

# 映像追跡技術を用いた再特定手法

吉田 諭史<sup>†</sup> 西村 祥治<sup>†</sup> 劉 健全<sup>†</sup>

<sup>†</sup> 日本電気株式会社 システムプラットフォーム研究所 〒211-8666 神奈川県川崎市中原区下沼部 1753  
E-mail: †{s-yoshida@ez, s-nishimura@bk, j-liu@ct}.jp.nec.com

あらまし 本論文では、複数のカメラ映像から、同一人物の追跡を行う方法を提案する。映像データから動いている人を追跡する技術には映像追跡技術があるが、現状では、複数のカメラを用いての追跡は難しい。また、複数の画像データ上で同じ人を特定する技術に再特定技術があるが、精度が低いという問題がある。そこで本論文では、映像追跡技術と再特定技術とを組み合わせることで、複数のカメラ間でも精度の高い追跡を行うことができる手法を提案する。

キーワード 映像分析, 映像追跡, 再特定

## 1. はじめに

IoT 技術やストレージ技術の発展により、膨大な数の監視カメラが利用可能となっており、莫大な量の映像データが蓄積されている [8]。映像データから、条件を満たす人が、どこから来て、何を、どこに行ったのかを抽出することは重要な問題となっている。例えば、犯罪捜査においては、防犯カメラ映像から、目撃情報を満たす人を探し出してきて、その人がどこに行ったのかを分析することが行われている。ところが、莫大な映像データの分析が人手で行われていることも多く、映像分析官に多大な負担がかかっている。

現状では、上記の問題を解く手法として、映像追跡技術と再特定技術とがある。映像追跡技術は、映像を入力として、映像中で移動している物体を追跡する技術である。映像追跡技術には様々な手法があるが、いずれの方法も前後のフレーム間で物体の移動を抽出するため、複数カメラ間での映像追跡は難しい。また、複数カメラで映像追跡する試みもあるが、カメラの視野を共有しなかったり、カメラ同士が近くなったりする場合は、物体の移動が抽出できなくなる。一方で再特定技術は、複数の映像や画像を入力として、同じ人物を発見する技術である。基本的には、画像や映像同士の類似度を計算して、類似度が高いものがあれば、それらを同一人物と判定する。静止画像を用いた再特定技術には精度があまり低いという問題がある。例となる顔画像を図 1 に示す。これは、同一人物から得られる特徴量はお互いの類似度が低くなる場合があるためである。特に、顔の特徴量に関しては、以下の性質が観察されている。

観察 1. NeoFace<sup>®</sup>(<sup>注1</sup>) を用いて抽出された人の顔の特徴量について、以下の 2 つの性質が経験的に知られている。

- (1) 同一人物の顔の特徴量間は、類似度が高いものと低いものがある。
- (2) 異なる人物の顔の特徴量間は、類似度が高くなることはない。

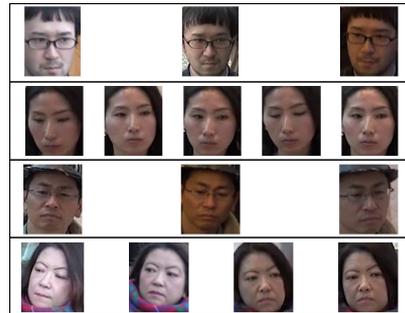


図 1 お互いの類似度が低くなる画像の例。枠内の画像はすべて同一人物のように見えるが、お互いの類似度が低い。そのため、単純に比較すると別人と判定される。

また、映像に対する再特定技術は、画像集合同士で同一人物判定を行う。既存の方法 [1, 6, 7, 11, 13–15, 26, 27, 29] では、機械学習によって画像集合のモデルを生成するため、一般に遅い。

そこで本論文では、映像追跡技術と再特定技術とを組み合わせることで、複数カメラ間でも精度が高い追跡を高速に行うことができる手法を提案する。この手法は、映像追跡技術によって抽出された同一人物の画像群より特徴量を抽出し、再特定技術によって画像群同士で同一人物を発見を行う。ここで抽出される特徴量は、観察 1 を満たすと仮定する。映像追跡技術を用いることで、特徴量の類似度以外の情報を用いて同一人物の特徴量群が得られる。そのため、再特定技術によって、同一人物が発見しやすくなると期待される。本論文ではさらに、再特定技術によって同一人物を発見する際に、類似検索を高速化するためのためのデータ構造である類似度木 [30] を用いることで再特定を高速に行う手法も提案する。2 つの画像群同士で同一人物かどうかを判定するには、単純な方法では 2 つの画像群に含まれる各画像同士で類似度を計算することになるため、画像群のサイズの 2 乗の計算量がかかる。そこで、類似度木を利用して、画像群同士が同一人物のものであるかの判定に不必要な類似度計算を省略することで、画像群のサイズ程度まで計算量を削減できる。

本論文は以下のように構成されている。第 2 節では、本論文

(注1): NEC による顔認証技術。 <https://jpn.nec.com/face/>.

表 1 本論文で用いる記号.

記号	意味
$f$	特徴量
$\text{sim}(\cdot, \cdot)$	2 つの特徴量間の類似度
$\delta$	類似度の閾値
$\delta(\cdot)$	類似度木中の代表データが持つ閾値
$n$	特徴量集合の数
$m$	特徴量集合に含まれる特徴量の数

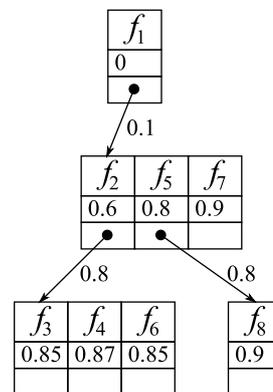


図 2 類似度木の例.

で用いる記号の説明を行い、類似度と類似検索の定義を行う。第 3 節では、従来技術である映像追跡技術と再特定技術、類似度木の概要を説明する。第 4 節では、本論文の提案手法である、映像追跡技術を用いた再特定技術について説明する。第 5 節では、計算機実験を行い、本手法の有効性を実験的に実証する。最後に第 6 節で、本論文の結論を述べる。

## 2. 記法と用語

本論文で用いる記号は表 1 のとおりである。本論文で扱うデータまたは特徴量は、それらに類似度が定義される。類似度は、以下の 3 つの条件を満たす。

- すべてのデータ  $f_i, f_j$  について、 $0 \leq \text{sim}(f_i, f_j) \leq 1$ 。
- すべてのデータ  $f_i$  について、 $\text{sim}(f_i, f_i) = 1$ 。
- すべてのデータ  $f_i, f_j$  について、 $\text{sim}(f_i, f_j) = \text{sim}(f_j, f_i)$ 。

また、本論文では、すべてのデータ  $f_i, f_j, f_k$  と  $\delta$  について、 $\text{sim}(f_i, f_j) \geq \delta$  かつ、 $\text{sim}(f_j, f_k) \geq \delta$  を満たすならば、 $\text{sim}(f_i, f_k) \geq \delta$  を満たすと類推し、以下では、これが真であると仮定して議論を進める。

類似検索は、データの集合  $D$  と、クエリとなるデータ  $q$ 、閾値  $\delta \in \mathbb{R}$  とが与えられ、 $D$  内の  $q$  との類似度が  $\delta$  以上となるデータをすべて出力する問題である。すなわち、出力されるデータの集合は、 $\{d \in D \mid \text{sim}(d, q) \geq \delta\}$  で表される。

## 3. 関連技術

本節では、本論文で提案する技術の関連技術である映像追跡技術と再特定技術、類似度木について簡単に述べる。

### 3.1 映像追跡技術

映像追跡技術は、映像を入力として、映像中に現れる特定の物体について、映像中の各フレームにおけるその物体の位置を特定することで、その物体を追跡する技術である。映像追跡技術は、いくつかの種類に大別される。特徴量ベースの追跡 [17, 22, 25] は、映像中の各フレームから、物体の特徴点を抽出し、フレーム間の特徴点の移動を算出することで物体を追跡する。カーネルベースの追跡 [4, 20, 24] は、追跡する物体を覆う楕円や四角などの単純な図形からなる領域の動きを前後フレーム間で計算することで追跡を行う。シルエットベースの追跡 [2, 3, 18] は、対象の物体の形状や領域を、前のフレームの形状や領域の情報を使いながら抽出する。

### 3.2 再特定技術

再特定技術は、複数の映像や画像を入力として、それらの中

から、同一人物を発見する技術である。例えば、静止画像の中の色情報を用いて、同一人物を発見する手法が提案されている [5, 7, 28]。また、色やテキストチャなどから属性を学習して同一人物を発見する手法も提案されている [10, 12, 21, 23]。深層学習を用いる手法も提案されている [9, 11, 16, 19, 27]。映像から同一人物を発見する手法も提案されている。これらの手法は、映像の各フレームを入力とし、深層学習 [6, 11, 15, 26, 27, 29] や統計的機械学習 [1, 7, 13, 14] によりモデルを作成して同一人物を発見する。

### 3.3 類似度木

類似度木 [30] は、類似検索問題を高速に解くためのデータ構造である。類似度木は、以下の 4 つの性質をすべて満たす木構造である。

- 各ノードには複数のデータが格納されている。
- 一部のデータは代表データと呼ばれ、子ノードを持つ。
- 各代表データ  $f$  は、閾値  $\delta(f)$  を持つ。
- 各代表データ  $f$  について、 $f$  の子ノードに含まれるすべてのデータ  $g$  は  $\text{sim}(f, g) \geq \delta(f)$  を満たす。

類似度木の例を図 2 に示す。各ノードは四角で表されており、データ  $f_2, f_5, f_7$  とデータ  $f_3, f_4, f_6$  はそれぞれ同じノードに格納されている。代表データとそれに対応する子ノードとの関係は矢印で表されている。また、各データの下に数字は、そのデータと、親となる代表データとの間の類似度を表す。矢印のそばにある数字は、代表データが持つ閾値を表す。この類似度木は、代表データ  $f_1$  の子ノードに、データ  $f_2, f_5, f_7$  が格納されている。代表データ  $f_1$  の閾値  $\delta(f_1)$  は 0.1 であり、 $f_1$  と  $f_2, f_5, f_7$  との間の類似度はそれぞれ 0.6, 0.8, 0.9 となっているため、上記 4 番目の性質を満たしている。代表データ  $f_2$  と  $f_8$  についても同様である。

類似度木を用いた類似検索は以下の要領で行われる。類似度木を深さ優先探索で辿っていき、類似度木内のデータとクエリデータとの間の類似度が所定の閾値以上であれば答えとして出力する。探索の際、ある代表データ  $f$  とクエリデータ  $q$  との類似度が代表データ  $f$  が持つ閾値未満、すなわち、 $\text{sim}(f, q) < \delta(f)$  ならば、代表データ  $f$  以下にクエリデータ  $q$  と類似するデータは存在しないと判断し、代表データ  $f$  以下の探索は行わない。

#### 4. 映像追跡技術を用いた再特定技術

本節では、本論文の提案手法である、映像追跡技術を用いた再特定技術の詳細を説明する。本手法の大まかな手順は次のとおりである。まず、各カメラの映像に対して、人物追跡技術を適用し、対象となる人物を追跡する。この手順によって、同一人物の画像の集合が延べの人数分得られる。次に画像集合から特徴量を抽出して、特徴量の集合とする。このときの特徴量は、観察 1 を満たすと仮定する。最後に、特徴量の集合同士で類似度を計算し、類似度が高ければ、それらの集合は同一人物から得られたものと判断する。以下では、特徴量の集合同士が同一人物から得られたか否かを判断する方法について述べる。

特徴量の集合同士での同一人物判定は、観察 1 より、画像集合間の各画像間で類似度の計算を行い、類似度が高いものがあれば同一人物とすればよいことがわかる。この比較を単純に行うと、各特徴量集合同士の組合せの各特徴量の組で特徴量間の類似度の計算をすることになる。そのため、最大で  $n^2m^2$  回の類似度計算をすることになる。計算回数を削減するためには 2 つのアプローチがある。一方は、2 つの特徴量集合を比較する際に類似度を計算する回数を削減するアプローチで、もう片方は、比較する特徴量集合の組合せの数を削減するアプローチである。以降では、計算量を削減するための方法を各アプローチについて 1 つずつ述べる。

特徴量集合同士を比較する際に、類似度を計算する回数を削減する手法について述べる。2 つの特徴量集合に含まれる特徴の組の中で、類似度が高いものがあるかないかがわかればよい。そのため、類似度が高くなると期待される部分のみを探索し、そうではない部分の探索は行わないようにする。これを実現するために、類似度木を用いる。すなわち、特徴量集合ごとに類似度木を作成しておき、2 つの類似度木の間で類似度が高い部分のみを探索することで、類似度の計算する回数を削減する。具体的には、以下のように行う。類似度木の各階層について、階層同士の代表データの組同士で類似度の計算をする。計算の結果、類似度が最大となる組について着目し、次の比較では、着目した代表データの子ノードのみで類似度の計算を行っていく。このアルゴリズムをアルゴリズム 1 に示し、図 3 を使って説明する。まず 5 行目から 7 行目で、2 つの類似度木で、それぞれ 1 つのノードの中に含まれるデータの組で類似度の比較をする。図ではデータ  $f_2, f_4$  と  $f_8, f_{11}$  の組合せで類似度の比較をしている。破線矢印は類似度を計算するデータの組を表す。すなわち、データ  $f_2$  と  $f_8, f_2$  と  $f_{11}, f_4$  と  $f_8, f_4$  と  $f_{11}$  の組で類似度を計算し、比較する。この中で、データ  $f_4$  と  $f_8$  の組が類似度最大であったと仮定する。その後、14 行目から 19 行目で類似度最大となるデータの子ノードに着目し、同様に類似度の比較を行う。図では  $f_4$  と  $f_8$  の子ノードに着目している。8 行目から 10 行目では、類似度が事前に決めた閾値を超過した場合に同一人物であると判定している。

ここで、本手法の計算量について述べる。本手法は、類似度木を親から子へ辿ることで同一人物の判定を行っている。1 ノードあたりの代表データの数を定数とみなすと、計算量の上

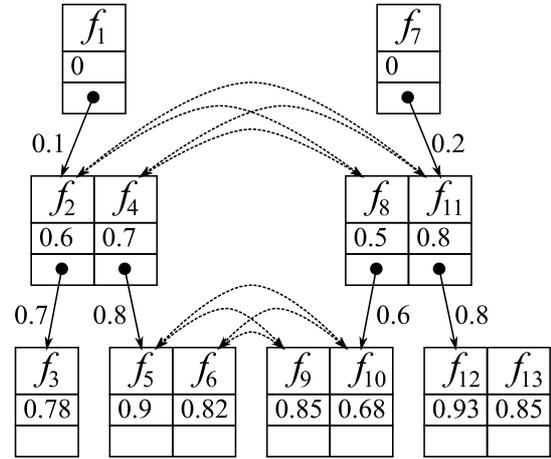


図 3 2 つの類似度木上での最も類似度が高い特徴量の検索。

---

#### Algorithm 1 最も類似度が高い特徴量の検索。

---

```

1: procedure FINDMOSTSIMILARFEATURE( $\mathcal{T}_1, \mathcal{T}_2, \sigma$ )  $\triangleright \mathcal{T}_1$  と  $\mathcal{T}_2$ 
   は類似度木,  $\sigma$  は閾値 .
2:    $\mathcal{N}_1 \leftarrow \mathcal{T}_1$  の根ノード .
3:    $\mathcal{N}_2 \leftarrow \mathcal{T}_2$  の根ノード .
4:   while True do
5:      $R_1 \leftarrow \mathcal{N}_1$  の代表データの集合 .
6:      $R_2 \leftarrow \mathcal{N}_2$  の代表データの集合 .
7:      $(f_1, f_2) \leftarrow \operatorname{argmax}_{f_1 \in R_1, f_2 \in R_2} \operatorname{sim}(f_1, f_2)$  .
8:     if  $\operatorname{sim}(f_1, f_2) > \sigma$  then
9:       return True
10:    end if
11:    if  $\mathcal{N}_1$  に  $f_1$  に対応する子ノードがない and  $\mathcal{N}_2$  に  $f_2$  に
      対応する子ノードがない then
12:      return False
13:    end if
14:    if  $\mathcal{N}_1$  に  $f_1$  に対応する子ノードがある then
15:       $\mathcal{N}_1 \leftarrow \mathcal{N}_1$  に  $f_1$  に対応する子ノード .
16:    end if
17:    if  $\mathcal{N}_2$  に  $f_2$  に対応する子ノードがある then
18:       $\mathcal{N}_2 \leftarrow \mathcal{N}_2$  に  $f_2$  に対応する子ノード .
19:    end if
20:  end while
21: end procedure

```

---

界は類似度木の高さでおさえられる。類似度木は、必ずしもバランスするとは限らないため、その高さは高々  $m$  となる。そのため、2 つの特徴量集合同士で同一人物の判定は  $O(m)$  時間かかる。また、2 つの特徴量集合の全ての組合せについて同一人物判定を行うため、全体として  $O(n^2m)$  時間かかる。

次に、比較する特徴量集合の組合せの数を削減する手法について述べる。この手法ではまず、特徴量集合ごとにカメラの位置と追跡開始時刻と終了時刻をメタデータとして保持しておく。保持しておいたメタデータ同士を比較し、同一人物として不適当な場合は、同一人物ではないと判定する。同一人物として不適当となる条件として、以下のものがある。

- (1) 1 つのカメラに同時に複数出現している
- (2) 遠い場所に短い時間で移動している

## Algorithm 2 メタデータによる類似性判定 .

```
1: procedure METADATA SIMILARITY EVALUATION( $\mathcal{F}, t, \sigma$ ) ▷  $\mathcal{F}$ 
   は特徴量集合の族,  $t$  は単調増加の関数,  $\sigma$  は類似度の閾値 .
2:   for all  $X, Y \in \mathcal{F}$  do
3:     if  $t(d(X, Y)) \leq s(Y) - e(X)$  and  $(s(Y) - e(X) > 0$  or
       $X$  と  $Y$  は異なるカメラ) then
4:        $\mathcal{T}_1 \leftarrow X$  内の特徴量から作成された類似度木 .
5:        $\mathcal{T}_2 \leftarrow Y$  内の特徴量から作成された類似度木 .
6:       FINDMOSTSIMILARFEATURE( $\mathcal{T}_1, \mathcal{T}_2, \sigma$ )
7:     end if
8:   end for
9: end procedure
```

この2つの条件を定式化すると、以下のとおりとなる。2つの特徴量集合を  $X, Y$  とする。以降では簡単のため、追跡開始時刻が前の方を  $X$ 、後の方を  $Y$  として議論を進める。なお、このようにしても一般性を損なうことはない。また、追跡開始時刻と追跡終了時刻、カメラ間の距離をそれぞれ  $s(\cdot)$  と  $e(\cdot)$ 、 $d(\cdot, \cdot)$  で表す。つまり、 $s(X) \leq s(Y)$  が成り立つ。このとき、1つ目の条件は、特徴量集合  $X$  と  $Y$  が同じカメラではないか、または、追跡している時間が重複しない、つまり、 $s(Y) - e(X) \geq 0$  である。また、2つ目の条件は、ある単調増加関数  $t: \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$  に対して、 $t(d(X, Y)) \leq s(Y) - e(X)$  が成立する。ここでの関数  $t$  は、カメラ間の距離から時間隔の閾値への関数となっている。特徴量集合  $X$  の追跡終了時刻と  $Y$  の追跡開始時刻の間隔  $s(Y) - e(X)$  が  $X$  と  $Y$  との間の距離から計算される閾値  $t(d(X, Y))$  以上ならば、 $X$  の地点から  $Y$  の地点へ移動するために十分な時間があると判断し、特徴量を用いた同一人物判定を行う。しかし、そうでない場合は、 $X$  の地点から  $Y$  の地点へ移動するために十分な時間がない、つまり、同一人物ではないと判断し、特徴量を用いた同一人物判定を行わない。メタデータを用いての類似性判定をアルゴリズム 2 に示す。3行目で上記2つの条件を満たす場合のみ、6行目で特徴量を用いた同一人物判定を行っている。

## 5. 計算機実験

本論文の提案手法である映像追跡を用いた再特定手法と前記手法を類似度木により高速化したもの、ナイーブな手法とを実装し、速度を比較した。ナイーブな手法では、特徴量集合を作成せず、すべての特徴量の組合せについて類似度を計算する。各プログラムのソースコードを C++ 言語で書き、g++5.4 でコンパイルした。また、実験は Intel Xeon 3.0 GHz の CPU と 8GB のメモリを搭載し、Ubuntu 16.04 を OS とする計算機上で行った。入力となるデータは、カメラ 30 台で撮影された延べ 140 分の映像データを解析して得られた特徴量データである。今回の実験では、NeoFace<sup>®</sup> を用いて、人の顔の特徴量を抽出した。また、この映像には、400 人の人が歩いている映像が収録されている。

実験の結果を図 4 に示す。提案手法は、ナイーブな手法よりも 30% 程度高速化されている。本実験で使用したデータでは、

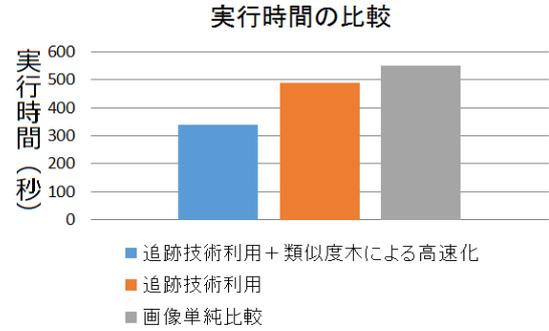


図 4 実行時間の比較。類似度木を使用することで 30% 程度高速化されている。

特徴量集合における特徴数が平均して 14 程度と比較的少数であったが、類似度木により 30% 程度高速化されている。

## 6. おわりに

本論文では、複数のカメラ映像から、同一人物の追跡を行う方法を提案した。提案手法は、映像追跡技術によって抽出された同一人物の画像群より特徴量を抽出し、再特定技術によって画像群同士で同一人物の発見を行う。この際に、類似検索を高速化するためのためのデータ構造である類似度木を用いることで再特定を高速に行うことができる。実験的にも再特定を高速化できることを示した。今回の実験に使用した入力データ中での特徴量集合における特徴数が比較的少数であったため、高速化が 30% 程度であったが、より大きな特徴量集合を用いることでさらなる高速化が期待できる。また、今回の実験では、再特定の精度や再現率の定量的な測定は行わなかったが、図 1 に掲げる 4 名分の画像は全員同一人物であると特定することができた。そのため、本論文での提案手法は、単純な方法よりもより正確に再特定が行えることが期待できる。

## 文 献

- [1] Loris Bazzani, Marco Cristani, Alessandro Perina, Michela Farenzena, and Vittorio Murino. Multiple-shot person re-identification by HPE signature. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pp. 1413–1416. IEEE, 2010.
- [2] Marcelo Bertalmio, Guillermo Sapiro, and Gregory Randall. Morphing active contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 7, pp. 733–737, 2000.
- [3] Andrew Blake and Michael Isard. The condensation algorithm-conditional density propagation and applications to visual tracking. In *Advances in Neural Information Processing Systems*, pp. 361–367, 1997.
- [4] Dorin Comaniciu, Visvanathan Ramesh, and Peter Meer. Kernel-based object tracking. *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 25, No. 5, pp. 564–577, 2003.
- [5] Abir Das, Anirban Chakraborty, and Amit K Roy-Chowdhury. Consistent re-identification in a camera network. In *European Conference on Computer Vision*, pp. 330–345. Springer, 2014.
- [6] Hehe Fan, Liang Zheng, and Yi Yang. Unsupervised person re-identification: Clustering and fine-tuning. *arXiv preprint arXiv:1705.10444*, 2017.
- [7] Michela Farenzena, Loris Bazzani, Alessandro Perina, Vit-

- torio Murino, and Marco Cristani. Person re-identification by symmetry-driven accumulation of local features. In *Computer Vision and Pattern Recognition, 2010 IEEE Conference on*, pp. 2360–2367, 2010.
- [8] John Gantz and David Reinsel. The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east. *IDC iView: IDC Analyze the future*, Vol. 2007, No. 2012, pp. 1–16, 2012.
- [9] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pp. 1097–1105, 2012.
- [10] Ryan Layne, Timothy M Hospedales, Shaogang Gong, and Q Mary. Person re-identification by attributes. In *Bmvc*, p. 8, 2012.
- [11] Wei Li, Rui Zhao, Tong Xiao, and Xiaogang Wang. Deep-reid: Deep filter pairing neural network for person re-identification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 152–159, 2014.
- [12] Xiao Liu, Mingli Song, Qi Zhao, Dacheng Tao, Chun Chen, and Jiajun Bu. Attribute-restricted latent topic model for person re-identification. *Pattern recognition*, Vol. 45, No. 12, pp. 4204–4213, 2012.
- [13] Bingpeng Ma, Yu Su, and Frédéric Jurie. Bicov: a novel image representation for person re-identification and face verification. In *British Machine Vision Conference*, pp. 11–pages, 2012.
- [14] Bingpeng Ma, Yu Su, and Frédéric Jurie. Local descriptors encoded by fisher vectors for person re-identification. In *European Conference on Computer Vision*, pp. 413–422. Springer, 2012.
- [15] Niall McLaughlin, Jesus Martinez del Rincon, and Paul Miller. Recurrent convolutional network for video-based person re-identification. In *Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference on*, pp. 1325–1334. IEEE, 2016.
- [16] Filip Radenović, Giorgos Tolias, and Ondřej Chum. CNN image retrieval learns from bow: Unsupervised fine-tuning with hard examples. In *European Conference on Computer Vision*, pp. 3–20. Springer, 2016.
- [17] Vali Salari and Ishwar K. Sethi. Feature point correspondence in the presence of occlusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 1, pp. 87–91, 1990.
- [18] Koichi Sato and Jake K Aggarwal. Temporal spatio-velocity transform and its application to tracking and interaction. *Computer Vision and Image Understanding*, Vol. 96, No. 2, pp. 100–128, 2004.
- [19] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 815–823, 2015.
- [20] Jianbo Shi, et al. Good features to track. In *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on*, pp. 593–600. IEEE, 1994.
- [21] Zhiyuan Shi, Timothy M Hospedales, and Tao Xiang. Transferring a semantic representation for person re-identification and search. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4184–4193, 2015.
- [22] Roy L Streit and Tod E Luginbuhl. Maximum likelihood method for probabilistic multihypothesis tracking. In *SPIE's International Symposium on Optical Engineering and Photonics in Aerospace Sensing*, pp. 394–405. International Society for Optics and Photonics, 1994.
- [23] Chi Su, Fan Yang, Shiliang Zhang, Qi Tian, Larry S Davis, and Wen Gao. Multi-task learning with low rank attribute embedding for person re-identification. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3739–3747, 2015.
- [24] Hai Tao, Harpreet S Sawhney, and Rakesh Kumar. Object tracking with bayesian estimation of dynamic layer representations. *IEEE transactions on pattern analysis and machine intelligence*, Vol. 24, No. 1, pp. 75–89, 2002.
- [25] Cor J Veenman, Marcel JT Reinders, and Eric Backer. Resolving motion correspondence for densely moving points. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 1, pp. 54–72, 2001.
- [26] Yichao Yan, Bingbing Ni, Zhichao Song, Chao Ma, Yan Yan, and Xiaokang Yang. Person re-identification via recurrent feature aggregation. In *European Conference on Computer Vision*, pp. 701–716. Springer, 2016.
- [27] Dong Yi, Zhen Lei, Shengcai Liao, and Stan Z Li. Deep metric learning for person re-identification. In *Pattern Recognition, 2014 22nd International Conference on*, pp. 34–39. IEEE, 2014.
- [28] Rui Zhao, Wanli Ouyang, and Xiaogang Wang. Learning mid-level filters for person re-identification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 144–151, 2014.
- [29] Liang Zheng, Zhi Bie, Yifan Sun, Jingdong Wang, Chi Su, Shengjin Wang, and Qi Tian. Mars: A video benchmark for large-scale person re-identification. In *European Conference on Computer Vision*, pp. 868–884. Springer, 2016.
- [30] 劉健全, 西村祥治, 荒木拓也. 類似度の階層関係を用いた類似検索の高速化とその応用. 電子情報通信学会技術研究報告 = IEICE technical report : 信学技報, Vol. 115, No. 316, pp. 123–128, nov 2015.