

二視点での人体追跡方式の検討

竹内 格 石倉 禅 高田 慎也

NTT セキュアプラットフォーム研究所

1. はじめに

監視映像を用いた防犯や行動監視のセキュリティ対策が広く使われている。また近年はクラウドを用いた監視ソリューションにより、導入コストや運用コストが抑えられるようになってきている。

現状の監視映像の用途は、撮影した監視映像を「証拠」として蓄積しておき、何か問題が起きた際にさかのぼって映像を目視で追跡することが一般的である。

蓄積された監視映像から「誰が何をしていたのか？」を自動抽出しログデータ化することができれば、セキュリティ対策の高度化が期待できる。例えば特定の人物が機微情報のあるエリアに接近する映像シーンを自動検索することができれば、機微情報管理における監視稼働の低減を期待できる。

上記の目標に向け、監視映像から人体追跡を行なうことで「いつ(WHEN)・どこで(WHERE)・誰が(WHO)」(ただし個人の識別はできない)をログデータ化する方式の研究に取り組んでいる。

人体追跡には市中製品を適用する構成だが、本報告では特にログデータ化に大きな影響を及ぼす人体追跡の誤検知に対して、二台のカメラ映像の人体追跡結果を組み合わせる構成とすることにより人体追跡の誤検知を回避する方式について示す。

2. 人体追跡

人体追跡とはカメラ映像のフレーム画像群を入力とし、フレームの日時・人体ID・そのフレーム画像上の座標(X座標・Y座標)を情報単位とする画像ログのデータを出力する機能である。人体IDは連続するフレーム画像で同じ人体には同じ値が付与される。

このような人体追跡には [1] などがある。

3. 人体追跡の誤検知

人体追跡は誤検知が生じることがある。誤検知が発生した際の人体IDとその座標位置の軌跡の例を図1に示す。

同じ人物が右から左奥へ歩いていくシーンの映像を、人体追跡処理したところ、検出された人体IDはH1→H2→H3と不連続に変化してしまった。本来ならば、人体IDはH1のままであり軌跡は連続的なものになるべきである。我々の環境では数名の人物で撮影した単純なテスト映像のうち、約30%の映像で同様の誤検知が確認された。

このような誤検知があると、人体追跡結果を使った各種の適用が期待通りに行えなくなってしまうため、誤検知を低減することを考える必要がある。

人体追跡の誤検知を、次の5パターンに大別した。

(誤検知1: 人体IDの消失) カメラにはその人物が映

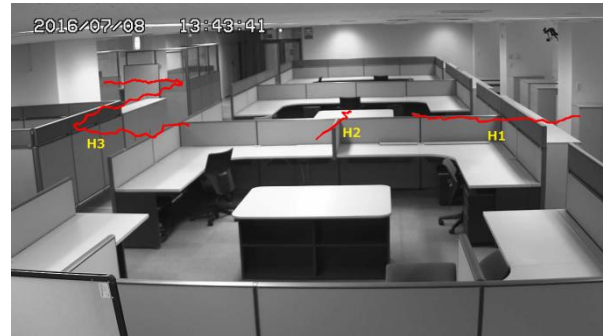


図 1

っているのにその人体が検出されなくなる
(誤検知2: 人体IDの変化) 同じ人物なのに別の人体として検出される
(誤検知3: 人体IDの発生) 新しい人体として検出される
(誤検知4) 人物のいない箇所で人体が検出される
(誤検知5) 非常に短い期間でのみ人体が検出される

誤検知4・5 は比較的簡単な方法でフィルタできることができる。例えば4 に対しては、人体追跡の入力となる監視映像のフレーム画像に対して、無人の背景画像との差分をとることで誤検知の回避が期待できる。5 に対しては、人体追跡の出力データから、極端に出現時間の短かった人体IDを除去する方法が考えられる。ここでは 1・2・3への対策を中心に考えることとする。

4. 誤検知回避の方式

多視点画像を使って人の動作を追跡する手法については多数の取り組みが存在している[2][3]。本方式では新たに画像処理や人体モデルを再構成する手法にフォーカスするのではなく、既存の人体追跡手法を部品として組み合わせ、そこで算出された人体追跡結果を評価する方式について考える方針とする。

今回はカメラ1画像の人体追跡結果と、カメラ2画像の人体追跡結果を組み合わせることで双方のカメラで生じうる誤検知を相互に回避する方式を示す。

カメラ1で同じ人物の人体IDがH1→H2→H3のように変化している一方でカメラ2では誤検知がなく、同じ人物の人体IDはH4として検出されたとする。

カメラのフレーム画像の一つ一つを追っていき、

$H1=H4 \Rightarrow H2=H4 \Rightarrow H3=H4$

の対応関係の変化を見いだせば、 $H1=H2=H3$ の関係にあることが推測可能である。

双方のカメラの画像から人体追跡処理によって得られた座標の組み合わせが対応関係であれば、意味のある組み合わせとみなすこととする。

双方のカメラ間の座標の対応関係については、その座標の対応を使用する。事前作業としてカメラ1カメラ2の撮影領域内を一つの人物が動き回っているシーンを撮影し、その際の画像データで人体追跡を行い、物理的に同じ位置となる、双方のカメラ画像上の座標の対応を記録しておく。以下この対応表を「座標対応表」と呼ぶ。

5. アルゴリズム概要

5.1. 全体フロー

カメラ1カメラ2それぞれにおいて、以下のステップ(1)~(4)で画像ログを生成し、(5)でこれらのマッチングを行う。

- (1) カメラ撮影：連続的に画像データを出力。時刻 T_i におけるカメラ1の画像データを画像データ $1i$ 、カメラ2は画像データ $2i$ と呼ぶ。
- (2) 背景差分処理：カメラからの画像データと事前に用意した背景画像データとの差分となる画像データを出力する。(誤検知4の対策)
- (3) 人体追跡処理：画像データから画像番号・人体ID・ピクセル座標からなる画像ログを出力する。
- (4) フィルタ処理：画像ログから、ある人体IDが継続して存在していた時間を保持しおき、閾値未満の画像ログを除外する。(誤検知5の対策)
- (5) マッチング処理：事前に作成した座標対応表を使用し、二台のカメラの画像ログの中から、同じ人体となる人体IDを抽出し対応づけ、共通人体IDを付与する。時刻・共通人体ID・ピクセル座標からなるマッチングログを出力する。

5.2. マッチング処理のフロー

マッチング処理のフローは以下の通りである。

- (1) 前回(時刻 T_{i-1})のマッチングログを取得する。前回のない初回は空のマッチングログとする。今回(時刻 T_i)の出力する空のマッチングログを用意。
- (2) 二つの画像ログから人体IDと座標の組み合わせのリストを作成

リストの例を示す。画像ログ $1i$ と画像ログ $2i$ が以下のとき、

画像ログ $1i = T_i, H11:X11$

画像ログ $2i = T_i, H21:X21$

組み合わせのリストは次のようになる

$\langle H11:X11, H21:X21 \rangle, \quad \langle H11:X11, H21:NULL \rangle,$

$\langle H11:NULL, H21:X21 \rangle$

ここで「NULL」は撮影領域に人物が映っていない状態かあるいは人体追跡の誤検知で検出されていないかのどちらかを示す。上の例は人物が一名もしくは二名の単純な場合だが、それ以上の複数人が映っている場合も同様にすべての組み合わせについてリストを作成する。

- (3) 対応座標データと近傍点のないものをリストから除外
- (4) 前回のマッチングログと比較し類似度の高いものを選択し、前回のマッチングログの共通人体ID

を引き継ぐ形で今回のマッチングログに追加。また同じ人体IDを持つ組み合わせはすべてリストから除外。

- (5) リストに残った組み合わせは、類似度は低いだが、対応座標との近傍点は存在するので、新たに人物が出現したとみなし、新しい共通人体IDを付与し、マッチングログに追加し、リストから除外
- (6) リストが空になったら、マッチングログを出力

5.3. 類似度の算出

類似度の算出方法の概要を記す。

ここでは、人体IDの一致性と座標の近傍性の二つについてスコアを計算し合計値を類似度とした。

- (1) 人体IDの一致性：前回のマッチングログと人体IDが一致するものが多いほどスコアを高くする。
例：双方一致、片方一致、一致なしでスコアを分けるなど
- (2) 座標の近傍性：前回のマッチングログの座標との距離が近傍のものほどスコアを高くする。
例：双方とも近傍、片方近傍、近傍なし、におうじてスコアを分けるなど

5.4. 誤検知への対処

誤検知1~3のそれぞれのケースは類似度により同じ共通人体IDが付与されることで回避される。

前回のマッチングログが $\langle H11:\langle H11:X11, H21:X21 \rangle$ のときに、それぞれで出現する組み合わせは以下のようになる：

(誤検知1：人体IDの消失) $\langle H11:X11', H21:NULL \rangle$

(誤検知2：人体IDの変化) $\langle H11:X11', H21':X21' \rangle$ (ただし $H21 \neq H21'$ のとき)

(誤検知3：人体IDの発生) $\langle H11:X11', H21':X21' \rangle$ (ただし $X21=NULL$ のとき)

いずれも適切な類似度が算出され、つまり、前回のマッチングログの人体IDとの一致性或座標の近傍性が確認できるならば、前回のマッチングログの共通人体IDを引き継ぐ組み合わせと判定される。このようにして誤検知が回避される。

6. おわりに

監視映像から人体追跡などを行なうことで人物の行動をログデータ化する方式について取り組んでいる。本報告では二台のカメラ映像の人体追跡結果を組み合わせる構成とすることで、人体追跡の誤検知を回避する方式とそのアルゴリズムについて示した。今後はアルゴリズムを実装評価していく予定である。

7. 参考文献

- [1] OMRON Corporation, "OKAO Vision," <http://plus-sensing.omron.com/technology/>
- [2] 佐藤・川田・大崎・山本, 「多視点動画像からの陣元動作の追跡と再構成」, 電子情報通信学会論文誌, 1997
- [3] 森・内海・大谷・谷内田・中津, 「非同期多視点画像による人物追跡システムの構築」, 電子情報通信学会論文誌, 2001