

ARCO による顔検出を併用した人誤検出の棄却について

山下 亮[†] 滝口 哲也^{††} 有木 康雄^{††}

[†] 神戸大学大学院システム情報学研究科 〒 657-8501 兵庫県神戸市灘区六甲台町 1-1

^{††} 神戸大学自然科学系先端融合研究環 〒 657-8501 兵庫県神戸市灘区六甲台町 1-1

E-mail: [†]ryo@me.cs.scitec.kobe-u.ac.jp, ^{††}{takigu,ariki}@kobe-u.ac.jp

あらまし 本研究では、人の特徴的なパーツである顔に着目した人検出器の精度改善手法を提案する。人検出には、特徴量を分散共分散行列とする Array of Covariance (ARCO) による人検出法を用いる。そして、ARCO による顔検出を人検出結果に併用する。この時、人検出器により識別が困難な画像に対してのみ顔検出を併用することで、人検出の精度改善を行う。評価実験では人の姿勢や向き、視点、照明の変動や背景テクスチャなど、様々なバリエーションを含んだ大規模な公開データセット INRIAPerson Data Set を使用し、Detection Error Tradeoff カーブで評価した。評価実験の結果従来手法では識別が困難であった画像の識別に成功し、提案手法の有効性を確認できた。

キーワード 人検出, 顔検出, 分散共分散行列, ARCO

1. はじめに

画像中の人物領域の検出は、人の形状の変化、複雑な背景、照明変動など、検出を困難とする要因が多く含まれているため、近年数多くの研究が行われてきている [1]-[13]。人物の正確な検出による位置情報の獲得は、監視システムや ITS、マーケティング、ロボットの視覚機能への応用など多岐にわたる分野で実現が期待されており、最も盛んに行われている研究のひとつである。また、人の領域を正確に検出できれば、人物の自動追跡や動作認識などにも適用できるため、その必要性はますます高まってきている。

検出技術の従来手法としては、検出対象の形状が事前に分かっている場合には、形状モデルとのマッチングで、検出する手法が提案されている。しかし、人物は形状の変化が激しく、形状モデルとのマッチングが難しいため、検出が困難になる。そのため、人検出においては、空間的特徴として、隣接画素との微分演算によって得られるエッジ特徴が検出によく用いられている。エッジ特徴には、影や天候などの照明変化の影響を受けにくい利点があるためによく利用されている。N.Dalal らは、Histograms of oriented gradients (HOG) [2] と呼ばれる高次元の特徴量を用いて、人検出を行う方法を提案している。HOG 特徴量は、局所領域の輝度の勾配方向をヒストグラム化したものであり、幾何学的変換に強いことや、照明変動に頑健であるといった利点がある。その後、HOG 特徴量を用いた人検出が数多く研究されている [11][12][13]。

近年では、画素ごとに複数の特徴量を抽出し、特定の領域ごとに分散共分散行列を求め、それをひとつの特徴量とする手法が高い精度を示している [1][3][4][5]。分散共分散行列を特徴量とする利点として、複数の特徴量を

特徴量間の関係性と共に表現することができるという点と、特徴量の次元数が大幅に削減されるという点がある。Tosato らは、対象の画像を複数個のパッチに分割することで、パッチごとに複数の特徴量を抽出して分散共分散行列を求め、パッチごとに人か人でないかの識別機を学習し、全パッチの多数決により最終的な識別結果を導く手法、Array of Covariance (ARCO) を提案している [1]。これは、複数クラスの識別にも適用できるものである。

本研究では、従来の人検出では識別が困難な画像に対して、人の特徴的なパーツである顔に着目し、誤検出を棄却することで人検出の高精度化を図る [1] で提案されている低解像度画像にも強い ARCO を用いた顔検出を人検出結果と併用することで、従来の人検出結果における誤検出率を減少させる手法を提案する。

2. 人検出における特徴量

近年、人検出の特徴量として注目されている複数の特徴量からなる分散共分散行列を、リーマン幾何学に基づいて特徴ベクトルとして表現する方法について述べる。

2.1 分散共分散行列を用いた人検出の特徴量

近年、精度の高い人検出の特徴量として、画素ごとに複数の特徴量を抽出し、その特徴量の性質とその関係性を共に表現することのできる分散共分散行列が数多く用いられてきている [1][3][4][5]。ある対象の画像領域の各画素ごとに d 次元の特徴ベクトルを抽出したとき、その領域に対応する分散共分散行列は d 行 d 列の行列 Σ として表現できる。図. 1 に、対象領域から特徴量の分散共分散行列を抽出する流れを示す。分散共分散行列は対角要素が分散、非対角要素は共分散を表し、対称行列となるので、 d 行 d 列の分散共分散行列は $d(d+1)/2$ 個の要素だけで表現が可能である。ある領域から各画素ごとに抽

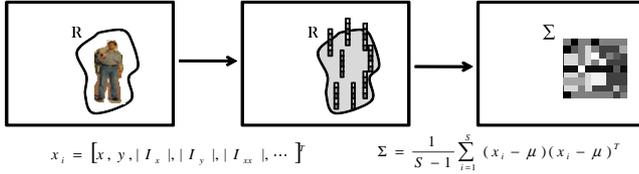


図1 Covariance matrix descriptor

出された d 次元の特徴ベクトルを $\{x_i\}_{i=1, \dots, S}$, その平均ベクトルを μ としたときの分散共分散行列 Σ は以下の式より求められる.

$$\Sigma = \frac{1}{S-1} \sum_{i=1}^S (x_i - \mu)(x_i - \mu)^T \quad (1)$$

2.2 リーマン幾何学に基づく共分散行列のベクトル化

前項で述べた分散共分散行列は正定値対称行列であるので、ベクトル空間を形成しない。したがって、行列間の距離計算はユークリッド空間上よりもリーマン空間上で行うほうが望ましい [1] [3] [4] [5] [15]。リーマン幾何学に基づいた分散共分散行列のベクトル化手法では、分散共分散行列 Σ の全体からなるリーマン多様体を M とおき、局所的にユークリッド空間とみなすことで、共分散行列をユークリッド空間上のベクトルに変換する。 T_Y をリーマン多様体上の点 $Y \in M$ における接ユークリッド空間とする。リーマン多様体上の点を局所的な接ユークリッド空間 T_Y へ写像する写像関数を $\log_Y : M \rightarrow T_Y$ とし、逆に T_Y 上の点からリーマン多様体 M へ写像する逆関数を $\exp_Y : T_Y \rightarrow M$ とする。図. 2 は、リーマン空間からユークリッド空間への写像の概念を示したものである。

リーマン多様体において、 $X, Y \in M$ とし、点 Y の接空間 T_Y 上の点を Z としたとき、点 Y の接空間上への写像関数 \log_Y と逆関数 \exp_Y はそれぞれ以下の式のように定義される。

$$\log_Y(X) = Y^{\frac{1}{2}} \log(Y^{-\frac{1}{2}} X Y^{-\frac{1}{2}}) Y^{\frac{1}{2}} \quad (2)$$

$$\exp_Y(Z) = Y^{\frac{1}{2}} \exp(Y^{-\frac{1}{2}} Z Y^{-\frac{1}{2}}) Y^{\frac{1}{2}} \quad (3)$$

また、行列の対数と指数はそれぞれ、以下の式により定義される。

$$\log(\Sigma) = U \log(D) U^T \quad (4)$$

$$\exp(\Sigma) = U \exp(D) U^T \quad (5)$$

ただし、 Σ は、 $\Sigma = U D U^T$ と固有値分解されるものとする。また、 $\log(D)$ と $\exp(D)$ はそれぞれの行列の固有値の対数または指数を対角要素として並べた対角行列である。[1] では、計算を簡単にするために、多様体上の接空間を形成する接点に単位行列 I_d を用いる。よって、式 (4) は、以下の式のように表わせる。

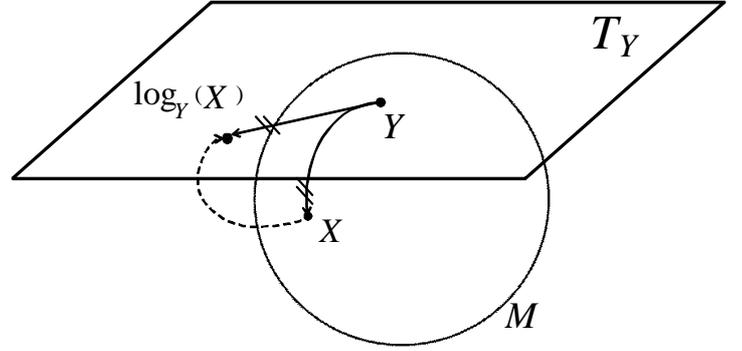


図2 Riemannian mapping to tangent space

$$\log_{I_d}(X) = \log(X) = U \log(D) U^T \quad (6)$$

このようにして、 d 行 d 列の分散共分散行列 X を接空間に写像した点 x も d 行 d 列の対象行列となる。この行列を式 (7) に示すように並べたベクトルを分散共分散行列から得られる特徴ベクトルとする。

$$\text{vec}_{I_d}(x) = \text{vec}(x) = [x_{1,1} \ x_{1,2} \ \dots \ x_{1,d} \ x_{2,2} \ x_{2,3} \ \dots \ x_{d,d}] \quad (7)$$

こうして得られた特徴ベクトルは、複数の特徴を関連性とともに表現でき、かつ次元数の少ない特徴ベクトルとなる。

3. ARCO (Array of Covariance)

ARCO (Array of Covariance) とは、Tosato らによって提案された分散共分散行列を用いた画像の特徴量表現である [1]。まず学習段階では、図. 3 のように画像の対象領域を重複を許しながら複数個のパッチに分割する。その後パッチごとに分散共分散行列を抽出する。分散共分散行列から変換された特徴ベクトルを基に、ブースティングを使って物体認識のための強識別器を得る。次に識別段階では、全パッチに対して強識別器を適用し、得られた識別結果の多数決によって最終的な画像の識別を行う。複数のクラスを持つ識別器を構築する場合、パッチごとに各クラス数分の強識別器を学習しておき、多数決によって複数クラスの識別が可能となる。上記の方法で学習される ARCO は、異なる複数の特徴量の有効な要素を考慮し、その相関関係を利用したものとなる。また、重複を許したパッチによる表現により、オクルージョンにも頑健となる。以下、ARCO による物体識別の詳細について述べる。

まず、一枚の画像を等間隔に分割し、重複を許しながら図. 5 のように N_P 個のパッチ $\{P_n\}_{n=1, \dots, N_P}$ に分割する。各パッチの記述子として画像特徴から得られた分散共分散行列を使用し、ブースティングにより J クラスの強識別器 $\{F_{P_n}\}_{n=1, \dots, N_P}$ を得る。それぞれのクラスの全パッチの識別結果による投票数は、式 (8) のように表わされる。

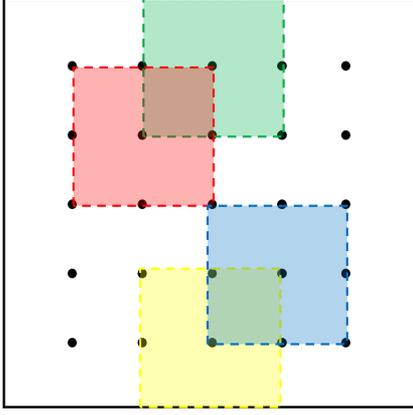


図 3 Array of Covariance matrices feature.

$$\Delta_j = \sum_{n=1}^{N_P} \theta(F_{P_n} == j) \quad (8)$$

θ は () の中が正しいときのみ 1 を返し、それ以外は 0 を返す関数である。ある画像のクラスを ARCO で識別したいとき、そのクラスは最終的に式 (9) によって識別される。

$$c = \arg \max_j \{\Delta_j\}, \quad (j = 1, \dots, J) \quad (9)$$

3.1 ARCO の学習方法

ARCO において、パッチごとに求められる強識別器について述べる。パッチごとに、特徴量として一つの分散共分散行列を求める。従って、 d 行 d 列の分散共分散行列全体がなすリーマン多様体 M 上の点を用いて、ブースティングにより識別器を学習する。ブースティング (boosting) とは、与えられた学習データを用いて学習を行い、その学習結果を踏まえて重みの調整を繰り返すことで複数の学習結果を求め、その結果を統合し組み合わせ、精度を向上させる手法である [18]。

ARCO の各パッチの強識別機は、Multi-class LogitBoost によって学習される [16]。Multi-class LogitBoost とは、Boosting アルゴリズムの一種である LogitBoost を複数クラスに拡張したものであり、クラス数を J としたとき、 J 個の強識別器 $\{F_1, \dots, F_J\}$ が得られる。弱識別器には重み付きの回帰木 [17] を使用する。Multi-class LogitBoost は学習データが対象クラスである確率 \Pr_j を見積もりながら学習を進めていくことで、高精度な識別器を構築できる。 $\Pr_j[\mathbf{x}_i]$ を、 j 番目のクラスに属する学習データ \mathbf{x}_i の事後確率とし、以下の式により計算する。

$$\Pr_j[\mathbf{x}_i] = \frac{e^{F_j(\mathbf{x}_i)}}{\sum_{k=1}^J e^{F_k(\mathbf{x}_i)}} \quad (10)$$

ただし $F_j(\mathbf{x}_i)$ は、

$$F_j(\mathbf{x}_i) = \sum_{l=1}^L f_{lj}(\mathbf{x}_i) \quad (11)$$

であり、 $\{f_{lj}\}_{l=1, \dots, L}$ はクラスごとの弱識別器の集合である。また、それぞれの学習データの重みは、クラスごとに以下の式で計算される。

$$w_{ij} = \Pr_j[\mathbf{x}_i](1 - \Pr_j[\mathbf{x}_i]) \quad (12)$$

次に、一つのクラスとその他のクラスのカテゴリを行う弱識別器 g_{lj} を、クラスごとに求める。2 クラスに対応した弱識別器 g_{lj} は、重み付きの回帰問題を解くことによって求められる関数である。このとき、回帰問題の目的変数 z_{ij} は $y_{ij}^* = \theta(j == y_i)$ としたとき、以下の式で定義される。ここで y_i は $y_i \in \{1, \dots, J\}$ である。

$$z_{ij} = \frac{y_{ij}^* - \Pr_j[\mathbf{x}_i]}{\Pr_j[\mathbf{x}_i](1 - \Pr_j[\mathbf{x}_i])} \quad (13)$$

そして、 J 個の 2 クラスに対応した弱識別器 g_{lj} は以下の式のように組み合わせられ、複数クラスの弱識別器 f_{lj} として計算される [16]。

$$f_{lj}(\mathbf{x}_i) = \frac{J-1}{J}(g_{lj}(\mathbf{x}_i) - \frac{1}{J} \sum_{k=1}^J g_{lk}(\mathbf{x}_i)) \quad (14)$$

図 4 に Multi-class LogitBoost のアルゴリズムを示す。パッチごとに抽出される特徴量を $X_i \in M$ とし、

- **Require :** $(X_1, y_1), \dots, (X_N, y_N)$ with $X_i \in M, y_i \in \{1, \dots, J\}$
- Map the data points to the tangent space T_{I_d} , by $\mathbf{x}_i = (\log_{I_d}(X_i))$
- Start with weights $w_{ij} = \frac{1}{N}$ and $i = 1, \dots, N, F_j(\mathbf{x}_i) = 0, \Pr_j[\mathbf{x}_i] = \frac{1}{J} \forall j$.
- for** $l = 1, 2, \dots, L$ **do**
- for** $j = 1, 2, \dots, J$ **do**
- Compute the response values $z_{ij} = \frac{y_{ij}^* - \Pr_j[\mathbf{x}_i]}{\Pr_j[\mathbf{x}_i](1 - \Pr_j[\mathbf{x}_i])}$ where $y_{ij}^* = \theta(j == y_i)$ and weights $w_{ij} = \Pr_j[\mathbf{x}_i](1 - \Pr_j[\mathbf{x}_i])$
- Fit the function $g_{lj}(\mathbf{x}_i) : \mathbf{R}^m \mapsto \mathbf{R}$ by weighted least-square regression of z_{ij} to \mathbf{x}_i using weights w_{ij} .
- Set $F_j(\mathbf{x}_i) \leftarrow F_j(\mathbf{x}_i) + f_{lj}(\mathbf{x}_i)$ where $f_{lj}(\mathbf{x}_i) = \frac{J-1}{J}(g_{lj}(\mathbf{x}_i) - \frac{1}{J} \sum_{k=1}^J g_{lk}(\mathbf{x}_i))$
- Update $\Pr_j[\mathbf{x}_i] = \frac{e^{F_j(\mathbf{x}_i)}}{\sum_{k=1}^J e^{F_k(\mathbf{x}_i)}}$
- Save $F_j = \{g_{lj}\}$
- end for**
- end for**
- Save the ensemble of classifiers $\{F_1, \dots, F_J\}$.

図 4 Multi-class LogitBoost on Riemannian Manifolds

$y_i \in \{1, \dots, J\}$ をクラスラベルとすると、 N 個の学習サンプル S は $S = \{X_i, y_i\}_{i=1, \dots, N}$ となる。このとき、 F_j は、ブースティングによって学習されたクラス j の L 個の弱識別器 $\{f_{lj}\}$ で定義される強識別器となる。そして、一つのパッチ P の識別器 $F_P(X_i) : M \mapsto \{1, \dots, J\}$ は、式 (15) のように定義される。

$$F_P(X_i) = \arg \max_j \{F_j(X_i)\}, \quad (j = 1, \dots, J) \quad (15)$$

3.2 ARCO による人検出

ARCO による人検出の実装について述べる．学習データには，人の姿勢や向き，視点，照明の変動や背景テクスチャなど，様々なバリエーションを含んだ大規模な公開データセットである INRIA Person DataSet [2] を使用した．

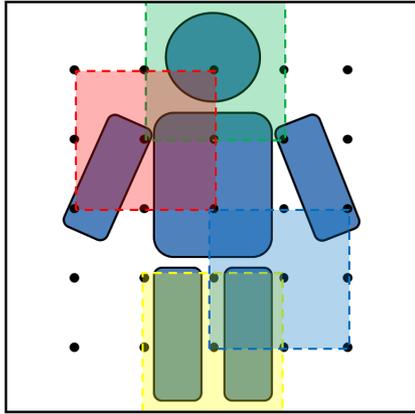


図 5 ARCO for Person.

図. 5 に示すように，各パッチから画素ごとに抽出される人検出用の特徴ベクトル Φ を式 (16) に示す．

$$\Phi = [x, y, |I_x|, |I_y|, m(x, y), |I_{xx}|, |I_{yy}|, \theta(x, y)]^T \quad (16)$$

x と y は各パッチ内の画素の座標を表す．画素の位置 (x, y) における輝度値を $I(x, y)$ としたとき I_x は x 方向の 1 次微分， I_{xx} は x 方向の 2 次微分を表している．1 次微分および 2 次微分は以下の式により算出する．

$$\begin{cases} I_x(x, y) = I(x+1, y) - I(x-1, y) & \forall x, y \\ I_y(x, y) = I(x, y+1) - I(x, y-1) & \forall x, y \end{cases} \quad (17)$$

$$\begin{cases} I_{xx}(x, y) = I(x+1, y) - 2I(x, y) + I(x-1, y) & \forall x, y \\ I_{yy}(x, y) = I(x, y+1) - 2I(x, y) + I(x, y-1) & \forall x, y \end{cases} \quad (18)$$

また， $m(x, y)$ はエッジの強さを， $\theta(x, y)$ はエッジの方向を表わし，以下の式より算出する．

$$m(x, y) = \sqrt{I_x(x, y)^2 + I_y(x, y)^2} \quad (19)$$

$$\theta(x, y) = \tan^{-1}(I_y(x, y)/I_x(x, y)) \quad (20)$$

次に，学習された ARCO による識別器を人検出器に拡張するために，Viola と Jones により提案された rejection cascade structure を適用する [14]．人と背景の 2 クラスからなる強識別器 $F(X_i)$ を直列につなぎ，図.6 に示すように，カスケード (Cascade) 型の検出器を構成する．

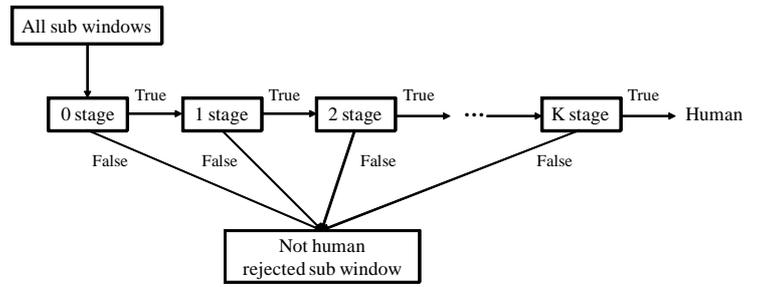


図 6 Rejection Cascade Structure

入力画像に対し，各段で人か人でないかの判定を行い，人と判定された画像だけが次の段へ進む．最後の段まで通過したものが最終的に人と判定される．このようにして学習されたカスケード型の検出器は，判別を行う場合，入力画像に対して，まず前の層の単純な識別器を使って識別する．初段で単純な識別器を用い明らかに人でない画像を除去し，識別の難しい画像に対して，より高い識別能力を有する後段の複雑な識別器を用いて識別を行う．大部分の人でない画像を計算量の少ない前段の層の識別器を使って識別するため，計算量を削減でき高速に演算可能である．図. 7 に，実際に ARCO を用いた人検出の実行結果の一部を示す．

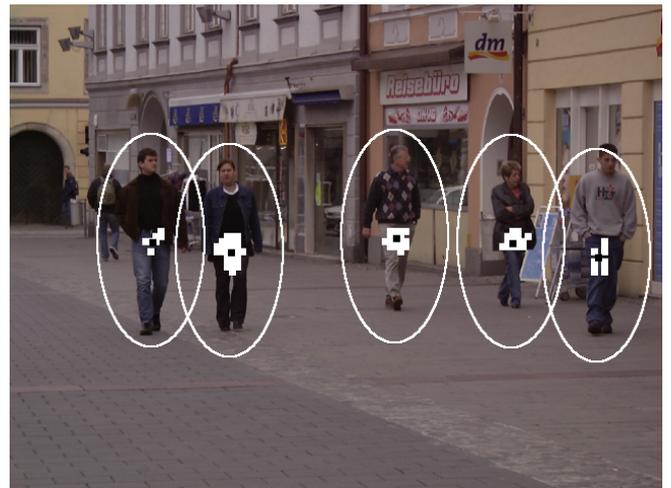


図 7 Detection result

3.3 ARCO を用いた顔検出

提案手法で使用する ARCO による顔検出について述べる [1]．学習用のデータセットには [1] で使用されているものと同様のものを使用した．このデータセットは 50×50 画素の顔画像からなり，実際の監視カメラ映像から得られた低解像度の画像データセットである．このデータセットは 4 つの前景クラスを含んでおり，それぞれ顔の向きが，後ろ，前，左，右の 4 種類からなる．これに，顔画像を含まない背景画像を加えることで 5 つのクラスからなるデータセットになっている．しかし，背景画像においてのみ，INRIA Person Dataset の背景

画像を用いた．これは，顔検出の対象が INIRA Person Dataset であるため，照明の変動や背景テクスチャなど様々なバリエーションを含んだ背景画像を学習したかったためである．図. 8 に例として学習画像の一部を示す．



図 8 Head pose dataset

顔検出に適した特徴量 Φ を以下の式に示す．この特徴量は対象画像の各画素ごとに抽出される特徴量である．

$$\Phi = [x, y, R, G, B, I_x, I_y, \theta(x, y), Gab_{\{0, \pi/3, \pi/6, 4\pi/3\}}]^T \quad (21)$$

x, y は各画素の座標を表し， R, G, B はカラー情報である． I_x, I_y は画像の位置 (x, y) における輝度値を $I(x, y)$ としたとき， x 方向と y 方向の 1 次微分である．最後に， Gab は，ガボール関数の縞模様の方向 θ の値を変えることで得られる特徴量である．特徴ベクトルを各画素ごとに抽出した後，図. 9 に示すように， $p/2$ 画素ごとにオーバーラップを許して得られる $p \times p$ 画素の全てのパッチから分散共分散行列を求める． p の値は， s を対象画像の画素数としたとき， $p = 0.32s$ によって定義される．よって， 50×50 の画像からは 25 個のパッチが得られることになる．この後，前章で述べた方法で ARCO の学習を行い顔検出用の識別器を得る．

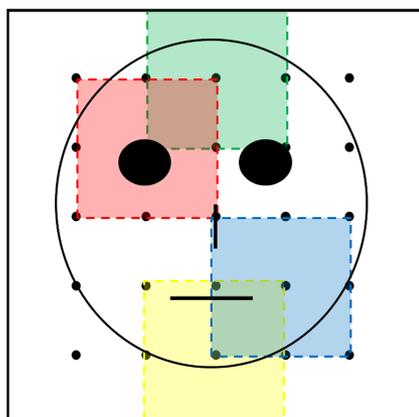


図 9 ARCO for face

4. 人検出結果への顔検出併用

4.1 顔検出器を用いた人認識器

顔検出器を用いた人認識器の流れを図. 10 に示す．図. 11 に示すように，人検出を行った後，検出された人領域の特定の位置に対して，ARCO を用いた顔検出を行う．顔検出を実行する場合，対象領域に対して検出ウイ

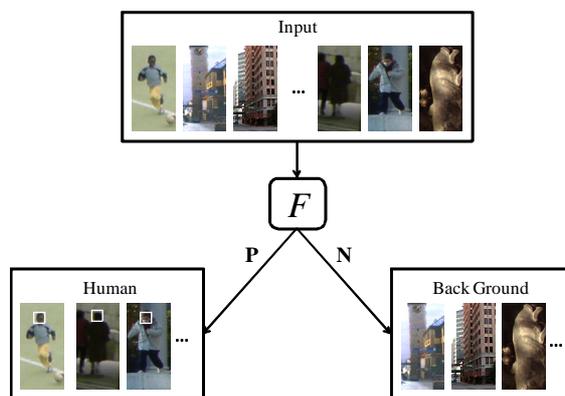


図 10 Person recognition with face detector

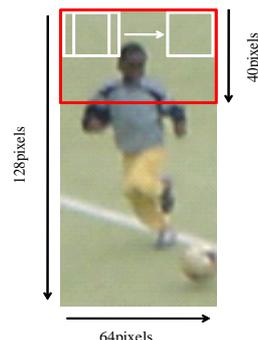


図 11 Face detection

ンドウをラスタスキャンさせる．もし，入力画像の対象領域において顔として検出されたウィンドウが得られたならば，入力画像を人 (Positive) とする．逆に入力画像から顔が検出されなければ，背景 (Negative) として棄却する．

4.2 人検出器と顔検出器の併用方法

人検出器と顔検出器の併用方法の流れを図. 12 に示す．ARCO を用いて学習された人検出器は各カスケードを

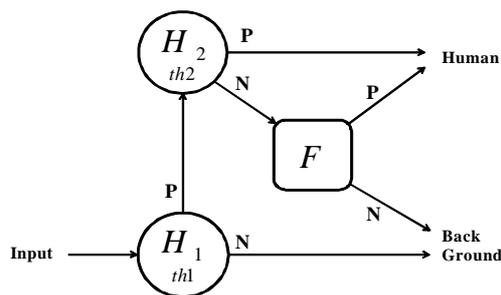


図 12 Detection Flow

通過させる割合を決める閾値 th の値を変化させることによって，誤検出率と未検出率を変化させることができる．人検出器 H_1 の閾値を，人の検出率は高いが背景の誤検出率 (背景を人と検出する率) も高くなるような閾値 th_1 に設定する．逆に，人検出器 H_2 の閾値を，人の検出率は低くなるが，背景の誤検出率も低くなるような閾

値 th_2 に設定する．人検出器 H_1 を用いて人検出を行い，人が検出されたとするとする画像全てに対して，人検出器 F を適用すると，正しく人と認識されていた人画像の多くから顔が検出されずに誤検出として棄却され，検出率が大きく下がってしまうという問題が生じる．そこで，人検出器 H_1 を用いた人検出の結果，人が検出された画像全てに対して人検出器 H_2 で再度検出を行う．人検出器 H_2 による人検出の結果，背景として認識された画像に対してのみ，顔検出を用いた人検出器 F による検出を行う．その結果，顔が検出された画像は，人検出器 H_2 によって背景と誤検出された人画像と判断し，顔が検出されなかった画像は人検出器 H_1 によって人と誤検出された背景画像として棄却する．このような構造をとることによって，検出が困難な人画像と人と誤検出しやすい背景画像に対してのみ，顔検出を用いた人検出器 F を適用することができる．この結果，人の検出率を下げすぎることなく，背景の誤検出率を下げる事が可能になる．この検出システムの概要を図. 13 に示す．

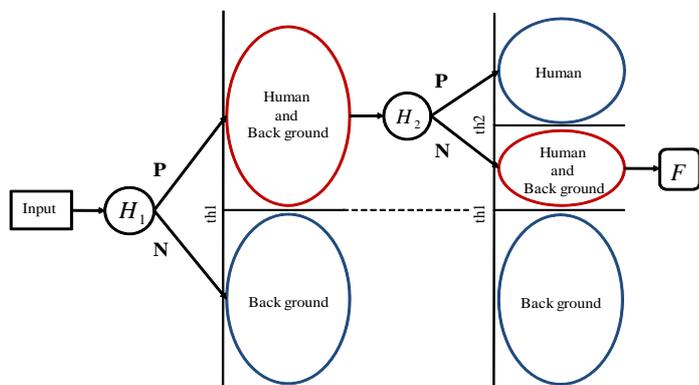


図 13 Detection system

5. 評価実験

5.1 データベース

実験に使用する人画像のデータベースには，汎用性のある画像データベースとして，Web 上で一般公開されている INRIA Person Dataset [2] を使用する．INRIA Person Dataset は，人検出法のベンチマークデータベースとして利用されており，人の姿勢や向き，視点，照明変動や背景など，様々なバリエーションを含む大規模データセットである．INRIA Person Dataset では，学習用データセットおよび評価用データセットのそれぞれについて，人物画像および背景画像が与えられている． 64×128 画素の人物画像 2416 枚と人物を含まない背景画像 1218 枚から成る．背景画像については用意された背景画像からランダムに矩形領域を切り出すことで背景サンプルとして用いる．また，評価用セットとしては，人物画像 1125 枚と背景画像 453 枚が提供されている．図. 14 に INRIA Person Dataset の一部を示す．

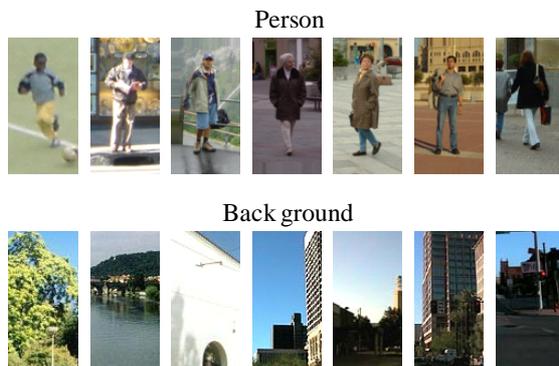


図 14 INRIA Person Dataset

5.2 評価方法

評価方法には，Detection Error Tradeoff (DET) カーブ [20] によって行う．DET カーブとは，横軸に誤検出率 (False positives per window)，縦軸に未検出率 (Miss rate) を両対数グラフによって表したものである．誤検出率 (FPPW) は式 (22) によって，未検出率 (Miss rate) は式 (23) で表わされる．識別器の閾値を変化させることによって，誤検出率に対する未検出率の比較を行うことが可能である．DET カーブでは，左下に向かうほど検出器の性能が高いことを意味する．

$$FPPW = \frac{\text{FalsePositive}}{\text{TrueNegative} + \text{FalsePositive}} \quad (22)$$

$$\text{Miss rate} = \frac{\text{FalseNegative}}{\text{FalseNegative} + \text{TruePositive}} \quad (23)$$

5.3 実験結果

図. 15 に実験結果による DET カーブを示す．実験の

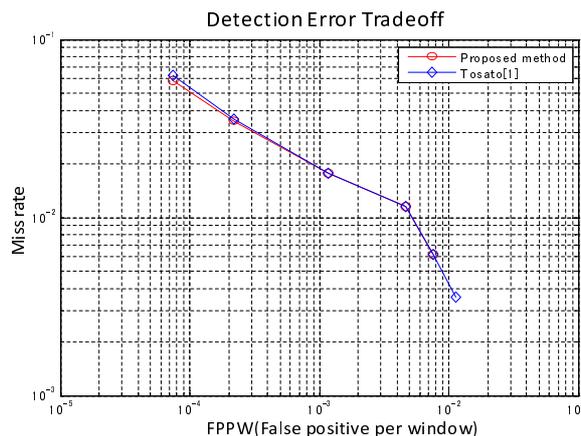


図 15 DET curve

結果，人誤検出率 (FPPW) が $10^{-4}\%$ のときの人検出率が 0.44% 向上していることが確認できた．

6. まとめと考察

顔検出を併用しない手法と提案手法とを比較すると、誤検出率が $10^{-4}\%$ の場合の人検出率が 0.44% と僅かながら向上した。しかし、顔検出器によって実際に顔を検出することのできた人画像はわずかであった。これは、テスト画像のサイズが 64×128 画素であるため、顔の領域は 20×20 画素程度しかなく解像度が低すぎたことが一つの原因として考えられる。また、INRIA Person Dataset は、様々な人の姿勢や向き、服装など様々なバリエーションを含んでいるために、姿勢によっては顔が半分程度隠れてしまっている人画像もあり、正確な顔検出結果が得られなかったことにも原因があると考えられる。また、提案手法では顔検出を併用することで学習データに無いような未知の姿勢にも対応でき、姿勢の変化に頑健であるといえる。

文 献

- [1] D. Tosato, M. Farenzena, M. Cristani, V. Murino, " Multi-class Classification on Riemannian Manifolds for Video Surveillance , "ECCV 2010, Part II, LNCS 6312, pp. 378-391, 2010.
- [2] N. Dalal, B. Triggs, " Histograms of oriented gradients for human detection , "CVPR, vol.1, pp. 886-893, 2005.
- [3] O. Tuzel, F. Porikli, P. Meer, " Pedestrian detection via classification on riemannian manifolds , "IEEE Trans. PAMI, pp. 1713-1727, 2008.
- [4] D. Tosato, M. Farenzena, M. Cristani, V. Murino, " PART-BASED HUMAN DETECTION ON RIEMANNIAN MANIFOLDS , "ICIP 2010, pp. 3469-3472, 2010
- [5] D. Tosato, M. Farenzena, M. Cristani, V. Murino, " An FPGA-based Classification Architecture on Riemannian Manifolds , "2010 Workshops on Database and Expert Systems Applications, pp. 215-220, 2010.
- [6] K. Levi, Y. Weiss, " Learning Object Detection from a Small Number of Examples: the Importance of Good Features , "IEEE Computer Vision and Pattern Recognition, vol. 2, pp. 53-60,2004.
- [7] B. Wu, 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.
- [8] P. Viola, M. Jones, D. Snow, " Detection Pedestrians Using Patterns of Motion and Appearance , "IEEE International Conference on Computer Vision, pp. 734-741, 2003.
- [9] N. Dalal, B. Triggs, C. Schmid, " Human Detection Using Oriented Histograms of Flow and Appearance , "IEEE European Conference on Computer Vision, vol. 2, pp. 428-441,2006.
- [10] P. Sabzmeydani, G. Mori, " Detecting Pedestrians by Learning Shapelet Feature , "IEEE Computer Vision and Pattern Recognition, pp. 511-518, 2007.
- [11] F. Suard, A. Broggi, " Pedestrian Detection using Infrared images and Histograms of Oriented Gradients , "IEEE Symposium on Intelligent Vehicule, pp. 206-212, 2006.
- [12] Q. Zhu, S. Avidan, M. Yeh, K. Cheng, " Fast Human Detection Using a Cascade of Histograms of Oriented Gradients , "IEEE Computer Vision and Pattern Recognition, vol. 2, pp. 1491-1498, 2006.
- [13] 小林拓也, 日高章理, 栗田多喜夫, " Histograms of Oriented Gradients を用いた対象識別での特徴選択 , "信学技報, vol. 106, pp. 119-124, 2007.
- [14] P. Viola, M. Jones, " Rapid Object Detection Using a Boosted Cascade of Simple Features , "IEEE Computer Vision and Pattern Recognition, vol. 1, pp. 511-518, 2001.
- [15] X. Pennec, P. Fillard, N. Ayache, " A Riemannian Framework for Tensor Computing , "International Journal of Computer Vission, vol. 66, no. 1, pp. 41-66, 2006.
- [16] J. Friedman, T. Hastie, R. Tibshirani, " Additive Logistic Regression: A Statistical View of Boosting , " Annals of Statistics, vol. 28, no. 2, pp. 337-407, 2000.
- [17] L. Breiman, J. Friedman, R. Olshen, C. Stone, " Classification and Regression Trees , "Ann. Math. Statist. 19, pp. 293-325, 1984.
- [18] Y. Freund, R. Schapire, " A decision-theoretic generalization of on-line learning and an application to boosting , "Journal of Computer and System Sciences 55, pp. 119-139, 1997.
- [19] C. P. Papageorgiou, M. Oren, and T. Poggio, " A general framework for object detection , "IEEE Computer Vision and Pattern Recognition, pp. 511-562, 1998.
- [20] F. Martin, "The DET Curve in Assessment of Detection Task Performance "Eurospeech '97, Rhodes, Greece, vol. 4, pp. 1899-1903, 1997.