

1. はじめに

人体パーツ識別には、距離画像と Random Forest を用いる手法が多く提案されている。Shotton 等によって提案された Random Forest を用いる手法 [1] は、詳細な人体パーツ位置を識別可能である。近年では、Regression Forest に頭の向きの条件を加え、複数の Forest を学習することにより、顔特徴点を検出する手法が提案されている。本アプローチは、人体パーツ識別の際に人体の向きを条件とすることで同様の効果が期待できる。そこで本研究では、Regression Forest により回帰した人体の向きを条件とした Conditional Classification Forests を用いた人体パーツ識別手法を提案する。

2. 提案手法

提案手法では、距離画像の画素毎に人体パーツのクラス確率を算出する際、Regression Forest の出力から得られる人体の向きに対する確率分布を条件として、Conditional Classification Forests により条件付き確率を算出する。人体の向きを考慮することで高精度な人体パーツ識別が期待できる。以下にそのアルゴリズムを述べる。

2.1 Regression Forest による人体の向き推定

人体向き推定のための Regression Forest の学習と回帰推定について述べる。

学習

人体モデルにより $\{-100 \sim +100\}^\circ$ 方向の範囲で 10° 毎に生成した距離画像データを図 1 に示すように、 $\omega \in \{-90, -45, 0, +45, +90\}^\circ$ の 5 方向にグループ化する。学習サンプルは、人体の距離画像とその向きのセットとなる。

Regression Forest の学習では、学習サンプルをサブセットに分割し、 T 個のサブセット毎に回帰木を作成する。回帰木の構築は、学習サンプルを分岐ノードにより分岐することで行う。各分岐ノード n に対し、入力サンプル集合 S_n 、左右に分岐するサンプル集合をそれぞれ S_l 、 S_r とすると分岐は次式により行われる。

$$S_l = \{i \in S_n | f_\theta(I) < t\} \quad (1)$$

$$S_r = S_n \setminus S_l \quad (2)$$

ここで、 $f_\theta(I)$ は入力画像 I から抽出された特徴量、 t はしきい値である。回帰木の分岐ノードにおける特徴量は、文献 [1] と同様に 2 点間の距離差を特徴量として用いる。図 1 に示すように、入力画像 I からピクセル x とオフセット量 $\theta = (u, v)$ をランダムに選択し、2 点の距離値 d の差分を次式より算出する。

$$f_\theta(I, x) = d_l \left(x + \frac{u}{d_l(x)} \right) - d_r \left(x + \frac{v}{d_l(x)} \right) \quad (3)$$

各分岐ノードでは、情報利得 (Information gain) ΔE により、その分岐を評価する。分岐ノードにおける分岐関数には、情報利得 ΔE が最大となる特徴量としきい値を用いる。末端ノードでは、トラバースルして到達したサンプルの人体の向きを表す角度の平均 $\bar{\omega}_l$ と分散 σ_l^2 を算出し、次式のように正規分布 N を求めることで、人体の向きの確率分布を保持する。

$$p(\omega | l) = N(\omega; \bar{\omega}_l, \sigma_l^2) \quad (4)$$

回帰推定

入力画像 I が与えられたとき、回帰木をトラバースルして到達した末端ノード l が保持する人体の向きの確率分布 $p(\omega | l)$ を (5) 式により統合する。

$$p(\omega | I) = \frac{1}{T} \sum_{t=1}^T p(\omega | l_t) \quad (5)$$

これにより、図 2 に示すような各末端ノードが持つ正規分布を重ね合わせた確率分布を得ることができる。以上により算出した確率分布を Conditional Classification Forests において人体パーツ識別に利用する条件の確率分布とする。

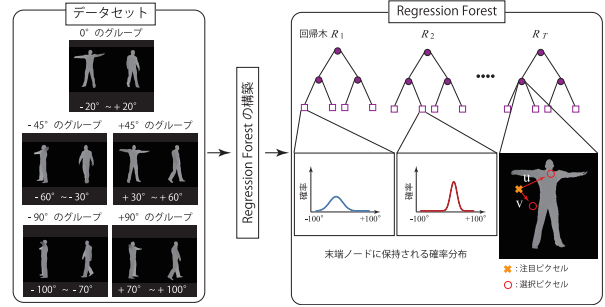


図 1: Regression Forest による学習の流れ

2.2 Conditional Classification Forests による人体パーツ識別

人体パーツ識別のための Conditional Classification Forests の学習と識別について述べる。

学習

Conditional Classification Forests の学習は、人体の向き $\omega \in \{-90, -45, 0, +45, +90\}^\circ$ の 5 方向毎に Classification Forest を構築することにあたる。人体パーツ識別における学習サンプルは、入力画像 I の各ピクセル x とそのピクセルが属するパーツのクラス番号 c のセットとなる。

Classification Forest の学習は、基本的に 2.1 節と同様である。異なる点は、学習サンプルと末端ノードの表現が異なる。まず、学習サンプルをサブセットに分割し、分岐は (1)(2) 式により行う。分類木の分岐ノードにおける特徴量も回帰木と同様に文献 [1] で用いられた 2 点間の距離差を用いる。2 点間の距離差を用いることで、各人体パーツの前後関係や形状を捉えることができる。サンプルを分岐させた後、情報利得 ΔE を求め、情報利得 ΔE が最大となる特徴量としきい値の組み合わせを分岐ノードのパラメータとして決定する。末端ノード l では、到達したサンプルのクラスラベルから、次式のようにクラス c のサンプル集合 S_c の確率を算出することで、事後確率 $p(c|l)$ を保持する。

$$p(c|l) = \frac{|S_c|}{|S|} \quad (6)$$

ここで S は全クラスのサンプル数である。以上のように、5 つの Classification Forest からなる Conditional Classification Forests を構築する。

識別

学習により構築した Conditional Classification Forests による人体パーツ識別の流れを以下に示す (図 2)。

Step1. Regression Forest より入力画像 I の人体向きを回帰推定し、(5) 式により確率分布 $p(\omega_i | I)$ を求める。

Step2. 人体の向き $\omega \in \{-90, -45, 0, +45, +90\}^\circ$ 毎の各 Classification Forest より入力画像 I の各画素 x の人体パーツクラスの事後確率 $p(c|\omega_i, I, x)$ を求める。

$$p(c|\omega_i, I, x) = \frac{1}{T} \sum_{t=1}^T p_t(c|l) \quad (7)$$

Step3. 人体向きの確率分布 $p(\omega_i | I)$ と 5 つの人体パーツクラスの事後確率 $p(c|\omega_i, I, x)$ を用いて、次式より条件付き確率を算出する。

$$p(c|I, x) = \sum_i p(c|\omega_i, I, x) p(\omega_i | I) \quad (8)$$

Step4. ナイブベイズ識別を用いて出力クラス C_i^* を決定する。

$$C_i^* = \arg \max_{c_i} p(c_i | I, x) \quad (9)$$

以上により、Conditional Classification Forests による人体の向きを考慮した人体パーツの識別を行う。

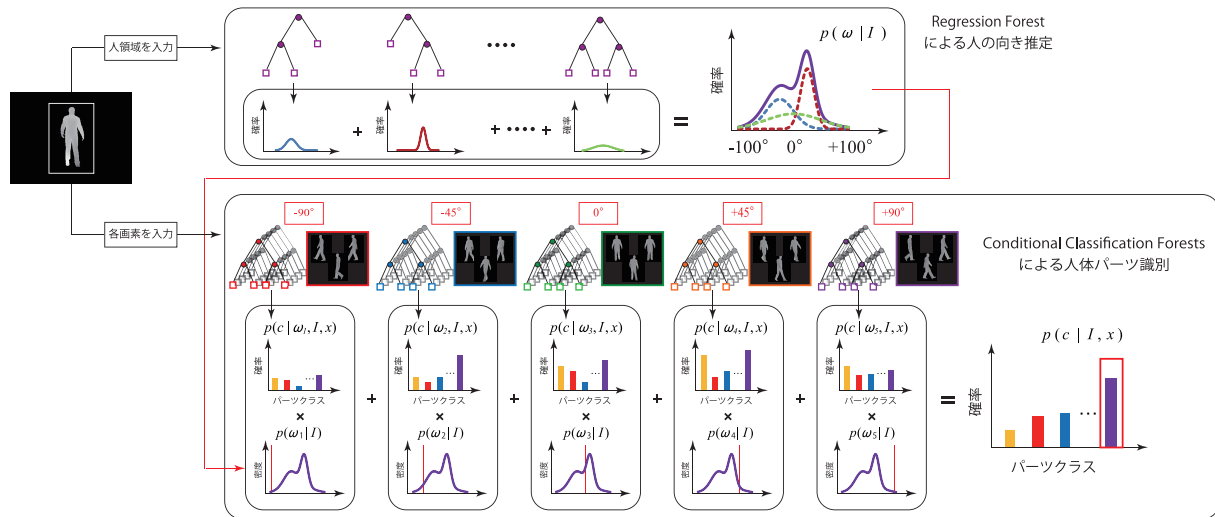


図 2: Conditional Classification Forests による人体パーツ識別の流れ

3. 評価実験

提案手法の有効性を示すために、人体モデルより作成したサンプルと実環境で撮影したサンプルにより評価する。従来法として全ての向きを含む学習データセットを Classification Forest により学習した手法と提案手法 (Conditional Classification Forests) を比較する。

3.1 人体向きの推定実験

Regression Forest により回帰推定した人体向きの推定精度を評価する。図 3 に人体の向きを回帰推定した際の角度毎の平均 2 乗誤差と分散を示す。実験により平均 4.96° の誤差で人体の向きを推定できることを確認した。入力距離画像の角度が $\pm 25^\circ$ や $\pm 65^\circ$ の周辺で平均誤差が高いことがわかる。これは、学習に用いていない角度のためだと考えられる。しかしながら、誤差 7° 以内で推定可能である。また、入力画像が $+65^\circ$ と推定された場合、分散が大きい確率分布となった。そのため、分散が大きい $+65^\circ$ のような確率分布を得た場合、Conditional Classification Forests では $+45^\circ$ と $+90^\circ$ の Classification Forest の重みが大きくなる。これにより、グループ化した 5 方向に含まれない向きの入力画像に対しては、複数の Classification Forest の識別結果を利用することで精度良い人体パーツ識別を可能とする。

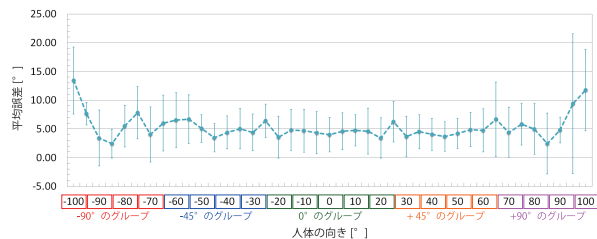


図 3: 人体向きの推定精度

3.2 人体パーツの識別実験

Conditional Classification Forests による人体パーツの識別精度を評価する。評価には、部位の識別が正解した画素の総数を TP、他の部位と間違えた画素の総数を FN として、 $\text{Precision} = \text{TP} / (\text{TP} + \text{FN})$ により識別率を算出する。図 4 に人体パーツ識別結果のコンフュージョンマトリクスを示す。対角成分が黒であるほど、そのパーツを正しく識別できたことを表す。図 4 から提案手法は、従来法と比べ識別率が高いといえる。提案手法の平均識別率は 89.1% であり、従来法の 75.9% と比較して、13.2% の精度向上を実現した。

3.3 実環境における人体パーツの識別実験

実環境における実験として、TOF カメラを用いて撮影した距離画像から人体パーツ識別を行う。図 5 に TOF カメラにより撮影した距離画像の識別結果を示す。図 5 より、実環境においても人の向きに関わらず精度良く識別できていることがわかる。また、表 1 に提案手法を構成する

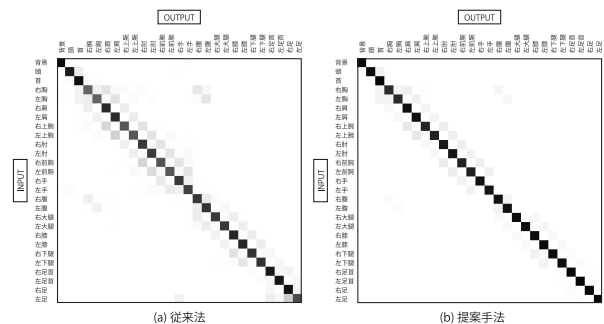


図 4: 人体パーツの識別精度

Regression Forest と Conditional Classification Forests, 従来法の Classification Forest の処理速度を示す。提案手法は、画像 1 枚につき平均 $298.8[\text{ms}]$ で人体の向き推定と人体パーツ識別を行うことができる。

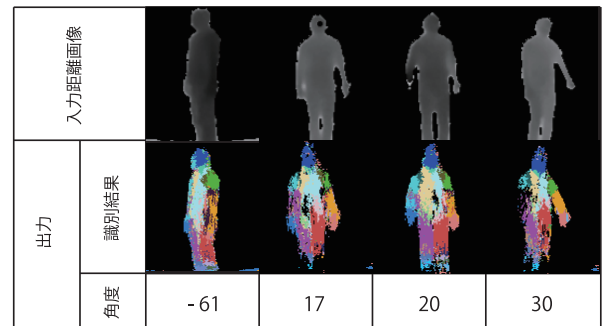


図 5: 実環境における人体パーツの識別例

表 1: 画像 1 枚の処理時間

手法	処理時間 [ms]
Classification Forest	71.4
Regression Forest	11.5
Conditional Classification Forest	287.3

4. おわりに

本研究では、人体の向きを考慮した Conditional Classification Forests を用いた人体パーツ識別手法を提案し、その有効性を確認した。

参考文献

[1] J. Shotton *et al.*, "Real-Time Human Pose Recognition in Parts from Single Depth Images", CVPR, pp.1297-1304, 2011.

研究業績

[1] 縣植輝, 三品陽平, 弓場竜, 藤吉弘亘, "Conditional Classification Forests による人体パーツ識別", 動的画像処理実用化ワークショップ 2013(DIA2013), 2013. (発表予定)
(他 研究会技術報告 2 件, 学会口頭発表 2 件)

受賞

[1] 情報処理学会 第 177 回 CVIM 研究会 卒論セッション 優秀賞