

# ハイブリッド型転移学習による物体検出における学習の効率化

土屋 成光<sup>†</sup> 山内 悠嗣<sup>†</sup> 山下 隆義<sup>††</sup> 藤吉弘亘<sup>†††</sup>

<sup>†,†††</sup> 中部大学 工学部 情報工学科 〒487-8501 愛知県春日井市松本町 1200

<sup>††</sup> オムロン株式会社 〒525-0025 滋賀県草津市西草津 2 丁目 2-1

E-mail: <sup>†</sup>{tsuchiya,yuu}@vision.cs.chubu.ac.jp, <sup>††</sup>takayosi@omm.ncl.omron.co.jp, <sup>†††</sup>hf@cs.chubu.ac.jp

**あらまし** 統計的学習法に基づく人検出では、学習サンプル収集に伴う人的コストと特定シーンに合わせた再学習のための時間的コストが大きな問題である。サンプル収集を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が、事前学習シーンと特定シーン間に大きな変化がある場合、その変化に適応するのは困難である。そこで本研究では、転移によって得られる特徴量と、再学習と同様の全特徴探索の2つの特徴空間を用意し、定義する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。提案手法は転移学習の持つ少数サンプルでの高速な識別器構築を特長として持ち、再学習に比べ 3.2 倍以上高速且つ、従来の転移学習に比べ最大で 8.35% の高精度化を実現した。

**キーワード** 転移学習, 生成型学習, ブースティング, 物体検出

## Hybrid Transfer Learning for Efficient Learning in Object Detection

Masamitsu TSUCHIYA<sup>†</sup>, Yuji YAMAUCHI<sup>†</sup>,

Takayoshi YAMASHITA<sup>††</sup>, and Hironobu FUJIYOSHI<sup>†††</sup>

<sup>†</sup> Dept. of Computer Science, Chubu University Matsumoto-cho 1200, Kasugai-shi, Aichi, 487-8501 Japan

<sup>††</sup> Omron Corporation 2-2-1 Nishikusatsu, Kusatsu, Shiga, 525-0025 Japan

E-mail: <sup>†</sup>{tsuchiya,yuu}@vision.cs.chubu.ac.jp, <sup>††</sup>takayosi@omm.ncl.omron.co.jp, <sup>†††</sup>hf@cs.chubu.ac.jp

**Abstract** In the detection of human from image using statistical learning methods, the labor cost of collecting training samples and the time cost for retraining to match the target scene are major issues. One method to reduce the work involved in sample collection is transfer learning based on boosting. However, if there is a large change between the auxiliary scene and target scene, it is difficult to apply the transfer learning. We therefore propose a hybrid transfer learning method in which two feature spaces are prepared, one of feature obtained by transfer and another of full feature search that is the same as retraining. The feature space is selectively switched on the basis of the defined training efficiency. The proposed method improving accuracy up to 8.35% compared to conventional transfer learning while accelerating training time by 3.2 times faster compared to retraining.

**Key words** transfer learning, generative learning, boosting, Object detection

### 1. はじめに

統計的学習法に基づく人検出法は、これまでに多くの研究が行われ、マーケティングなどで既に実用化が始まっている [1]. 統計的学習法に基づく人検出法では、まず大量の人画像と背景画像を用意し、これらを識別するように Support Vector Machines(SVM), Boosting [2] などの識別器を学習する。特徴量としては Dalal 等が提案した Histograms of Oriented Gradients(HOG) [3] のように、勾配から人らしいシルエットを抽出する手法が用いられている [1]. また、学習用データベースのポ

ジティブサンプルとして人画像を数千枚、ネガティブサンプルとして背景画像を数万枚というオーダーで収集して学習に利用するのが一般的であり、サンプル全てにおいて人の位置ずれがなく、大きさや高さなどについて統一されていることが求められる。これは、特徴量に局所領域の勾配特徴を用いるため、人の位置ずれや画像中における人の大きさの違いが発生すると、人らしい共通の特徴が統計的に選択できなくなるからである。

この問題に対して、簡易に良質なサンプルを大量に採取する方法として、CG によるサンプル画像生成の研究が行われている [4], [5]. これらの研究では、ポジティブサンプルの位置を均

一にして大量のサンプルを得ることができるため、位置ずれがない良質な学習サンプルを準備することができる。しかし、学習サンプルを収集した環境と人検出システムが稼働するシーンが異なる場合、人の見え方が異なるため人検出性能が低下するという問題がある。例えば、カメラの俯角が異なれば同一人物であっても縦横比や部位ごとの比率など、大きく見えが変化する。この問題を解決するためには、システムが稼働する環境から人画像を収集し、識別器を再学習する必要がある。しかしながら、シーン毎に人検出器を学習するためのデータセットを作成し、学習することは大きな労働力と多大な時間を必要とする。このような問題を解決するアプローチが転移学習の利用である。転移学習は異なるバリエーションについて学習する際のサンプル収集を省力化する手法として提案されており、この手法では多数の標準的なサンプルで事前学習を行い、得られた識別器を少数の目的サンプルに適合するよう転移させることで省力化を実現している [9]。しかし、シーン間に大きな差がある事前学習から目標学習へ転移する際、再学習に比べ大きく性能が低下する。

そこで我々は、転移により得られる特徴量と、再学習と同様の全特徴探索の2つの特徴空間を用意し、定義する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、転移学習では対応困難な大きな隔たりがあるシーン間であっても高精度に、且つ再学習するより高速な識別器の構築が可能となる。

## 2. 共変量シフトによる転移学習

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 [6] の論文募集要項で定義された、「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。転移学習を用いた研究は近年になって増えており、SOINN [8] や TRAdaBoost [7], CovBoost [9] などの論文が報告されている。本研究では、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) を転移学習を用いる。

### 2.1 Covariate-shift Boost(CovBoost)

Covariate-shift Boost(CovBoost) とは、事前学習によって取得した弱識別器の情報と、学習に用いた学習サンプルを再学習に用いることで、新たなシーンにおける学習サンプルが少量でも大量の学習サンプルを用いた時と変わらない検出精度を保持する Boosting 手法である。標準的な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とに俯角変化がある場合の省力化として Pang らにより提案され、半教師型オンライン学習などにも展開されている [9], [10]。Boosting 手法は一般的に式 (1) に示す損失関数を最小化する強識別器  $H(x)$  を求めることが目標となる。ここで、 $\Omega$  は学習サンプルの総数を示す。また、ここで、 $x$  と  $y$  はそれぞれ学習サンプルの特徴量  $x$  とそのクラスラベル  $y$  を示す。

$$L = \sum_{\Omega} e^{-yH(x)} \quad (1)$$

CovBoost では入力するサンプルが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) が存在する。ここで、特徴量により観測される事前ドメインの確率密度分布  $p_a(x)$  は、一般的に目標ドメインの確率密度分布  $p_t(x)$  と異なる分布を持つ。すなわち、 $p_a(y|x) \neq p_t(y|x)$  であると考えられる。そこで、 $\frac{p_t(y|x)}{p_a(y|x)}$  で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  $\frac{p_t(y|x)}{p_a(y|x)}$  は共変量と呼ばれ、この値により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よぶ。また、共変量  $\frac{p_t(y|x)}{p_a(y|x)}$  は記号  $\lambda_j$  で表す。共変量シフトを用いた転移学習の目的関数を式 (2) に示す。ここで、 $(x_i, y_i) \in T_t$  と  $(x_j, y_j) \in T_a$  はそれぞれ目標ドメインと事前ドメインの学習サンプルの特徴量  $x$  とそのクラスラベル  $y$  を示す。

$$\tilde{L} = \sum_{(x_i, y_i) \in T_t} e^{-y_i H_t(x_i)} + \sum_{(x_j, y_j) \in T_a} \lambda_j e^{-y_j H_t(x_j)} \quad (2)$$

また、ベイズ推定を用いることで、 $\lambda$  は次のように変形することができる。

$$\begin{aligned} \lambda &= \frac{p_t(x, y)}{p_a(x, y)} \\ &= \frac{p(x, y|t)}{p(x, y|a)} \\ &= \frac{p(t|x, y)p(x, y)}{p(t)} \frac{1}{\frac{p(a|x, y)p(x, y)}{p(a)}} \\ &= \frac{p(t|x, y)p(a)}{p(a|x, y)p(t)} \end{aligned} \quad (3)$$

$p(a) = p(t)$  と仮定すると式 (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。

$$p(t|x, y) = \frac{1}{1 + e^{yH_t(x)}} \quad (4)$$

$$p(a|x, y) = \frac{1}{1 + e^{yH_a(x)}} \quad (5)$$

ここで  $H_a(x)$  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 (4) と式 (5) から  $\lambda$  は式 (6) のように変形できる。

$$\lambda = \frac{1 + e^{-yH_a(x)}}{1 + e^{-yH_t(x)}} \quad (6)$$

$\lambda$  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。

### 2.2 CovBoost の問題点

Pang 等は、視点等の異なる環境において、新規に採集する学習サンプルを3分の1に削減しても CovBoost により識別器を転移することで、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1に示すように俯角の変化につ

れてその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その性能の低下を招く。図2は、図1のようにカメラ俯角を変更した際の転移学習による識別性能の変化を Detection Error Trade-off(DET) により示したものである。事前学習には INRIA person dataset を、特徴量としては HOG を用いた。図2より、俯角を大きく変化させるにつれ性能が低下していることがわかる。これはサンプルの見えが変化していくためであり、転移した特徴のみでは大きく変化した目標ドメインにうまく適合できないことがわかる。



図1 俯角による学習サンプルの変化

Fig. 1 Variation of appearance of training samples by an angle of depression .

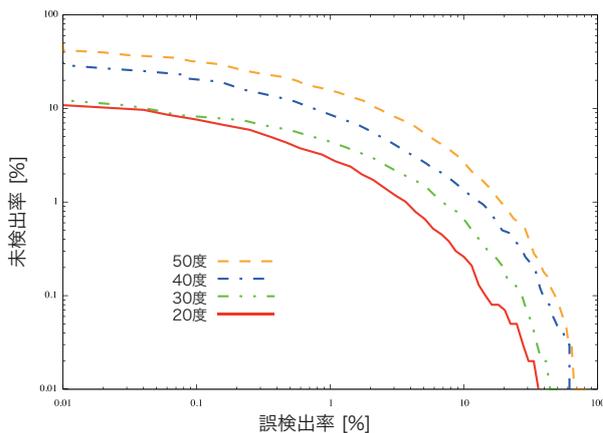


図2 俯角の変化による転移学習の性能低下

Fig. 2 Performance decrement of the transfer learning by change of an angle of depression.

### 3. ハイブリッド型転移学習

転移学習を用いることで少量の目標学習サンプル採集で高精

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適応できない。そこで我々は、図3に示すように転移により得られる特徴量と、再学習により得られる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探索空間を選択的に切り替える。これをハイブリッド型転移学習として提案し、再学習より高速に、従来の転移学習より高精度な識別器の構築を目的とする。

#### 3.1 本研究の問題設定

本研究では、標準的な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、標準的な大量のデータを使用する。本研究の事前ドメインには、標準的な人検出器用データセットである INRIA person dataset [11] の人画像 2,416 枚を使用し、HOG 特徴量による AdaBoost で学習したものを識別器として設定する。INRIA person dataset は歩行者検出のベンチマークとして広く知られており、直立した人を正面から検出するのに有効である。

- 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 0°, 正面を中心とした人画像

##### 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、サンプルが少数で済むことが望ましい。本研究では文献 [5] に基づき人画像を CG で生成することでサンプル取得を省力化する。カメラも俯角を 20°, 30°, 40°, 50° とそれぞれ変化した画像を特定シーンとして各 800 枚生成し、目標ドメインとする。

- ポジティブサンプル数：各 800 枚
- ネガティブサンプル数：12,180 枚
- シーン：俯角 20°, 30°, 40°, 50° の CG 生成画像

本研究では、以上のように定義した問題に対して転移学習を行う際、再学習時の特徴探索を選択的に取り入れることで、少数(文献 [9] より事前ドメインの 1/3) の新規サンプルを用いて精度を維持した高速な学習(識別器の構築)を行う。

#### 3.2 転移尤度に基づく Feature Shift

CovBoost は、学習の前準備として特徴の転移を行う。この特徴の転移とは、事前学習で選択された弱識別器と同様な特徴量を持つ弱識別器を再学習を行うシーンのサンプル群(目標ドメイン)から選択することである。まず、図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を求める際に正規化を行った範囲の中心座標を求める。この座標を中心に、図4(2)のように正規分布に従い  $L$  個の候補領域を発生させる。文献 [9] では、 $L = 50$  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転移尤度を求める。ヒストグラムの類似度には次式に示すバタチャリヤ距離を用いる。

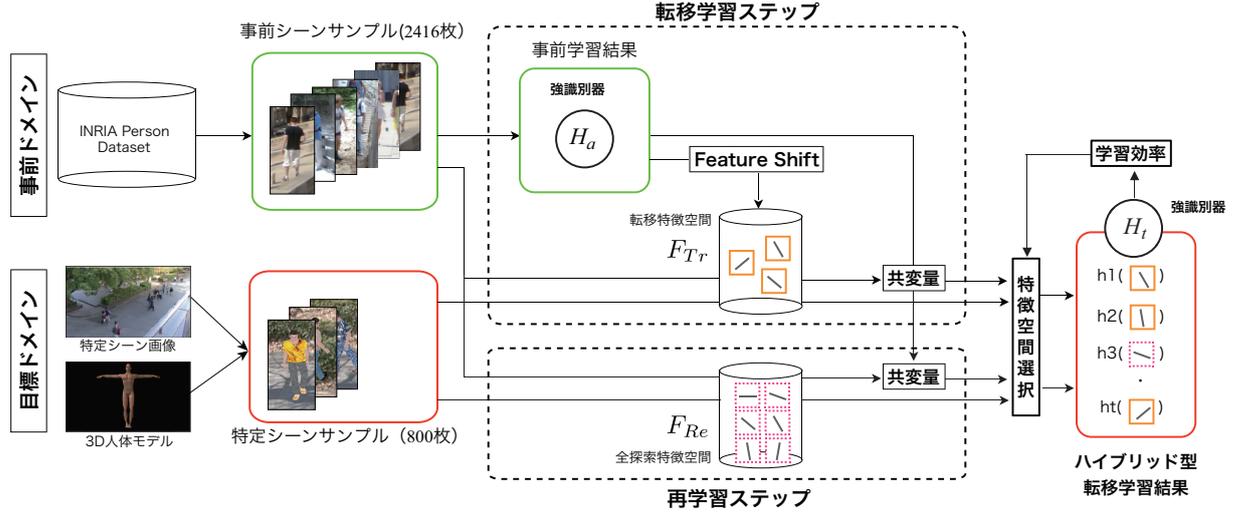


図3 ハイブリッド型転移学習

Fig.3 Hybrid transfer learning.

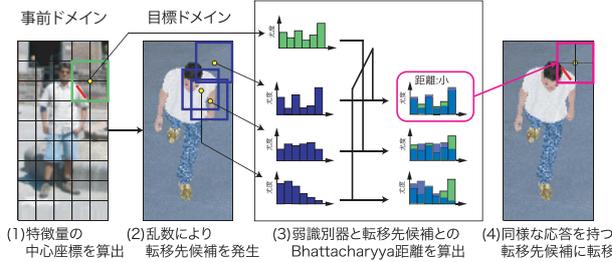


図4 HOG 特徴量の転移

Fig.4 Transferring HOG feature for Pedestrian detection.

表1 俯角による転移尤度の変化

Table 1 Variation of the transfer-likelihood by an angle of depression.

| カメラの俯角 [°]    | 20    | 30    | 40    | 50    |
|---------------|-------|-------|-------|-------|
| Bhattacharyya | 0.975 | 0.974 | 0.970 | 0.967 |

$$Bhattacharyya = \sum_{i=1}^n \sqrt{p(x)q(x)} \quad (7)$$

ここで  $p(x)$  と  $q(x)$  はそれぞれ異なるドメインの確率密度分布である。最後に、事前学習で選択された弱識別器を最も高い転移尤度を持つ転移候補へ転移させ、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探索特徴空間  $F_{Re}$  と定義する。表1に各俯角ごとの転移尤度の平均値を示す。問題設定で触れたように、事前ドメインの学習データから離れるほど転移尤度は低くなる。それらの差が図2のように識別性能を低下させる一因であるといえる。

### 3.3 ハイブリッド型転移学習による学習

ハイブリッド型学習には、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。これらのサンプルは全てクラスラベルを持ち、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  $D_t(x_i)$  と  $D_a(x_j)$  と表現する。弱識別器の選択は、式(2)を最小化するような  $h(x)$  を求めることである。ここで、式(2)を  $h(x) = 0$  の一次のテイラー展開により式(9)の近似式を求める。この式を最小にする  $h(x)$  を求めることで弱識別器を選択する。

$$h_m = \arg \min_{h_t} \left( \sum_{(x_i, y_i) \in T_t} e^{-2y_i D_t(x_i)} y_i h_t(x_i) + \sum_{(x_j, y_j) \in T_a} \lambda_j e^{-2y_j D_a(x_j)} y_j h_t(x_j) \right) \quad (8)$$

この際、各  $h()$  は転移特徴空間  $F_{Tr}$  を探索して求める。次に、式(9)でエラー率  $\epsilon_m$  を算出する。

$$\epsilon_m = \frac{\sum_{h(x_i) \neq y_i} e^{-2y_i D_t(x_i)} + \sum_{h(x_j) \neq y_j} \lambda_j e^{-2y_j D_a(x_j)}}{\sum_i e^{-2y_i D_t(x_i)} + \sum_j \lambda_j e^{-2y_j D_a(x_j)}} \quad (9)$$

ここで、学習効率  $\zeta$  を算出し、その値が閾値以下のとき、全探索特徴空間  $F_{Re}$  において弱識別器の再選択が行われる。 $\zeta$  の算出方法は次節で示す。次に選択した弱識別器に対する重み  $\alpha_m$  を式(10)で算出する。

$$\alpha_m = \frac{1}{4} \ln \frac{1 - \epsilon_m}{\epsilon_m} \quad (10)$$

次に、学習サンプルの重みを更新する。

$$D_t(x_i) = D_t(x_i) e^{-2y_i \alpha_t h_m(x_i)} \quad (11)$$

$$D_a(x_j) = D_a(x_j) e^{-2y_j \alpha_t h_m(x_j)} \quad (12)$$

以上の処理を事前学習の学習回数と同数繰り返す。最終的に、すべての弱識別器に重みを付けて多数決を取ることで識別を行う強識別器を構築する。強識別器を式(13)に示す。

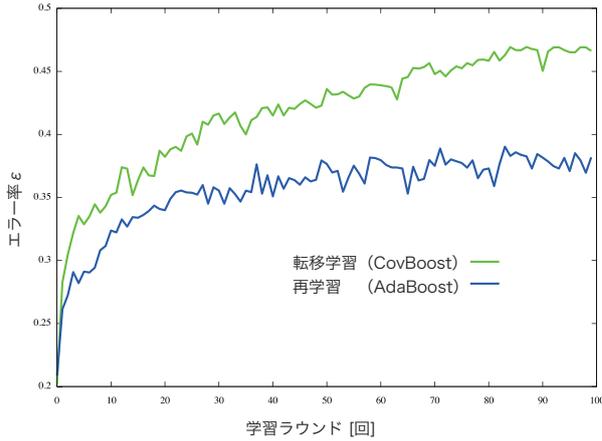


図 5 弱識別器のエラー率  $\epsilon$   
Fig. 5 Error of weak-classifier( $\epsilon$ ).

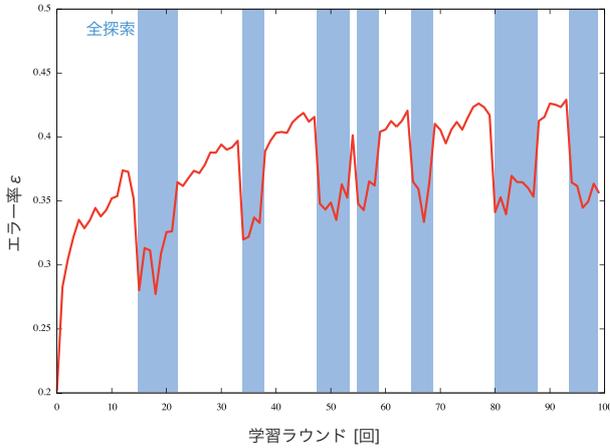


図 6 学習効率による特徴空間選択

Fig. 6 Feature space selection using value of learning efficiency.

$$H_t(x) = \text{sign}\left[\sum_{t=1}^M \alpha_t h_t(x) - th\right] \quad (13)$$

ここで  $th$  は閾値であり、 $M$  は事前学習の学習回数である。

### 3.4 学習効率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）が可能であるが、事前学習データと目標学習データに大きな変化がある場合、そもそも同様の特徴が観測されない等の場合が考えられる。そこで、提案する尤度に基づいた転移特徴空間と、再学習同様の全探索特徴空間を選択的に切り替え、転移が有効な場合には尤度に基づく高速な転移学習を、転移が困難な場合には全探索特徴空間を用いて学習を行う。それぞれの特徴空間は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）
- 弱識別器の閾値探索：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

#### 全探索特徴空間

- 特徴次元：3,780
- 弱識別器の閾値探索：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、転移学習で補いきれるか否かを判定する値である。図 5 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  $\epsilon$  の推移を示す。  $\epsilon$  は式 (9) で求めることができ、転移学習、再学習共に学習が進むにつれ高くなる。これは、学習が進行しても適応困難なサンプルの重みが高くなるためである。特に、転移学習ではシーンに大きな差がある場合識別性能の限界が低いため上昇しやすい。この値が定常的になると、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、我々はこのエラー率  $\epsilon$  の勾配を以下に示すように学習効率  $\zeta$  として定義し、切り替えの指標として利用する。

$$\zeta = \text{grad}(\epsilon) \quad (14)$$

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、その絶対値が閾値を下回る際に全探索を適用する。全探索により有効な特徴を発見できれば大幅に  $\epsilon$  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図 6 に示す。図より、全探索へ移行する度に大幅にエラー率を改善できていること、充分に下がれば転移学習に再度移行することがわかる。

## 4. 実験

提案手法の有効性を示すために、識別精度と速度の観点から検証を行う。

### 4.1 実験概要

HOG 特徴量と AdaBoost による識別器を事前学習とする。この際、HOG 特徴量をセルサイズを 8、ブロックサイズを 2 として計 3780 次元とする。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。事前学習サンプルとしては INRIA person dataset をポジティブサンプルとして人画像 2,416 枚使用し、ネガティブサンプルとして人画像 12,180 枚使用する。目標学習としてはポジティブサンプルとして俯角 20~50° 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特定シーンの俯角に設定した 3D 人体モデルを設置した CG 生成画像 10,000 枚をポジティブサンプルとして、ネガティブサンプルも背景画像を 10,000 枚使用する。性能評価は EER (Equal Error Rate) により行う。EER は誤検出と未検出が等しくなる値であり、値が低いほど高精度であることを表す。

### 4.2 特徴空間選択法による識別精度変化

ハイブリッド型転移学習の有効性を示すため、事前学習シーンと隔たりのある目標シーンに対して転移学習、再学習を行い比較する。

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

表 2 EER による性能評価.

Table 2 Accuracy evaluation by EER .

| カメラの俯角 [°]      | 20   | 30    | 40    | 50    |
|-----------------|------|-------|-------|-------|
| ハイブリッド型転移学習 [%] | 2.26 | 6.18  | 8.37  | 6.37  |
| 転移学習 [%]        | 3.85 | 10.56 | 16.72 | 15.61 |
| 再学習 [%]         | 0.08 | 1.07  | 1.45  | 1.02  |

表 3 計算時間の比較.

Table 3 Comparison of training cost.

| カメラの俯角 [°]        | 20   | 30   | 40   | 50   |
|-------------------|------|------|------|------|
| ハイブリッド型転移学習 [min] | 15.0 | 13.8 | 18.6 | 14.4 |
| 再学習 [min]         | 60   |      |      |      |
| 計算時間比 [n 倍高速化]    | 3.99 | 4.34 | 3.22 | 4.16 |

下することがわかる. 一方, このように大きなシーンの変化を受けても, 提案手法は再学習には及ばないが, 転移学習に比べ 1.59%~8.35%と大きく性能が向上していることがわかる. 再学習は事前ドメインに頼らず潤沢なサンプルで目標ドメインに適応できたため, 最も性能が高い.

### 4.3 学習速度の比較

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の計算コストも存在する. そこで, 提案手法と再学習を比較し, その学習コスト削減率を表 3 に示す. 再学習に必要な学習時間に対して, 提案手法では転移学習法における特徴の転移により 3.2~4.16 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, 大きな時間を必要とする高精度化を適用しやすいといえる.

### 4.4 考察

提案手法は転移学習で対応しきれない大きな変化に対して, 全探索で特徴量を補うことで高い精度を維持する. 図 7 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 7(a) より, 転移特徴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 7(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部部分の見えの変化に適応した特徴が選択されている. 図 7(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している.

## 5. おわりに

提案手法は, 転移学習において特徴の転移で生成される転移特徴空間と, 再学習と同様の全探索特徴空間の 2 つの特徴空間を設定し, 弱識別器の誤差  $\epsilon$  の勾配に基づく学習効率に従って適応的に特徴空間を選択することにより, 従来の転移学習より 1.59%~8.35%性能を向上しつつ再学習に比べ 3.2 倍以上高速

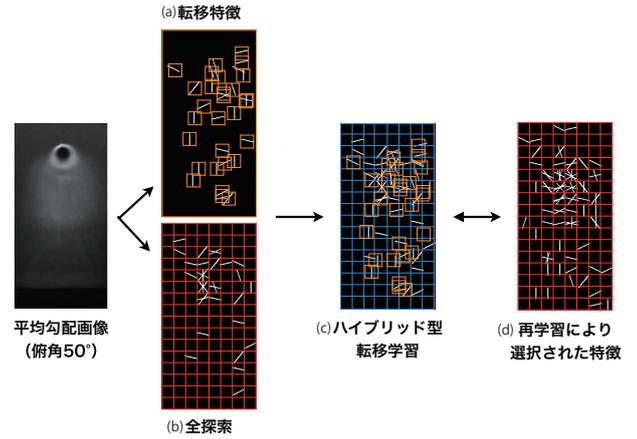


図 7 選択された HOG 特徴量 : (a) 転移特徴量, (b) 全探索, (c)(a)+(b), (d) 再学習  
Fig.7 Selected HOG Feature : (a)selected transfer-step, (b)selected retrain-step, (c)(a)+(b) selected hybrid transfer learning, (d)selected retrain

に学習可能である. 今後は, 共変量シフト以外の転移学習に対する拡張や, Real AdaBoost などより高精度な識別器への展開を行う予定である.

## 文 献

- [1] 山内悠嗣, 山下隆義, 藤吉弘亘, “[サーベイ論文] 統計的学習手法による人検出”, 電子情報通信学会 パターン認識・メディア研究会 (PRMU) 技術報告, Vol. 112, No. 197, pp. 113–126, 2012.
- [2] Y. Freund and R. E. Schapire: “Experiments with a New Boosting Algorithm”, In Proceedings of International Conference Machine Learning, (1996).
- [3] N. Dalal and B. Triggs: “Histograms of oriented gradients for human detection”, IEEE Computer Vision and Pattern Recognition, pp. 886–893, 2005.
- [4] F. M. Tur, D. Vazquez, D. Geronimo and A. M. Lopez, “Learning Appearance in Virtual Scenarios for Pedestrian Detection”, IEEE Computer Vision and Pattern Recognition, 2010.
- [5] 土屋成光, 山内悠嗣, 藤吉弘亘 “人検出のための生成型学習と Negative-Bag MILBoost による学習の効率化”, 画像の認識・理解シンポジウム (MIRU2012), 2012.
- [6] NIPS 2005 Workshop - Inductive Transfer: 10 years later. <http://iitrl.acadiau.ca/itws05/>.
- [7] W. Dai, Q. Yang, G. Xue and Y. Yu “Boosting for Transfer Learning”, International Conference on Machine Learning, 2007.
- [8] F. Shen, H. Yu, K. Sakurai and O. Hasegawa, “An Incremental Online Semi-Supervised Active Learning Algorithm based on Self-Organizing Incremental Neural Network”, Neural Computing & Applications, Springer London, pp.1–14, (2010)
- [9] J. Pang, Q. Huang, S. Yan, S. Jiang and L. Qin, “Transferring Boosted Detectors Towards Viewpoint and Scene Adaptiveness”, IEEE Transactions on Image Processing 20(5): pp.1388–1400, 2011.
- [10] G. Le, L. Qin, Q. Huang, J. Pang and S. Jiang, “Treat Samples Differently: Object Tracking with Semi-supervised Online CovBoost” 2011 IEEE Computer Vision Pattern Recognition. pp. 627–634, 2011.
- [11] INRIA Person Dataset: “<http://pascal.inrialpes.fr/data/human/>”.