

## 身体と外見の共起属性を用いた人物対応付け

中野 翔太<sup>†</sup>            四元 辰平<sup>†,††</sup>            吉村 宏紀<sup>†</sup>  
西山 正志<sup>†</sup>            岩井 儀雄<sup>†</sup>            菅原 一孔<sup>†</sup>

Co-occurrence Attributes of Physical and Adhered Human Characteristics  
to Improve Person Re-identification

Shota NAKANO<sup>†</sup>, Tappei YOTSUMOTO<sup>†,††</sup>, Hiroki YOSHIMURA<sup>†</sup>,  
Masashi NISHIYAMA<sup>†</sup>, Yoshio IWAI<sup>†</sup>, and Kazunori SUGAHARA<sup>†</sup>

あらまし 本論文では、複数カメラで得た人物画像を検索するために、身体と外見の共起属性を用いて人物を対応付ける手法について述べる。既存手法は性別や服装といった単一属性を用いて人物対応付けを行っていたが、提案手法は身体と外見の共起属性を用いてより詳細な手がかりを抽出し対応付けの性能向上を狙う。共起属性は、例えばスカートの女性など、身体的な手がかりと外見の手がかりを組合せることで表される。提案手法は、共起属性を推定する識別器を事前に学習しておき、人物画像が与えられると共起属性の確信度を算出し、得られた確信度をその人物画像の特徴量とする。対応付けに有効な共起属性を重視するために計量学習を用いて特徴量間の距離を算出する。複数の公開データベースを用いた人物対応付けの実験により、単一属性を用いる既存手法に比べて共起属性を用いる提案手法が有効であることを確認した。また、対応付けの性能向上に寄与する共起属性を解析する手法を設計し、身体と外見の共起属性が有効であることを確認した。

キーワード 人物対応付け, 特徴抽出, 共起属性, 身体的手がかり, 外見の手がかり

### 1. はじめに

安心安全な社会の実現に向けて、屋外で複数の防犯カメラを用いた広域の見守りが求められている。このような広域の見守りを実現することで、例えば迷子になった子供がどのような移動経路を辿ったかなど、人物の行動履歴を検索できるようになる。しかし、視野を共有しない複数の防犯カメラを渡って移動する人物を見つけ出すには人手だと大きな負担が伴う。そこで、クエリとして人物画像が入力されると、一致する可能性のある人物画像を精度よく検索する手法が求められている。

人物をカメラ間で対応付ける手法はこれまで多く提案されてきたが、これまでに得られた知見により人物画像から抽出する特徴量の設計が性能向上に大きく寄

与することが分かっている。人物対応付けの特徴量として、勾配や色など画像特徴を抽出する手法 [1] ~ [4], 深層学習における中間層を特徴量とした手法 [5], 性別や年代など人物属性を特徴として抽出する手法 [6] ~ [11] が提案されてきた。特に人物属性の特徴はユーザが直観的に理解しやすいため生体認証の研究分野からも期待されている。属性を特徴として用いる人物対応付けの応用は次の二つがある。一つ目として属性そのものをクエリとしてユーザが指定し人物画像を検索する応用があり、属性認識の性能を高めることを目的とした手法 [6], [7] が提案されている。二つ目として属性ではなく人物画像をクエリとし画像から推定される属性らしさをシステム内の特徴量として用いる応用があり、画像同士を精度よく対応付けることを目的とした手法 [8] ~ [11] が提案されている。いずれの応用でも、文献 [6], [7] で指摘されているように、複数の属性が同時に出現する共起の考え方が重要となる。しかし、これらの既存手法は以下で述べるように、人が人を探す際に用いる手がかりと比較すると、人物対応付けに用いる共起の取り方として十分とは言えなかった。

<sup>†</sup> 鳥取大学大学院 工学研究科 情報エレクトロニクス専攻  
Department of Information and Electronics, Graduate  
School of Engineering, Tottori University

<sup>††</sup> メルコ・パワー・システムズ株式会社  
System Engineering Department, Melco Power Systems  
Co.,Ltd.

ここで実際に人が人を探す時に用いる手がかりである属性の組合せについて考察する．例えば、警察庁が公開している行方不明者の情報 [12]、警視庁の公開捜査ファイル [13]、Police Scotland の Missing Persons [14]、Most Wanted [15] では、性別や年代などの身体的な手がかりや、服の種類やサングラスの有無など外見の手がかりが組合せられた形で掲載されている．これらには、身体的な手がかりと外見の手がかりの組合せ、身体的な手がかり同士の組合せ、外見の手がかり同士の組合せがある．これらの組合せは人が人を探すうえで有効な特徴であると考えられ、計算機で人物対応付けの手法を実現する上で有効な特徴量になると期待できる．

そこで、本論文では、身体的な手がかりや外見の手がかりを組合せることで共起属性を定義し、人物画像から抽出された共起属性を用いて人物対応付けの性能を向上させる手法について述べる．人が人を探す際に用いる手がかりを考察することで、共起属性をどのように表現すべきかを決定する．提案手法は、人物画像が与えられると複数の共起属性の確信度を推定し、それらの確信度から人物対応付けの特徴量を抽出し、計量学習を用いて特徴量間の距離を算出する．提案手法の主な貢献点を以下に挙げる．

- 人が人を探す時の知見に基づき身体的な手がかりと外見の手がかりを用いて共起属性を定義
- 手がかりのどの組合せが人物対応付けの性能向上に寄与するかを明らかにするために計量学習で得た計量行列を解析

さらに、提案手法を用いることで既存手法と比べて人物対応付けの性能が向上することを複数の公開データベースを用いた評価で確認した．

以下、2. で関連研究、3. で共起属性の定義、4. で共起属性を用いた人物対応付けについて述べる．次に、5. で人物対応付けの性能を評価する．最後に 6. でまとめる．

## 2. 関連研究

ここでは人物対応付けに用いる特徴量について既存手法を紹介する．まず、勾配や色などの画像特徴を抽出する手法について述べる．文献 [2] では勾配方向ヒストグラム、文献 [1] では RGB 値ヒストグラム、文献 [3] ではそれらの組合せを用いる手法を提案している．また、文献 [4] では、勾配や色のフィルタバンクから学習で有効な特徴を選択する手法を提案してい



図 1 身体と外見の共起属性の例

Fig. 1 Examples of co-occurrence attributes of physical and adhered human characteristics

る．しかし、勾配を用いる手法はカメラに映る人物の姿勢など見え方の変化の影響を受けやすいという問題があり、色を用いる手法は照明色の変動の影響を受けやすいという問題がある．一方、姿勢や照明の変化に影響を受けにくい特徴量を目指し人物属性を用いる手法が提案されている．以下では、性別や髪型や服装などそれぞれの人物属性を単一属性と呼ぶことにする．文献 [6], [7] では、複数の属性が同時に出現する共起の制約を入れることで単一属性の認識性能を向上させる手法が述べられている．文献 [10] では、単一属性の識別性能と人物対応付けの性能を同時に最適化する手法を提案している．さらに文献 [11] では、ファッション人物画像に付与された単一属性を転移学習により防犯カメラ映像の人物対応付けに利用する手法を提案している．また、文献 [8], [9] では、どの単一属性が人物対応付けに有効であるかを明らかにしている．しかし、これらの既存手法は人が人を探す際に用いる身体的な手がかりと外見の手がかりの組合せを十分に考慮しているとは言えなかった．

## 3. 共起属性の定義

### 3.1 身体の手がかりと外見の手がかりの組合せ

ここでは提案手法で取り扱う人物の共起属性について述べる．人が人を探す際、文献 [16], [17] でも述べられているように、身体、外見、行動の 3 種類の手がかりが利用されている．身体的な手がかりとはその人固有の特徴であり、性別や年代のように時間的な影響をほぼ受けにくいものや、髪型のように数か月単位で個人性に影響を与えるものが含まれる．外見の手がかりとはその人の見た目の特徴であり、服装やサングラスのようにその人物に何らかの物が一時的に取り付けられることで変化していく．行動的な手がかりとはその人の仕草や癖や身振りなど時間方向の変化に含まれる個人性である．

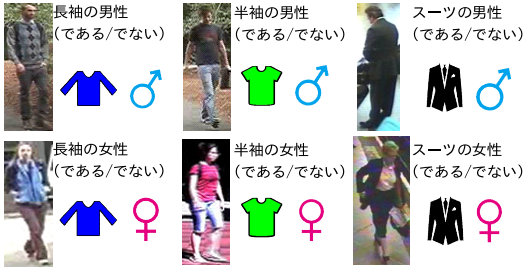


図 2 身体 (性別) と外見 (トップス) の共起属性の二値表現の例

Fig. 2 Binary representation of co-occurrence attributes of physical (gender) and adhered human (tops) characteristics

提案手法では、公開 WEB サイト [12] ~ [15] にも掲載されているように、人が人を探す際に特に利用される特徴である身体的な手がかりと外見の手がかりの組合せから共起属性を定義する。二つの手がかりを組合せる場合、身体的な手がかり同士、外見の手がかり同士、身体的な手がかりと外見の手がかりのように 3 種類が存在する。提案手法は例えば図 1 のように、身体的な手がかり同士の共起属性である性別と年代の組合せの“20 代の男性”，外見の手がかり同士の共起属性であるトップスとボトムスの組合せの“半袖で半ズボン”，身体的な手がかりと外見の手がかりの共起属性である性別とボトムスの組合せの“スカートの女性”を用いる。

### 3.2 二値表現

身体的な手がかりと外見の手がかりには、性別 (男性/女性) やサングラス (有/無) のように二値で表される属性と、年代 (20 代/30 代/40 代) やトップス (長袖/半袖/スーツ) のように多値で表される属性がある。二つの手がかりのクラス数を  $L, M$  とすると、これらを単純に組合せた後のクラス数は  $LM$  となる。提案手法では、組合せの各要素が独立であると仮定し共起属性性を二値で表すこととする。例えば、図 2 のように性別 ( $L = 2$ ) とトップスの種類 ( $M = 3$ ) を組合せる場合、二値で表される共起属性性が  $LM = 6$  種類ほど出現する。それぞれの共起属性は、“長袖の男性である/でない”，“長袖の女性である/でない”，“半袖の男性である/でない”，“半袖の女性である/でない”，“スーツの男性である/でない”，“スーツの女性である/でない”となる。文献 [17] でも述べられているように、生体認証の研究分野では二値の属性が多値や連続値の属性と比べて人間にとって直観的に理解しやすいことが知ら

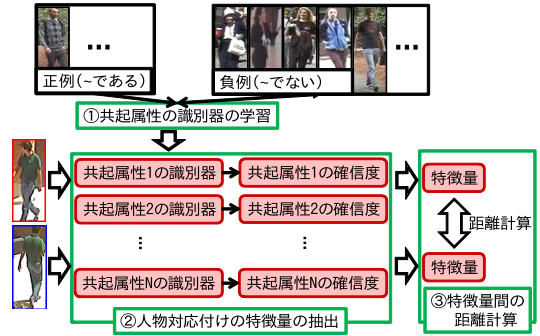


図 3 身体と外見の共起属性を用いた人物対応付けの流れ  
Fig. 3 Flow of person re-identification using co-occurrence attributes

れている。この知見に基づき二値の共起属性を提案手法でも用いることで、人物対応付けの実験結果を解析し易くすることを狙う。なお提案手法では共起属性の推定を集団学習 [18] における弱識別器として扱う。各共起属性を独立と仮定すると、共起属性の種類によっては弱識別器の出力が矛盾する場合も発生するが、ここでは弱識別器の設計を単純にしその数を増やすことを考える。多数の弱識別器を用いることで、人物対応付けに用いる強識別器の性能を高めることを狙う。

## 4. 共起属性を用いた人物対応付け

### 4.1 提案手法の流れ

身体と外見の共起属性を特徴量として用いる人物の対応付け手法について述べる。その流れを図 3 に示す。まず、共起属性ラベルが指示された訓練サンプルを用いて属性推定の識別器を事前に学習する。共起属性ラベルは、3. で述べた定義に従って正例と負例を決定する。ただし、共起属性の組合せ数は大量になるため、その訓練サンプルを手動で用意することは容易ではない。よって、共起属性を生成する際に組合せる二つの単一属性ラベルを用いて共起属性ラベルの正例と負例を自動で割り振る。二つの単一属性の組合せ数が  $LM$  種類とすると、共起属性では 1 種類を正例とし、残りの  $LM - 1$  種類を負例とする。例えば図 2 の場合、正例を長袖の男性とすると、負例には長袖の女性、半袖の男性、半袖の女性、スーツの男性、スーツの女性が属する。それぞれの共起属性について学習を行い二値の識別器を生成する。

次に、人物対応付けの画像が入力されると、学習した識別器を用いて各共起属性の確信度を算出する。算

出した確信度を結合することで人物対応付けの特徴量を決定する．人物画像から抽出された特徴量の間で距離を求め対応付けの候補を決定する．以下では提案手法の詳細について述べる．

#### 4.2 共起属性の識別器の学習

共起属性ラベルが付与された訓練サンプルについて、正例の枚数は負例の枚数に対し極端に少なくなる場合がある．一般的に機械学習の手法は、正例と負例の間でサンプル数が大きく異なる場合、サンプル数が多い方に学習が偏り、識別精度が低下する．訓練サンプルの枚数の不均衡に対応するために、枚数の逆数に応じた重みを学習時に与える手法が提案されているが、ここでは正例と負例の間で枚数が少ないほうにサンプル数を揃える手法を採用する．これは予備実験で性能が高かったためである．次に、共起属性の識別のための特徴量を Ensemble of Localized Features(ELF) [4] を用いて訓練サンプルから抽出する．共起属性の識別器として二値を判別する線形サポートベクタマシンを用い、以下では  $N$  種類の共起属性を  $SVM_i$  ( $i \in 1, \dots, N$ ) と表記する．

#### 4.3 人物対応付けの特徴量の抽出

人物を対応付ける際に用いる特徴量の算出手法について述べる．まず、人物画像が与えられると、各共起属性の識別器を適用し、識別境界面からの符号付き距離  $h_i$  を式 (1) で算出する．

$$h_i = SVM_i(e) \quad (1)$$

ここで  $e$  は人物画像から算出された ELF 特徴量を表す．符号付き距離の値が正であれば入力画像が正例に属し値が負であれば負例に属する．提案手法では、一般的な機械学習の手法と同様に符号付き距離の絶対値を信頼度とし、その値が大きいかほど信頼度が高く、小さいほど信頼度が低いとする．ただし、得られた符号付き距離をそのまま人物対応付けの特徴量の要素とすると、各要素で値の取りうる範囲が揃っていないため偏りが生じる．ここでは一般的に知られている手法で共起属性毎に符号付き距離をスケールリングする．確信度  $x_i$  は、式 (2) によって得られた  $h'_i$  を用いて最小値  $-1$  から最大値  $1$  の範囲に式 (3) で変換することで求まる．

$$h'_i = \frac{h_i - \min h_i^j}{\max h_i^j - \min h_i^j} \quad (2)$$

$$x_i = 2h'_i - 1 \quad (3)$$

ただし、正例と負例の訓練サンプル数の合計を  $T$  とすると  $\max, \min$  は  $h_i^j$  ( $j \in 1, \dots, T$ ) の最大値と最小値を返す．最後に、各共起属性の確信度  $x_i$  を式 (4) でベクトル化し人物対応付けの特徴量  $x$  を決定する．

$$x = [x_1, x_2, \dots, x_N]^T \quad (4)$$

#### 4.4 特徴量間の距離計算

人物画像から求めた特徴量間の距離を求める手法について述べる．特徴量の要素である共起属性の確信度は、人物対応付けに寄与するものと寄与しないものが存在する．そのため距離計算の際、寄与する要素に大きな重みを与え対応付けの性能を高める手法が提案されている．文献 [9] では欲張り法を用いて単一属性に重みづけを行っていたが収束するまでの速度に問題があった．近年では文献 [19] のように計量学習を用いて人物対応付けの特徴量に重みを与える手法が提案されている．提案手法では計量学習の手法としてよく知られている Large Margin Nearest Neighbor (LMNN) [20] を使用する．この手法は本人の特徴量を近づけ、別人物の特徴量を遠ざけるように計量行列  $M$  を学習する．特徴量  $x_a, x_b$  間の距離  $d$  は行列  $M$  を用いて式 (5) で表される．

$$d^2 = (x_a - x_b)^T M (x_a - x_b) \quad (5)$$

距離  $d$  が小さいほど人物画像  $a$  と人物画像  $b$  が似ていることを表す．

## 5. 実験

### 5.1 人物対応付けの基本性能の評価

#### 5.1.1 評価方法

提案手法の共起属性を用いた人物対応付けの有効性を確認するために実験を行った．この実験では公開されている人物画像データベースの PETA Dataset [21] を用いた．PETA Dataset では複数の既存データベース (3DPeS [22], CAVIAR4REID [23], CUHK [24], GRID [25], i-LID [26], MIT [27], PRID [28], SARC3D [29], Town-Centre [30], VIPeR [31]) に対して様々な単一属性ラベルが付加されている．評価に用いる身体と外見の手がかりとして、公開 WEB サイト [12] ~ [15] と文献 [8] の双方に示されている単一属性を選択した．二値の身体的な手がかりとして性別、二値の外見的な手がかりとしてサングラス有無、多値の身体的な手がかりとして髪型と年代、多値の外見的な手がかりとしてトップ

表 1 既存手法の実験に用いた単一属性

Table 1 Single attributes for experiments of existing method

	二値属性	多値属性
身体 ( )	性別 (男性/女性)	髪型 (短髪/長髪/禿) 年代 (30 以下/45 以下/60 以下/60 超)
外見 ( )	サングラス (有/無)	トップス (半袖/長袖/スーツ上) ボトムス (短パン/スカート/スーツ下)

とボトムスの種類を選択し表 1 の合計 15 種類とした。表中の ( ) は身体的な手がかり、( ) は外見の手がかりを表す。なお、公開 WEB サイト [12] ~ [15] では色を使った属性が多く用いられているが、文献 [17] で述べられているように色はカメラ間での再現性が難しいため本実験では用いないこととした。次に、単一属性の組合せから共起属性を表 2 の 96 種類で定義した。表中の ( ) は身体的な手がかり同士の共起属性、( ) は外見の手がかり同士の共起属性、( ) は身体的な手がかりと外見の手がかりの共起属性を表す。ただし、単一属性 15 種類の間における全ての組合せを共起属性として採用するのではなく“トップスがスーツでボトムスが短パン”や“禿げた女性”のように、公開 WEB サイト [12] ~ [15] を参考に実シーンであまり出現しない組合せは採用しなかった。このような組合せの人物画像は PETA dataset でも存在しないか、極端に枚数が少なかった。共起属性の内訳として、身体同士の組合せが 22 種類、外見同士の組合せが 15 種類、身体と外見の組合せが 59 種類であった。人物対応付けの評価データと属性識別の訓練サンプルを区別するために、図 4 のように PETA Dataset を VIPeR とそれ以外のデータベースに分けた。以下では VIPeR を用いて人物対応付けの性能を評価し、共起属性の識別器はそれ以外のデータベースから学習した。VIPeR に含まれる 632 人から 316 人をランダムに選択することで 2 つに分割し、1 つを計量学習に用いて残りを人物対応付け用のテストセットとした。このランダム分割を 10 回行い、各回毎に人物対応付けの性能を評価した。対応付けの際は、各人物につき一枚を辞書画像とし異なる一枚を入力画像とした。評価指標として、Cumulative Match Characteristic (CMC) 曲線と normalised area under the CMC curve (nAUC) を用いた。CMC は入力画像と距離が近い上位  $n$  番目

表 2 提案手法の実験に用いた共起属性

Table 2 Co-occurrence attributes for experiments of proposed method

45 以下-短パン ( )	男性-長髪 ( )
スーツ上-長袖 ( )	60 以下-半袖 ( )
女性-60 超 ( )	男性-30 以下 ( )
短髪-スーツ下 ( )	禿-半袖 ( )
短髪-半袖 ( )	女性-スカート ( )
サングラス有-短髪 ( )	45 以下-短髪 ( )
60 以下-スーツ上 ( )	男性-長袖 ( )
女性-スーツ下 ( )	サングラス無-短パン ( )
サングラス有-半袖 ( )	60 以下-禿 ( )
長袖-スーツ下 ( )	サングラス無-長袖 ( )
45 以下-長髪 ( )	女性-45 以下 ( )
女性-スーツ上 ( )	45 以下-スーツ下 ( )
30 以下-半袖 ( )	女性-長袖 ( )
短髪-スカート ( )	男性-60 超 ( )
男性-スーツ上 ( )	30 以下-スーツ上 ( )
60 以下-長髪 ( )	男性-短髪 ( )
女性-短パン ( )	短パン-半袖 ( )
60 以下-長袖 ( )	サングラス有-女性 ( )
女性-60 以下 ( )	長髪-短パン ( )
30 以下-スーツ下 ( )	女性-短髪 ( )
サングラス無-スーツ下 ( )	長髪-スーツ下 ( )
女性-30 以下 ( )	スーツ上-半袖 ( )
サングラス無-スーツ上 ( )	45 以下-長袖 ( )
スカート-長袖 ( )	サングラス有-45 以下 ( )
30 以下-短パン ( )	30 以下-短髪 ( )
サングラス有-長袖 ( )	長髪-スカート ( )
男性-半袖 ( )	60 以下-短髪 ( )
長髪-長袖 ( )	60 超-半袖 ( )
禿-長袖 ( )	45 以下-半袖 ( )
男性-スーツ下 ( )	女性-半袖 ( )
サングラス無-60 以下 ( )	女性-長髪 ( )
サングラス有-男性 ( )	スーツ上-スーツ下 ( )
60 超-禿 ( )	サングラス無-女性 ( )
サングラス無-スカート ( )	男性-45 以下 ( )
60 超-スーツ上 ( )	60 以下-スーツ下 ( )
短髪-スーツ上 ( )	長髪-半袖 ( )
サングラス無-45 以下 ( )	サングラス有-30 以下 ( )
60 超-長髪 ( )	サングラス無-30 以下 ( )
45 以下-スカート ( )	長髪-スーツ上 ( )
サングラス無-短髪 ( )	男性-60 以下 ( )
男性-禿 ( )	サングラス無-半袖 ( )
男性-短パン ( )	30 以下-スカート ( )
サングラス有-長髪 ( )	サングラス無-60 超 ( )
サングラス無-禿 ( )	45 以下-スーツ上 ( )
サングラス無-長髪 ( )	60 超-長袖 ( )
30 以下-長髪 ( )	サングラス無-男性 ( )
30 以下-長袖 ( )	短髪-長袖 ( )
短パン-長袖 ( )	短髪-短パン ( )

以内の辞書画像に正しい人物が含まれている割合を表す。nAUC は CMC 曲線下の面積であり 0 から 100 までの値をとる。

### 5.1.2 評価結果

評価結果について、CMC 曲線を図 5 に  $n$  位正解率と nAUC を表 3 に示す。比較対象の既存手法として

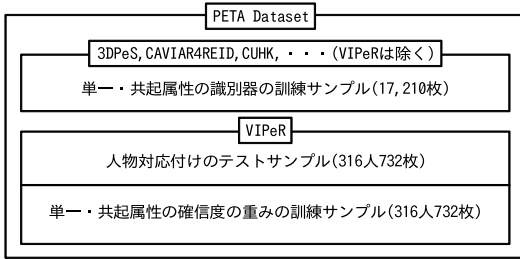


図 4 実験に用いた PETA Dataset の内訳  
Fig. 4 Details of PETA Dataset for experiments

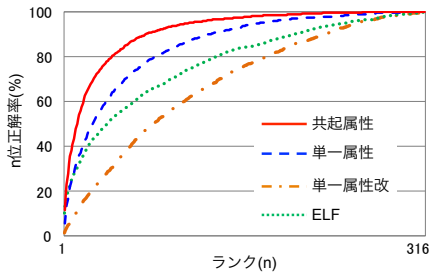


図 5 VIPeR の CMC 曲線  
Fig. 5 CMC curve of VIPeR

表 3 共起属性と単一属性の n 位正解率と nAUC の比較  
Table 3 Comparing nAUC between single or co-occurrence attributes

	$n = 1$	$n = 5$	$n = 10$	$n = 20$	nAUC
共起属性	11.3	32.0	45.6	63.9	91.6
単一属性	5.5	19.2	29.8	43.9	85.1
単一属性改	1.2	5.4	9.4	17.8	68.5
ELF	10.1	21.7	28.2	38.5	77.8

単一属性 [8] と画像特徴量の ELF [4] を用いた。ただし、本実験で用いた単一属性の実装は文献 [8] と基本的な流れは同じであるが、表 1 に記載されている単一属性を用いたこと、属性識別器の特徴量が ELF であること、人物対応付けの重み計算が LMNN であることが異なる。このように実装は異なるものの表 3 の単一属性を用いた人物対応付けの性能は、文献 [8] の Table 2 の条件 W.AIR に記載されている結果とほぼ同等であった。さらに比較のため、文献 [7] で述べられている context-level attribute classifier の共起の取り方を単一属性に組み込んだ実験を行った (単一属性改)。図 5 より、提案手法の共起属性は既存手法と比較すると人物対応付けの性能が向上していることが分かる。また、単一属性改の共起の取り方では性能が大きく低下することが分かった。表 3 より、 $n = 20$  の時に注目すると、単一属性を用いる既存手法と比較して 20.0

表 4 人物対応付けの正解数の比較

Table 4 Comparing the number of correct answers for person re-identification

	単一属性で成功	単一属性で失敗
共起属性で成功	1100	718
共起属性で失敗	144	882

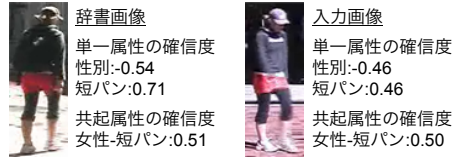


図 6 共起属性で対応付けに成功した例  
Fig. 6 Examples of correct matching using co-occurrence attributes

ポイント、ELF と比較して 25.4 ポイント性能が向上している。また、nAUC が単一属性を用いる既存手法と比較して 6.5 ポイント、ELF と比較して 13.8 ポイント向上していることが確認できる。以上より、提案手法の共起属性を用いた人物対応付けが有効であることが分かった。

次に、提案手法と既存手法で対応付けに成功した人物数の違いについて比較した。共起属性を用いた手法でのみ成功した人数、単一属性を用いた手法でのみ成功した人数、共に成功した人数、共に失敗した人数を表 4 に示す。ここでは CMC が  $n = 20$  の時を対象とした。この結果から単一属性で失敗し共起属性で成功した人数が、単一属性で成功し共起属性で失敗した人数と比較しておよそ 5 倍あることが分かった。図 6 に単一属性で失敗し共起属性で成功した人物画像の例を挙げる。この人物画像は短パンを着用した女性であることが特徴的であるが、それらの属性に対する確信度をみると、単一属性と比べて共起属性の方が辞書と入力との間で差が小さいことから、対応付けに共起属性が有効に働いたと考えられる。

最後に、身体同士、外見同士、または、身体と外見の組合せの内どれが人物対応付けに有効であるかを検証した。全ての組合せを用いた表 3 の共起属性 96 種類から特定の組合せを除くことで以下の 4 パターンを設定した。

- P1: 除外なし (共起属性 96 種類)
- P2: 身体と外見の組合せ 59 種類を除外
- P3: 身体同士の組合せ 22 種類を除外
- P4: 外見同士の組合せ 15 種類を除外

パターン P4 は表 3 の共起属性と同じである。実験結



表 5 手がかりの組合せ方の違いによる  $n$  位正解率と nAUC (全 96 種類から除外)Table 5 Comparison of  $n$ -th accuracy rate and nAUC due to differences in combinations of clues selected from all 96 attributes

	$n = 1$	$n = 5$	$n = 10$	$n = 20$	nAUC
P1	11.3	32.0	45.6	63.9	91.6
P2	7.9	24.5	37.4	54.1	89.2
P3	9.9	29.5	42.1	58.6	90.9
P4	11.6	30.6	44.0	60.5	91.4

表 6 手がかりの組合せ方の違いによる  $n$  位正解率と nAUC (全 45 種類から除外)Table 6 Comparison of  $n$ -th accuracy rate and nAUC due to differences in combinations of clues selected from 45 attributes

	$n = 1$	$n = 5$	$n = 10$	$n = 20$	nAUC
P'1	5.7	20.0	31.2	46.7	87.0
P'2	4.5	16.1	25.7	39.8	83.9
P'3	5.3	17.7	28.6	43.4	85.7
P'4	4.1	14.7	23.7	36.5	82.3

果について  $n$  位正解率と nAUC を表 5 に示す。ただし、この結果は除外する共起属性の数の違いに影響されることも考えられるため、共起属性の数を揃えた場合についても実験を行った。外見同士の組合せが 15 種類と最小であったため、他の組合せをそれぞれランダムに 15 種類ほど選択した。これら共起属性 45 種類から以下の 4 パターンを設定した。

- P'1: 除外なし (共起属性 45 種類)
- P'2: 身体と外見の組合せ 15 種類を除外
- P'3: 身体同士の組合せ 15 種類を除外
- P'4: 外見同士の組合せ 15 種類を除外

ランダム選択の影響を避けるために試行を 5 回行い、 $n$  位正解率と nAUC の平均を求めた結果を表 6 に示す。表 5 の P1 と P4 を比較すると性能がほとんど低下していないのに対し、表 6 の P'1 と P'4 を比較すると性能が大きく低下していた。このことから、外見同士の組合せは対応付けの性能に寄与するが、表 5 の実験のように他の組合せと比べて数が小さい場合は影響が少なくなることが分かった。一方で、P1 と P2 および P'1 と P'2 を比較すると、共に対応付けの性能が低下していることが分かる。このことから、身体的な手がかりと外見の手がかりの組合せは、人物対応付けの性能に影響が大きいことが分かる。提案手法の共起の取り方では、身体と外見の組合せ数が他の組合せ数に対して大きくなる場合が多いため、性能への影響が増えてくると考えられる。ただし、全ての組合せを

使用する場合と比較すると、表 5 と表 6 のいずれのパターンも性能が低下している。従って、人物対応付けの性能には、身体的な手がかりと外見の手がかりの組合せの影響が大きい、身体的な手がかり同士や外見の手がかり同士の組合せも影響があることが分かった。

## 5.2 計量行列の解析

### 5.2.1 解析方法

人物対応付けに有効な単一属性と共起属性を調査するために、LMNN で得られた計量行列  $M$  を解析する。まず解析の流れについて述べる。行列  $M$  は半正定値行列であるため式 (6) に変形できる。

$$M = \sum_{i=1}^N \lambda_i q_i q_i^T \quad (6)$$

ここで、 $\lambda_i$  は固有値を表し  $q_i$  は固有ベクトルを表す。固有値  $\lambda_i$  が大きいほどその固有値に対応する  $i$  番目の固有ベクトルが行列  $M$  をよく表しており、特徴量間の距離を算出する際に大きな影響を及ぼす。 $m$  番目に大きい固有値  $\lambda_m$  までの累積寄与率  $C_m$  は式 (7) で決まる。

$$C_m = \frac{\sum_{i=1}^m \lambda_i}{\sum_{j=1}^N \lambda_j} \quad (7)$$

さらに、固有ベクトル  $q_i$  中の  $n$  番目の要素  $q_{i,n}$  がその固有ベクトルの中でどの程度の距離計算に役割を果たすかについて重み割り合い  $w_{i,n}$  を式 (8) で求める。

$$w_{i,n} = \frac{|q_{i,n}|}{\sum_{j=1}^N |q_{i,j}|} \quad (8)$$

降順ソートされた  $w_{i,1}$  から  $w_{i,n}$  までの和である累積重み割り合い  $W_{i,n}$  を式 (9) で求めることで、 $n$  番目までの要素がある固有ベクトルの中で占める割合がわかる。

$$W_{i,n} = \sum_{j=1}^n w_{i,j} \quad (9)$$

### 5.2.2 解析結果

共起属性の計量行列と単一属性の計量行列との間で、累積寄与率と累積重み割合が同等になる場合について比較した。共起属性の計量行列として、組合せ数を揃えた表 6 のパターン P'1 の実験で生成したものをを用いた。単一属性の計量行列として、表 3 の実験で生成し

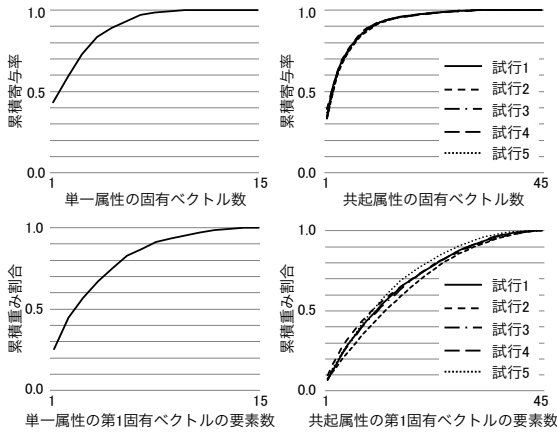


図 7 単一属性の計量行列と共起属性の計量行列に対する累積寄与率と累積重み割合  
 Fig. 7 Cumulative contribution ratio and cumulative weight ratio of attributes.

たものを用いた。ただし、前節までの実験では VIPeR に含まれる 316 人のデータから計量行列を生成していたが、この実験では VIPeR の全データ 632 人から計量行列を生成した。単一属性の計量行列、および、パターン P'1 の試行毎の計量行列について、累積寄与率と累積重み割合を図 7 に示す。この実験では、累積寄与率  $C_m$  が 0.4 を超えるまでの固有値に対応する固有ベクトルを解析することとした。単一属性では第 1 固有ベクトルのみが該当し、共起属性では各試行とも第 1 から第 2 固有ベクトルが該当した。累積重み割合  $W_{i,n}$  は 0.2 を超えるまでの要素を対象とし、単一属性では第 1 固有ベクトルから 1 要素、共起属性では第 1 固有ベクトルから平均で 4 要素、第 2 固有ベクトルから平均で 4 要素が該当した。選ばれた共起属性の各要素については表 7 に示す。なお、解析対象となる単一属性の 1 種類は短髪であった。この結果から、8 種類の共起属性が 1 種類の単一属性に相当する重みを有しており、共起属性が人物の特徴をより詳細に表現していると考えられる。次に表 7 の各要素について、身体と外見の組合せ数、身体同士の組合せ数、外見同士の組合せ数を比較した。その結果、身体と外見の共起属性が 16 種類、身体同士の共起属性が 4 種類、外見同士の共起属性が 22 種類であった。各組合せの総数は  $15(\text{共起属性数}) \times 2(\text{固有ベクトル数}) \times 5(\text{試行数}) = 150$  であった。表 8 に Fisher の直接確率法による  $p$  値とオッズ比を示す。その結果、身体と外見の共起属性と身体同士の共起属性との間、外見同士の共起属性と身

表 7 計量行列の解析で得られた性能に寄与する共起属性の例

Table 7 Co-occurrence attributes contributed to the performance of person re-identification

(a) 試行 1

第 1 固有ベクトル	第 2 固有ベクトル
スーツ下-長袖 ( )	サングラス無-長袖 ( )
サングラス無-スーツ下 ( )	短パン-長袖 ( )
45 以下-長髪 ( )	短髪-スーツ下 ( )
サングラス無-スーツ上 ( )	60 超-サングラス無 ( )

(b) 試行 2

第 1 固有ベクトル	第 2 固有ベクトル
スーツ下-長袖 ( )	短パン-長袖 ( )
サングラス無-スーツ下 ( )	スーツ下-半袖 ( )
短髪-半袖 ( )	サングラス無-スカート ( )
短髪-スカート ( )	サングラス無-長袖 ( )
短髪-スーツ上 ( )	

(c) 試行 3

第 1 固有ベクトル	第 2 固有ベクトル
スーツ下-長袖 ( )	短髪-スーツ下 ( )
45 以下-長髪 ( )	45 以下-サングラス無 ( )
女性-スカート ( )	60 以下-サングラス無 ( )
サングラス無-短パン ( )	30 以下-短髪 ( )

(d) 試行 4

第 1 固有ベクトル	第 2 固有ベクトル
スーツ下-長袖 ( )	秃-半袖 ( )
サングラス無-スーツ下 ( )	30 以下-サングラス無 ( )
60 超-半袖 ( )	短髪-短パン ( )
	スーツ下-半袖 ( )
	女性-半袖 ( )

(e) 試行 5

第 1 固有ベクトル	第 2 固有ベクトル
スーツ下-長袖 ( )	スーツ下-スーツ上 ( )
45 以下-短パン ( )	サングラス無-長袖 ( )
サングラス無-スーツ下 ( )	短パン-長袖 ( )
サングラス無-スーツ上 ( )	30 以下-長髪 ( )
	45 以下-スカート ( )

体同士の共起属性との間で有意な差がみられた。なお、 $C_m$  が 0.6 を超えるまでかつ  $W_{i,n}$  が 0.3 を超えるまでとした実験を行った結果、身体と外見の共起属性が 41 種類、身体同士の共起属性が 22 種類、外見同士の共起属性が 44 種類となり、種類の数は増えたものの同様の傾向が見られた。以上より、組合せ数を揃えた場合には、身体的な手がかりと外見的な手がかりの組合せ、および、外見的な手がかり同士の組合せが人物対応付けに有効であることが分かった。

次に、手がかりの組合せ数が異なる表 5 の P1 の実験で生成した計量行列について同じ  $C_m, W_{i,n}$  の設定で解析した。その結果、身体同士の共起属性が 6 種類、



表 8 Fisher の直接確率法による  $p$  値およびオッズ比の比較Table 8 Comparing  $p$ -value using Fisher's exact test and odds ratio

	該当数	非該当数
身体と外見の組合せ	16	134
身体同士の組合せ	4	146
$p$ 値=0.0092, オッズ比=4.34		
	該当数	非該当数
外見同士の組合せ	22	128
身体同士の組合せ	4	146
$p$ 値=0.0003, オッズ比=6.24		
	該当数	非該当数
外見同士の組合せ	22	128
身体と外見の組合せ	16	134
$p$ 値=0.6861, オッズ比=1.44		

表 9 複数の公開データベースを用いた人物対応付けの  $n$  位正解率と nAUC の比較

Table 9 Comparing nAUC on multiple public datasets for person re-identification

Dataset	手法	$n = 1$	$n = 5$	$n = 10$	$n = 20$	nAUC
3DPeS	共起	<b>20.0</b>	<b>39.8</b>	<b>51.0</b>	<b>65.0</b>	<b>78.6</b>
	単一	14.1	31.2	42.2	56.9	74.2
CAVIAR	共起	<b>24.7</b>	<b>55.8</b>	<b>72.3</b>	<b>90.1</b>	<b>81.7</b>
	単一	15.6	40.7	59.2	80.2	72.9
4REID	共起	<b>18.9</b>	<b>46.4</b>	<b>65.1</b>	<b>80.3</b>	<b>89.9</b>
	単一	10.1	30.9	45.2	62.5	81.9
GRID	共起	<b>20.9</b>	<b>46</b>	<b>60.7</b>	<b>77.6</b>	<b>81.2</b>
	単一	10.7	32.8	48.4	68.3	73.7
i-LID	共起	<b>8.9</b>	<b>28.5</b>	<b>44.2</b>	<b>61.3</b>	<b>78.7</b>
	単一	4.1	19.5	31.4	50.6	71.3
PRID	共起	<b>60.9</b>	<b>94.5</b>	<b>98.9</b>	<b>99.5</b>	<b>96.5</b>
	単一	50.0	90.3	98.4	99.9	94.7
SARC3D	共起	<b>32.3</b>	<b>49.9</b>	<b>57.8</b>	<b>67.6</b>	<b>82.5</b>
	単一	18.4	35.1	44.2	55.1	74.6
Town Centre	共起	<b>32.3</b>	<b>49.9</b>	<b>57.8</b>	<b>67.6</b>	<b>82.5</b>
	単一	18.4	35.1	44.2	55.1	74.6

外見同士の共起属性が 3 種類, 身体と外見の共起属性が 13 種類であった。身体と外見の共起属性の数は多いものの Fisher の直接確率法で検定したところ組合せ数の違いにより有意な差はみられなかった。このため, 表 2 の 96 種類の共起属性を用いた場合はどの組合せ方も性能に寄与していると考えられる。

### 5.3 複数データベースにおける性能向上の確認

これまでの実験で用いた VIPeR 以外のデータベースでも提案手法が有効に働くことを確認するために, 3DPeS, CAVIAR4REID, GRID, i-LID, PRID, SARC3D, TownCentre で人物対応付けの性能を評価した。VIPeR を用いた実験と同様に人物対応付けの評価データベースと属性識別の訓練サンプルを分けた。実験結果について  $n$  位正解率と nAUC を表 9 に

示す。この結果より, 全てのデータベースにおいて提案手法の共起属性を用いた人物対応付けが既存手法の単一属性を用いた人物対応付けよりも有効であることが分かった。特に人物対応付けの評価に広く用いられている i-LID や PRID においても提案手法が有効であることが分かった。

## 6. おわりに

本論文ではカメラ間の人物対応付けを目的として身体と外見の共起属性を用いる手法を述べた。人が人を探す際の知見をもとに身体的な手がかりと外見の手がかりの組合せから共起属性を設計した。実験結果より, 共起属性を用いる提案手法が単一属性を用いる既存手法と比較して人物対応付けの性能向上に有効であることを複数の公開データベースを用いて確認した。また, 人物対応付けに寄与する共起属性を明らかにするために, 計量行列の解析を行い, 身体的な手がかりと外見の手がかりのどの組合せが有効であるか確認した。

今後の課題として, 性能に寄与する共起属性の種類解析, 3 種類以上の属性の組合せから共起をとる場合の検証, 共起属性と低次特徴量との組合せによる人物対応付けの性能向上が挙げられる。

## 文 献

- [1] C. Nakajima, M. Pontil, B. Heisele, and T. Poggio. Full-body person recognition system. *Proceedings of Pattern recognition*, Vol. 36, No. 9, pp. 1997–2006, 2003.
- [2] X. Wang, G. Doretto, T. Sebastian, J. Rittscher, and P. Tu. Shape and appearance context modeling. *Proceedings of International Conference on Computer Vision*, pp. 1–8, 2007.
- [3] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani. Person re-identification by symmetry-driven accumulation of local features. *Proceedings of Computer Vision and Pattern Recognition*, pp. 2360–2367, 2010.
- [4] D. Gray and H. Tao. Viewpoint invariant pedestrian recognition with an ensemble of localized features. *Proceedings of European Conference on Computer Vision*, pp. 262–275, 2008.
- [5] W. Li, R. Zhao, T. Xiao, and X. Wang. Deep-reid: Deep filter pairing neural network for person re-identification. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 152–159, 2014.
- [6] J. Roth and X. Liu. On the exploration of joint attribute learning for person re-identification. *Proceedings of 12th Asian Conference on Computer Vision*,

- pp. 673–688, 2014.
- [7] L. Bourdev, S. Maji, and J. Malik. Describing people: A poselet-based approach to attribute classification. *Proceedings of International Conference on Computer Vision*, pp. 1543–1550, 2011.
- [8] R. Layne, T. Hospedales, and S. Gong. Person re-identification by attributes. *Proceedings of British Machine Vision Conference*, pp. 93–117, 2012.
- [9] R. Layne, T. Hospedales, and S. Gong. Person re-identification. *chapter Attributes-Based Re-identification*, pp. 93–117, 2014.
- [10] K. Sameh, K. Cheng-Hao, K.S. Vivek, Vinay D.S., and Larry S.D. Joint learning for attribute-consistent person re-identification. *Proceedings of ECCV Workshop on Visual Surveillance and Re-Identification*, pp. 134–146, 2014.
- [11] Z. Shi, T.M. Hospedales, and T. Xiang. Transferring a semantic representation for person re-identification and search. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4184–4193, 2015.
- [12] 警察庁. 生活安全の確保. <https://www.npa.go.jp/safetylife/seianki/fumei/index.html>.
- [13] 警視庁. 公開捜査ファイル. <http://www.keishicho.metro.tokyo.jp/jiken/jiken.htm>.
- [14] Police Scotland. Missing persons. <http://www.scotland.police.uk/whats-happening/missingpersons/james-barr>.
- [15] Police Scotland. Most wanted. <http://www.scotland.police.uk/whats-happening/most-wanted/>.
- [16] 川西康友, 新村文郷, 出口大輔, 村瀬洋. サーベイ論文: 画像からの歩行者属性認識. 信学技報, pp. 117–127.
- [17] A. Dantcheva, C. Velardo, A.D. Angelo, and J. Dugelay. Bag of soft biometrics for person identification. *Journal of Multimedia Tools and Applications*, Vol. 51, No. 2, pp. 739–777, 2011.
- [18] 麻生英樹, 津田宏治, 村田昇. パターン認識と学習の統計学 新しい概念と手法. 岩波書店, 2003.
- [19] S. Liao and S.Z. Li. Efficient psd constrained asymmetric metric learning for person re-identification. *Proceedings of International Conference on Computer Vision*, pp. 11–18, 2015.
- [20] K.Q. Weinberger, J. Blitzer, and L.K. Saul. Distance metric learning for large margin nearest neighbor classification. *Journal of Machine Learning*, pp. 207–244, 2009.
- [21] Yubin Deng, Ping Luo, Chen Change Loy, and Xiaou Tang. Pedestrian attribute recognition at far distance. pp. –, 2014.
- [22] D. Baltieri, R. Vezzani, and R. Cucchiara. 3dpes: 3d people dataset for surveillance and forensics. 2011.
- [23] D.S. Cheng, M. Cristani, M. Stoppa, L. Bazzani, and V. Murino. Custom pictorial structures for re-identification. *Proceedings of British Machine Vision Conference*, p. 6, 2011.
- [24] W. Ouyang and X. Wang. A discriminative deep model for pedestrian detection with occlusion handling. *Proceedings of Computer Vision and Pattern Recognition*, pp. 3258–3265, 2012.
- [25] C.C. Loy, T. Xiang, and S. Gong. Time-delayed correlation analysis for multi-camera activity understanding. *International Journal of Computer Vision*, Vol. 90, No. 1, pp. 106–129, 2010.
- [26] T. Wang, S. Gong, X. Zhu, and S. Wang. Person re-identification by video ranking. *Proceedings of European Conference on Computer Vision*, pp. 688–703, 2014.
- [27] C. Papageorgiou and T. Poggio. A trainable system for object detection. 2000.
- [28] M. Hirzer, C. Beleznaï, P. M Roth, and H. Bischof. Person re-identification by descriptive and discriminative classification. pp. 91–102. 2011.
- [29] D. Baltieri, R. Vezzani, and R. Cucchiara. Sarc3d: a new 3d body model for people tracking and re-identification. pp. 197–206. 2011.
- [30] B. Benfold and I. Reid. Stable multi-target tracking in real-time surveillance video. pp. 3457–3464, 2011.
- [31] D. Gray, S. Brennan, and H. Tao. Evaluating appearance models for recognition, reacquisition, and tracking. *Proceedings of Performance Evaluation of Tracking and Surveillance*, pp. 1–7, 2007.

(平成 xx 年 xx 月 xx 日受付)

#### 中野 翔太

2014 鳥取大学工学部卒業 . 2016 鳥取大学大学院工学研究科情報エレクトロニクス専攻博士前期課程了 . 2016 三菱電機インフォメーションネットワーク株式会社に勤務 .

#### 四元 辰平

昭和 50 年生 . 平成 8 年国立鹿児島工業高等専門学校情報工学科卒業 . 同年 , メルコ・パワー・システムズ株式会社入社 . 平成 15 年放送大学卒業 . 平成 26 年より鳥取大学大学院工学研究科情報エレクトロニクス専攻博士後期課程に在学中 . 主に映像蓄積配信・人物追跡に関する研究開発に従事 . 情報処理学会会員

#### 吉村 宏紀 (正員)

1993 年鳥取大学工学部知能情報工学科卒業 . 1997 年鳥取大学大学院工学研究科

博士後期課程情報生産工学専攻修了。博士(工学)。現在鳥取大学大学院工学研究科情報エレクトロニクス専攻, 助教。画像, 音声信号処理の研究に従事。電気学会, 情報

処理学会の会員。

西山 正志 (正員: シニア会員)

2002 岡山大学大学院博士前期課程了。株式会社東芝研究開発センターを経て, 現在鳥取大学大学院工学研究科准教授。学際情報学博士(東京大学)。画像認識, 拡張現実感の研究に従事。2006 本会パターン認識・メディア理解研究会研究奨励賞, 2007 画像センシングシンポジウム優秀論文賞 2009 山下記念研究賞などを受賞。情報処理学会会員

岩井 儀雄 (正員)

1992 年(平成 4 年)大阪大学基礎工学部情報工学科卒業。1997 年(平成 9 年)大阪大学大学院基礎工学研究科博士課程後期修了。同年同大学院助手。2003 年(平成 15 年)同大学院助教授。2004 年(平成 16 年)5 月~2005 年(平成 17 年)3 月英国ケンブリッジ大学客員研究員。2007 年(平成 19 年)同大学院准教授。2011 年(平成 23 年)鳥取大学大学院工学研究科教授。コンピュータビジョン, パターン認識の研究に従事。博士(工学)

菅原 一孔 (正員)

昭和 31 年生。昭和 56 年東京工業大学大学院理工学研究科電子物理工学専攻修士課程修了。同年神戸市立工業高等専門学校電気工学科講師。同校助教授を経て平成 6 年鳥取大学工学部電気電子工学科助教授, 現在, 同大学大学院工学研究科情報エレクトロニクス専攻教授。計算機工学に関する研究に従事。平成 22 年 日刊工業新聞社 モノづくり連携大賞, 平成 21 年 船井ベストペーパー賞, 平成 21 年 総務大臣賞 産学官連携功労者表彰, 平成 20 年 総務大臣表彰 U-Japan 大賞など受賞。工学博士。情報処理学会会員