移学習 [min]	再学習 [min]
30	12.1	0.58	60
40	15.2	0.58	60
50	13.6	0.58	60

表 6.3 より提案手法は、従来の転移学習に比べ 20 倍以上の時間を必要とする。これは、全探索ステップ 1 回が転移学習ステップ 1 回に比べ大きく処理時間を必要とするためである。しかし、表 6.2 で示した通り提案手法は従来法より高精度である。一方、再学習に対しては、提案手法では転移学習法における特徴の転移により、同等精度のまま 4~5 倍の高速化を実現している。これにより、さ

らなるデータ追加やカスケード構造など、長時間の学習を必要とする問題に、本手法を適用することで効率化が期待できる。

6.4 まとめ

本章では、まず、効率的な再学習として転移学習の導入とその有効性を述べた後、転移学習には、事前ドメインと目標ドメインの相違が大きい場合に深刻な性能の低下を招くという問題が存在することを示した。

次に、この問題に対し、転移学習において特徴を転移させた転移特徴空間と、再学習と同様の全特徴空間の2つの特徴空間を設定し、適応的に特徴空間を選択するハイブリッド型転移学習法を提案した。弱識別器のエラー率 ϵ の勾配に基づく学習効率に従った空間選択により、従来の転移学習より同等 ϵ 2.26%性能を向上しつつ、従来の再学習に比べ約3.9倍高速な学習が可能となる。これにより、高速性と高精度を両立した効率的な学習を実現した。また、転移学習で得られた特徴量と提案手法の全探索で得られた特徴を組み合わせ可視化したところ、再学習により得られる特徴量と非常に類似した勾配が得られた。これは、転移した特徴量のみでは大きな変化に対応できないことと、その問題を全探索で選択した特徴量で補うことが可能であることを示している。この点から、転移学習では異なるドメイン間の転移に失敗するような困難な問題に対して、提案手法は全探索とのハイブリッド戦略を導入することで再学習と同様の改善を得ることができたといえる。

