

進行方向制約を用いた大域的最適な複数対象追跡

岡田亜沙美[†] 浮田 宗伯^{†a)} 萩田 紀博[†]

Globally-Optimal Multi-Object Tracking with Constraints on Moving Directions

Asami OKADA[†], Norimichi UKITA^{†a)}, and Norihiro HAGITA[†]

あらまし 大域的最適解を得る複数対象追跡では、各検出結果をノードに対応付けたグラフにおける最小コスト経路探索が行われる。従来法では、隣接 2 フレーム間でそれぞれの静的特徴量（位置・サイズなど）を抽出し、それらの差分から計算できるコストのみが探索に利用できた。提案法では、進行方向のような動的特徴量の差分をコストに利用するため、各ノードが追跡経路に応じて複数の動的特徴量をもつ場合、そのノードを仮想的に分割する。各分割ノードが異なる進行方向を經由した追跡経路に対応つき、生成されたグラフに対して従来法とおりの解探索法によって大域的最適解を得ることができる。実験では、Person Re-ID2011 dataset を使い、従来法と比較して、提案法により追跡成功率が向上していることを定量的に確認する。また、提案法の効果の典型的な例を画像で示し、目標としていたすれ違い時の追跡成功を実現できていることを確認する。

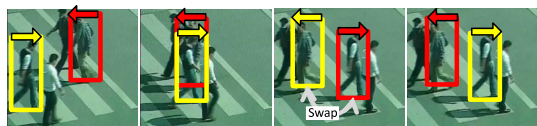
キーワード 複数対象追跡, 大域的最適解, 進行方向制約

1. ま え が き

監視カメラでの追跡技術は、犯人追跡や車や人などの流量解析に用いられる重要な技術である。Mean shift tracker [1] や Particle tracking [2] に代表されるオンライン追跡では、過去フレームの追跡結果から次フレームにおける対象位置を予測して追跡性能を向上させる。これに対し、Data association を応用した大域的最適追跡に代表されるオフライン追跡では、全フレーム中の全ての人領域候補から同一対象の領域を時間方向に繋げ合わせ、全追跡結果のコスト総和が最適となる追跡結果を得る。オフライン追跡の計算量の高高速化は進み、1分程度の動画であれば1秒足らずで結果が得られる [3]。よって、追跡対象に追従する能動カメラ制御のような実時間・オンライン追跡を除けば、犯罪捜査や交通量解析などに十分応用可能である。本論文では、このようなオフライン追跡を研究対象とする。

しかし、グラフィカルモデルにおいて効率的に大域

的最適解を得る従来法では、図 1 (左) のように人がすれ違った際、追跡結果が交換される誤りを得やすい問題があった (図 1 (Previous))。検出結果を時系列に繋げるとき、隣接 2 フレーム間の位置情報や領域のサイズといった静的特徴量のみしか利用できないことに起因する。これは、効率的に解を得るため、対象となるグラフにはマルコフ性を仮定した「前後 1 フレームのノード間のみ連結可能」という制約があるためである。一方、「移動対象の進行方向は急激に変化しない」という進行方向制約は、最低でも 2 フレームの撮影像が無ければ各フレームの進行方向が計算できないため、



Passing each others at $t-1$ (Previous) Swap at t (Proposed) Success at t

図 1 提案法の効果。位置などの静的情報しかコストにできない従来の大域的最適追跡では、すれ違い時の誤追跡が簡単に発生する (図中央の“Previous”)。提案法では、進行方向制約を導入し、主にすれ違い時の追跡頑健性を向上させる (“Proposed”)。

Fig.1 While previous methods might fail tracking people passing each others, they can be tracked by constraints on moving directions by the proposed method.

[†] 奈良先端科学技術大学院大学, 生駒市
Nara Institute of Science and Technology, Ikoma-shi, 630-0192 Japan

a) E-mail: ukita@is.naist.jp

DOI:10.14923/transinfj.2014JDP7091

静的特徴量からだけでは計算できない。よって、オンライン追跡では一般的に利用される進行方向制約は、マルコフ性を仮定したグラフでの大域的最適追跡には利用できなかった。この問題の解決のため、本研究では、グラフ構造を拡張し、大域的最適な複数対象追跡において進行方向制約を用い、図 1 (Proposed) に示すようにすれ違い時の追跡失敗を回避する。

2. 関連研究

大域的最適追跡の初期の研究 [4]～[6] では、追跡対象数や各対象の追跡開始・終了フレームを既知とすることにより、現実的な計算時間で解を得ていた。しかし、実際の問題において追跡対象数などを既知と仮定することは現実的ではなく、これらのパラメータも未知とする手法も研究されてきた。これらの研究 [7]～[12] はいずれも、大域的最適解のために線型計画法を用いた手法であり、 N を動画のフレーム数としたとき、計算量は $O(N^3 \log^2 N)$ である。これはフレーム数に大きく依存するため、現実的な時間で解を得るために比較的短いシーケンスに対してしか適用できなかった。

これに対し、文献 [3] では計算量を $O(KN \log N)$ (K は追跡対象数) に削減した。しかし、こうしたグラフィカルモデルにおける大域的最適追跡では、動的計画法などによる高速解探索のために隣接 2 フレームの検出領域 (グラフ中のノード) 間しかエッジで結ぶことができない。このため、エッジに与える追跡コストには、隣接フレーム間の対象の位置・領域サイズなどの静的な情報しか利用できない。すなわち、オンライン追跡では一般的に行われている進行方向による対象領域対応付けの頑健化が、大域的最適追跡では実現できていない。本研究では、グラフにおけるエッジ連結を隣接フレーム間に制限したまま進行方向制約をグラフに導入することによって、全フレームにおける大域的最適な複数対象追跡を効率的に実現する。

大域的最適な追跡において進行方向制約を利用した従来研究には、手法 [7], [13] がある。手法 [7] では、複数フレーム間の追跡結果をノードとすることによってノード数を大幅に削減しているため、その追跡結果中の誤り (局所追跡誤差) には対処できない。手法 [13] では、隣接フレーム間の任意の対象検出結果を 1 ノードに変換して、各ノードが隣接フレーム間の進行方向を情報としてもつグラフを得ている。このグラフにおける解探索は複雑な制約条件付きの問題を解くことになり、得られる解は近似最適解である。また、計算コ

ストも手法 [3] の約 120 倍^(注1)もかかってしまう。

3. 人検出

人検出のため、まず画像中の各領域から HOG 特徴量 [15] を抽出する。この HOG 特徴量を Support Vector Machine (SVM) [16] によって人、またはその他の領域に識別する。

SVM で学習する正例は、汎用的な INRIA Person Dataset [17] と各カメラの撮影環境に特化した情報とともに学習する。環境特化の学習データには、監視カメラ視野内では、撮影方位も固定されているという仮定の下、各カメラ視野内で撮影された画像から人の矩形領域を抽出して与えた。INRIA Person Dataset [17] の人画像は、人の例外的な動きや多様な服装などに対応するための汎化性向上のために与えた。

一般的な人検出において、負例には多様な背景画像が与えられる。これに対し、監視カメラは設置されたあとに位置や角度の変化が生じにくいいため、本問題においては負例には各カメラの背景画像を与える。また、複数人が接近して移動するときに、人の間にできる HOG 特徴量が人らしい特徴を示すことが多々あるので、人と人之間を囲んだ矩形領域も負例として与える。

4. 大域的最適な人追跡と問題点

図 2 に大域的最適追跡のためのグラフを示す。赤線

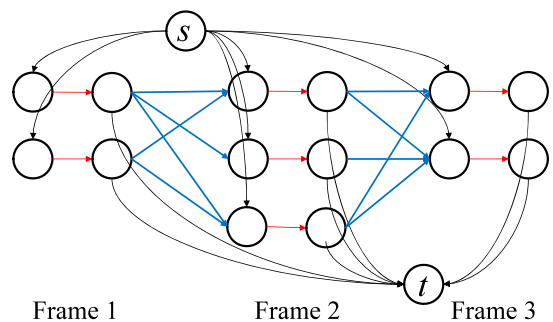


図 2 大域的最適追跡のためのグラフ (文献 [3] 参照)。赤線で結ばれたノードペアが各検出結果に対応し、赤線と青線のエッジは、それぞれ各検出結果を人とみなすコストとフレーム間で領域が同一対象として対応付けたためのコストをもつ。

Fig. 2 Graph for globally-optimal tracking [3]. Each node corresponds to a detected human region. Red and blue links have respectively costs of each detected region and similarity between detected regions.

(注1)：論文 [3], [13] 中の ETHMS セット [14] による結果より算出。

のエッジで結ばれた二つのノードは、同一フレームの一つの検出結果に対応しており、このフレーム内エッジにはその検出結果が人らしいほど小さいコストが与えられる。この二つのノードは、フレーム内エッジに人らしさのコストを与えるために仮想的に分割されたものであり、ともに等しい属性（例：対象の検出座標など）をもつ。隣接フレーム間のノードを結ぶ青線のエッジには、検出位置やサイズの類似度に基づいて、同一対象らしいほど小さいコストが与えられる。sノードとtノードは人追跡の始点と終点であり、これらのノードと任意フレームのノードをつなぐパスの最小コスト経路探索によって、任意のフレームでの追跡開始及び終了が可能になる。

本研究では、人らしさをもつフレーム内エッジに、 $C_H - S_H$ をコストとして与えた。ただし、 S_H は SVM による出力（人らしいほど正の大きな値をもつ）、 C_H は定数で、本論文の実験では文献 [3] の実装に倣い 0.2 とした。フレーム間エッジに関しては、以下 3 種のコストを計算し、全てのコストがしきい値以下のノード間のみエッジで結び、全てのフレーム間エッジにはコストとして定数を与えた [3]^(注2)。

- (1) フレーム間の矩形の中心座標の距離
- (2) フレーム間の矩形の面積の差の絶対値
- (3) フレーム間の色情報の差（矩形の中心 25% の面積の領域から抽出される RGB ヒストグラム間の Bhattacharyya 係数の負値）

手法 [3] の概要を以下に示す。

Step1: 全フレームにおける全ての人検出領域を、前述したコストのしきい値処理に基づいて隣接フレーム間で連結し、図 2 に示す有向非循環グラフを生成する。
Step2: この有向非循環グラフにおいて、反復的に最小コスト経路探索を適用し各追跡結果を抽出することで、複数人追跡を実現する。反復時、追跡結果に含まれたエッジは連結方向を反転させ、コストを負にして（初期に定数コストの負値）与える。このグラフに対して次の追跡結果を得るために最小コスト経路探索を行うと、ある追跡済みの経路（経路 A と呼ぶ）に対応する反転エッジが探索結果（経路 B と呼ぶ）に含まれる可能性がある。このとき、この反転エッジを「どの追跡結果にも含まれない部分パス」として、経路 A と

経路 B を修正する。この処理により、逐次的な最適解探索によって上位複数個の最適解を得ている（詳細は文献 [3] の 5 節、及び [18] を参照されたい）。

手法 [3] を含む大域的最適追跡のためのグラフ構成は、効率的な解探索のために隣接フレーム間にしかエッジをもたない。よって、より高次の連結からしか得られない情報は、エッジのコスト計算（手法 [3] ではエッジの有無の判定）に利用できない。例えば、フレーム i とフレーム $i+1$ の間で進行方向の差分を計算しようとする、フレーム i における進行方向はフレーム $i-1$ における検出座標も必要となる。よって、フレーム i の検出領域がフレーム $i-1$ のどの検出領域と連結するかによって進行方向は変わってしまうため、フレーム i の検出領域固有の情報として進行方向を与えることができず、その結果、進行方向をフレーム間エッジのコストに利用できない。図 5 の例では、frame2 の $P_{2,1}$ ノードが frame3 の $P_{3,1}$ と $P_{3,2}$ の二つのノードと連結するが、 $P_{2,1}$ のノード一つに異なる二つの進行方向を保持することはできない。 $P_{t,i}$ は人領域の検出座標であり、 t はフレーム番号、 i はフレーム中のノード番号である。

このように進行方向を参照できないと、すれ違い時に追跡失敗しやすい。図 3 に例を示す。図 3 (上) の例は、すれ違うタイミングでピンク枠が消失したのが追跡失敗の主な原因である。このケースの解決には、隣接フレームだけでなく消失フレームをまたいだエッジ接続か、後処理で複数の追跡結果を繋げ合わせる必

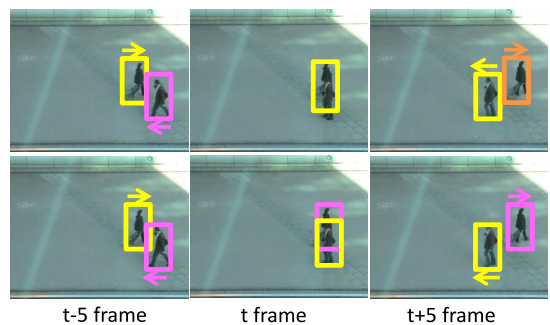


図 3 手法 [3] の失敗例。(上) すれ違うタイミングでピンク枠が消失した上、元のピンク枠の人領域に黄色枠が移ったため、2 人も追跡失敗する。(下) すれ違うタイミングで黄枠はピンク枠に、ピンク枠が黄枠に移り、2 人も追跡失敗している。

Fig. 3 Unsuccessful tracking by [3]. Upper: A region depicted by a pink rectangle was disappeared. Lower: Two tracking results were swapped when passing each other.

(注2)：こうしたコストのしきい値処理ではなく、コストをそのままエッジに与えることもできる。これにより、しきい値処理により発生しうるエッジの不足は回避できる。しかし、提案法では、手法 [3] に倣い、高速処理のため、エッジ数を減らして最小コスト経路探索を効率化させた。

要があり、提案法による解決の対象外となる。(下)の例では、人検出は成功しているが、すれ違うタイミングで追跡結果が交換されてしまっている。提案法では、このような追跡失敗例の解決を目標とする。

5. 人追跡の進行方向制約を用いたグラフ

5.1 進行方向制約を表すコスト

進行方向から計算されるコストには、進行方向の差分の絶対値を与える。

$$\|\overrightarrow{P_{t+1,j}P_{t+2,i}} - \overrightarrow{P_{t,k}P_{t+1,j}}\| \quad (1)$$

t はエッジの始点フレーム番号, i, j, k はフレーム中のノード番号であり, $\overrightarrow{P_{t,j}P_{t+1,i}}$ は t フレームの j ノードから $t+1$ フレームの i ノードへ進んだ場合の進行方向を表す。図 4 に、各ノードの進行方向と、ノード間を結ぶエッジに与えられるコストの例を示す。

5.2 進行方向制約を加えた追跡法

提案法の処理概要は以下のとおりである。

(1) 進行方向以外の特徴量を用いて、従来法 [3] より図 2 の有向非循環グラフを作成する。すなわち、4. で挙げた位置、サイズ、色の差のしきい値処理で、ノード間をエッジで結び、全エッジには定数コストを与える。

(2) 生成されたグラフに対して、進行方向コストに基づいたノード分割を適用してノードとエッジを追加する。生成されたグラフの全エッジにおいて、進行方向コストもしきい値処理で低コストのエッジのみ残し、最終的に全エッジに定数コストを与える。

(3) このグラフから、ノード分割を考慮した最小コスト経路探索により追跡結果を得る。

提案するノード分割によるコスト保持は、「隣接フ

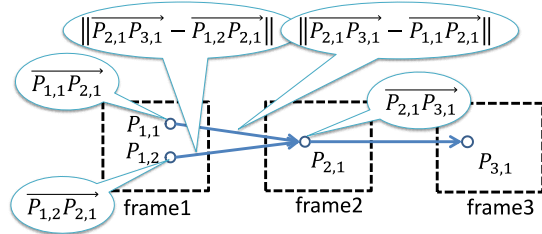


図 4 進行方向制約のためのコスト。ノードを指す吹き出しには、そのノードの進行方向を、エッジを指す吹き出しにはそのエッジの進行方向コストを記す。

Fig. 4 Cost for a constraint on moving directions. Balloons with a node and an edge show a moving direction of that node and the cost given to that edge, respectively.

レームの検出結果を連結して生成される」という基本構造が同じグラフに基づいている範囲内で、その他の追跡手法にもそのまま適用可能である。

5.2.1 進行方向制約を加えたグラフの作成方法

従来法 [3] によるグラフ作成後の、各ノードの進行方向計算、進行方向に基づくノード分割、進行方向差分によるコスト計算、の詳細な手順を以下に示す。

(1) 進行方向の求め方を図 5 に示す。注目フレームを $f1$ 、その一つ先のフレームを $f2$ とする。まず、図 5 (上) のように $frame1$ と $frame2$ に $f1$ と $f2$ を当てはめ、進行方向の計算をする。計算された進行方向は $f1$ のノードに保持する。以降、図 5 (下) の例で $frame2$ と $frame3$ に $f1$ と $f2$ を当てはめているように 1 フレームずらしつつ、進行方向を計算をする。

(2) $f1=frame2$, $f2=frame3$ の進行方向計算時、 $frame2$ の $P_{2,1}$ のノードは、 $frame3$ の $P_{3,1}$ と $P_{3,2}$ の二つのノードにエッジが連結しているため、 $P_{2,1}$ のノードには二つの進行方向を保持しなければならない。そこで、図 5 (下) のように $P_{2,1}$ のノードを分割する。以降、検出座標 $P_{t,i}$ のノードを $N_{t,i}$ と表記し、この分割ノードを $N_{t,i}^d$ とする (d は分割番号)。分割されたノードは、進行方向以外は全て元ノードと等しい属性 (例: 検出座標など) をもつ。 $f1$ の各分割ノードと $f2$ のノードの間では、分割ノードが保持する進行方向のノードとの間のみエッジをもつ (図中、 $N_{2,1}^1$ は $N_{3,1}$ と、 $N_{2,1}^2$ は $N_{3,2}$ とのみ)。 $f1$ の過去のフレームとは、分割ノードのいずれもが、分割元のノードが連結していた全てのノードとエッジをもつ ($N_{2,1}^1, N_{2,1}^2$ とともに、 $N_{1,1}, N_{1,2}$ の両方とエッジをもつ)。

(3) 進行方向のコストは、連続 3 フレームで計算する。図 6 の例では、 $frame1$, $frame2$, $frame3$ が対象である。進行方向のコストは $frame1$ と $frame2$ 間のエッジに保持する。図 6 の例において、 $frame1$ から $frame3$ を通る経路は 4 通りある ($N_{1,1} \rightarrow N_{2,1}^1 \rightarrow N_{3,1}$, $N_{1,1} \rightarrow N_{2,1}^2 \rightarrow N_{3,2}$, $N_{1,2} \rightarrow N_{2,1}^1 \rightarrow N_{3,1}$, $N_{1,2} \rightarrow N_{2,1}^2 \rightarrow N_{3,2}$)。 $frame1$ と $frame2$ 間の 4 本のエッジに対して、図 6 に示すようにそれぞれに対応する進行方向コスト (式 (1)) を与える。この処理を 1 フレームずつ進めつつ適用し、全てのエッジに対して、進行方向のコストを与える。

(4) 上記処理を全てのフレームに適用後、グラフの進行方向のコストに対してしきい値を設定し、進行方向コストの小さいノードのエッジを残すように、グラフ全体のエッジを削除する。

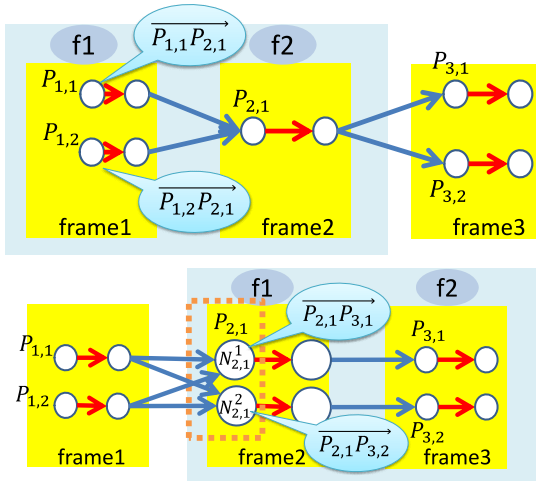


図5 進行方向保持のためのノード分割. 上図が frame1 と frame2 間を対象としたとき (ノード分割不要の例), 下図を frame2 と frame3 間を対象 (ノード分割必要の例) とした進行方向計算. 簡略化のため, 赤線のフレーム内エッジ左側の分割ノード以外は, ノード番号 $N_{t,i}^d$ を省略する.

Fig. 5 Node division for maintaining moving directions. Upper: no division is required between frame1 and frame2. Lower: division is required between frame2 and frame3.

5.2.2 進行方向制約を加えたグラフによる人追跡結果の取得

生成したグラフから, 人追跡の結果を得る. 従来法 [3] では, 最小コスト経路探索によりコストが小さい追跡結果から反復的に取得していく. この反復時に, 追跡結果に含まれたノードにつながっているエッジは, 続く最小コスト経路探索で繰り返し選択されにくいようにコスト修正される.

提案法でも, 従来法 [3] と同様の処理で追跡結果を得るが, ノード分割によって仮想的に取得可能な経路は増えている. 図5と図6の例では, frame1とframe2間で, 分割前は $P_{1,2} \rightarrow P_{2,1}$ と $P_{1,1} \rightarrow P_{2,1}$ の2経路しかないが, 分割後は $P_{1,1} \rightarrow P_{2,1}^1$, $P_{1,1} \rightarrow P_{2,1}^2$, $P_{1,2} \rightarrow P_{2,1}^1$, $P_{1,2} \rightarrow P_{2,1}^2$ の4経路になる. しかし, 同一検出領域に対応するノードが分割されているだけなので, これらの分割ノードが異なる追跡結果にそれぞれ含まれることは許されない. そこで, 分割ノードが追跡結果に含まれた場合, この分割ノードにつながるエッジだけでなく, 同一検出結果に対応する他の分割ノードにつながるエッジにもコスト修正を適用する. その後の, 反復探索は従来法のとおり行う.

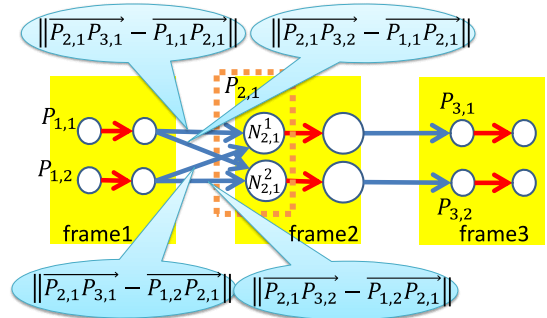


図6 進行方向のコスト計算. 連続3フレームを対象として, 第1と第2フレームの間の進行方向と, 第2と第3フレームの間の進行方向との差分をコストとして計算し (式 (1)) エッジに与える.

Fig. 6 Costs of moving directions given to edges. Costs between frame1 and frame3 are computed by Eq. (1) and given to edges between frame1 and frame2.

最後に, 提案法における計算量に関して述べる. 手法 [3] の計算量は $O(KN \log N)$ である (K は追跡対象数, N はフレーム数). よって, 提案法におけるノード分割によって各フレームにおける検出領域数が仮想的に増えても, 計算量には影響を与えない.

6. 実験

6.1 Person Re-ID2011 dataset

実験に用いるデータセットは, Person Re-ID2011 dataset [17] に含まれる二つのカメラの動画である. データの詳細を以下に示す.

フレームレート: 25 フレーム/秒

フレーム数: Camera a: 92825 フレーム, Camera b: 99997 フレーム

追跡人数: Camera a: 385 人, Camera b: 749 人

これらの二つの動画は, 図3などに示すよう, 上方から見下ろすような固定カメラで撮影された. 全フレームにおいて, 全ての人領域を矩形領域としてアンテーションし, その x, y 座標と x, y サイズを既知とした. また, フレーム間で同一人の領域には同一の ID が割り当て, 追跡結果も既知とした.

6.2 実験結果

実験は二つの動画に対して独立に行い, それぞれを9分割した後に, 八つを人領域検出及び4.に示した三つのコスト (矩形位置の差, 矩形サイズの差, 色情報の差) と進行方向制約によるノード間のエッジ連結のしきい値決定のための学習データとし, 残り一つでテ

ストする交差検定を実施した。

提案法の特徴を、文献 [3] でも評価されている検出成功率により示す。追跡結果中の各フレームの検出領域が、正解の人領域としきい値以内のずれで重なっていれば検出成功とみなし、全フレームの全領域を対象として正検出と誤検出を評価した。提案法と従来法 [3] の ROC (Receiver Operating Characteristic) 曲線を図 7 に示す。提案法については、進行方向コストのしきい値を変化させた複数の結果を示している。グラフから分かるように、赤バツ印で示される従来法の結果周辺に、提案法の結果が多く分布している^(注3)。このことから、提案法はしきい値の設定に対して頑健であることが分かる。一方、従来法と比べて結果がほとんど改善していないことも分かる。これは、対象検出率が、フレーム独立に検出領域の位置・スケールを評価するだけで、継続的な対象追跡 (同一対象 ID の割り当て) を評価していないためである。すなわち、すれ違い時に対象 ID の切り替わりが発生しても (追跡が失敗しても) 対象検出は成功とみなされるため、提案法の効果は確認しにくい。

そこで、検出成功率 (全フレームの全領域の検出成功を独立に評価) ではなく、追跡成功率で提案法の効果の評価する。追跡成功率は、(追跡成功人数) / (全観測人数) とした。ただし、各人の追跡成功は、追跡結果と正解を比較して、各追跡対象について初期フレーム (最初に対象が画像内で観測されるフレーム)

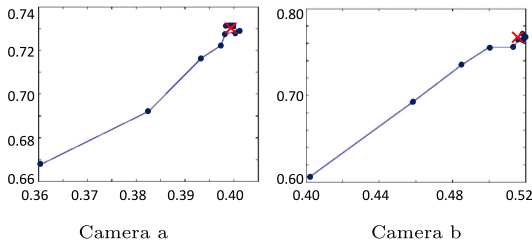


図 7 従来法と提案法の対象検出性能の比較。図の ROC 曲線の縦軸は正検出率、横軸が誤検出率を表す。青点は進行方向コストのしきい値を変化させた提案法の結果で、赤のバツ印は進行方向コストを利用していない従来法 [3] の結果である。

Fig. 7 Comparison of detection accuracy. Vertical and horizontal axes of a ROC curve show true-positives and false-positive, respectively. A red cross in each graph shows the result of [3].

(注3)：「少しでも進行方向が違ってればノードとノードを連結しない」ような極端に厳しいしきい値を与え際には、検出成功率が各グラフの左下の方向に向かって悪化している。

から最終フレーム (対象が最後に画像内で観測されるフレーム) まで正しく同一 ID が割り振られ、かつ、全検出フレームに渡って検出領域の正解人領域のずれがしきい値以内 (前述の検出成功と同基準) であることとした。検出成功率と比べ、より適切に提案法の効果を確認することが、追跡成功率評価の目的である。

進行方向制約の有無による追跡成功率の変化を、表 1 に示す。二つのカメラともに、追跡成功率が若干向上したが、大きな差は得られなかった。この原因を示すため、すれ違いにおける追跡結果が本手法によって改善されなかった例を図 8 に示す。T frame において、ID15 がそれまで ID104 だった領域に割り振られると同時に、ID15 だった領域は追跡失敗して ID15 は消えてしまった。その後、新たに ID42 が、ID15 だった領域に割り振られて追跡が行われた。この例では、フレーム間で検出結果の位置変化が少ないために、進行方向を安定に計算できなかったことが追跡失敗の主な原因である。前述のとおり、提案法では、動的計画法などによって最適解を得る。この制約上、グラフ中では隣接フレームとの間にしかエッジをはれず、進行方向は隣接フレーム間の検出領域の移動量のみから推定することになる。そのため、進行方向を安定に得ることが難しい。

進行方向を安定に得るため、動画中のフレームをスキップして、すなわち、元の動画の N_S フレーム間隔で動画を再構成し、その再構成動画に対して提案手法を適用する。これにより、隣接フレーム間の人の移動量が大きくなり、進行方向が安定に求まるようになる。しかし、 N_S を大きくするほど、4. で示したフレーム間コスト (1) の「フレーム間の矩形の位置の差」が同

表 1 従来法と提案法の追跡成功率。
Table 1 Comparison of tracking accuracy.

追跡成功率 (%)	カメラ A	カメラ B
従来法 [3] (進行方向制約なし)	65.8	60.5
提案法 (進行方向制約あり)	68.5	61.0

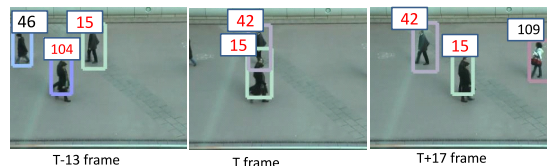


図 8 提案法による ID15 と ID104 のすれ違いの失敗例。
Fig. 8 Unsuccessful tracking results of the proposed method.

一対象のノード間で大きくなってしまい、追跡に失敗しやすくなる。よって、スキップフレーム数 N_S は、進行方向計算の安定性と矩形の位置の差のトレードオフに基づいて決めなければならない。このトレードオフの関係を、実験的に示す。

フレームスキップの効果は、追跡成功率により評価する (図 9)。追跡成功率は、(追跡成功人数) / (全観測人数) とした。ただし、各人の追跡成功は、追跡結果と正解を比較して、各追跡対象について初期フレーム (最初に対象が画像内で観測されるフレーム) から最終フレーム (対象が最後に画像内で観測されるフレーム) まで正しく同一 ID が割り振られ、かつ、全検出フレームに渡って検出領域の正解領域のずれがしきい値以内 (前述の検出成功と同基準) であることとした。グラフ横軸は、スキップをしない場合の結果を 0 とし、1 フレーム、2 フレーム、3 フレームごとにスキップしたときに対応する。

本実験では、交差検定において、スキップフレーム数ごとにノードを結ぶ進行方向コストのしきい値を最適化して与えた。スキップフレーム数が増えれば再構成動画のフレーム間の進行距離が増え、式 (1) で定義される正しい追跡結果に対応する進行コストの振幅も大きくなるため、スキップフレーム数ごとのしきい値調整が望ましい。交差検定を行うため、スキップフレーム数ごとに最適なしきい値を与えてもユーザの労力は変化なく、実用上の問題はない。スキップフレーム数 0, 1, 2, 3 に対して、カメラ A ではしきい値 10, 30, 40, 50 となり、カメラ B ではしきい値 20, 40, 50, 60 となった。スキップフレーム数増加に伴いしき

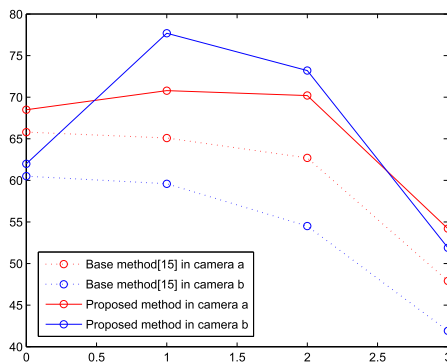


図 9 画像をスキップしたときの追跡結果. 横軸がスキップした画像の枚数, 縦軸が追跡成功率を示す.

Fig. 9 Tracking results with sequences of skipped frames. Horizontal and vertical axes indicate the interval of images and tracking accuracy.

い値も大きくなるという、直感的な理解と等しい。

図 9 から、このデータセットでは、1 フレームスキップの結果が最も良く、3 フレームスキップからは大きく成功率を下げていることがわかる。この結果から、進行方向計算の安定性と対象位置の差のトレードオフを考慮してスキップフレーム数を決定する必要性が確認できた。このスキップフレーム数の決定には、対象検出座標の誤差と、フレーム間で対象が移動する画素数が重要な基準である。その移動画素数は、空間中での対象の移動速度、カメラから対象までの距離、カメラの撮影時間解像度などに依存する。これら多くの要因を、各カメラにおいて全てパラメータ化することは現実的には難しい。実際に撮影された画像を検証用画像として、実験的にもっとも高い追跡成功率を得るフレームスキップ数を交差検定により得るのが、実現容易な方法であると思われる。

図 9 のグラフで示したように、統計的には最高の成功率を得るスキップフレーム数 N_S (この例では $N_S = 1$) が、いずれのシーケンスでも最高の結果を得る保証はない。例を図 10 に示す。スキップフレームなしの場合、この例では、 $T - 10$ frame における ID15 と ID104 のすれ違い時に追跡が失敗した。 $N_S = 1$, $N_S = 2$ の場合も、両者の追跡に失敗したが、 $N_S = 3$ では、成功した。今後は、より適切なコストやコストしきい値の決定法などについて検討していく必要がある。

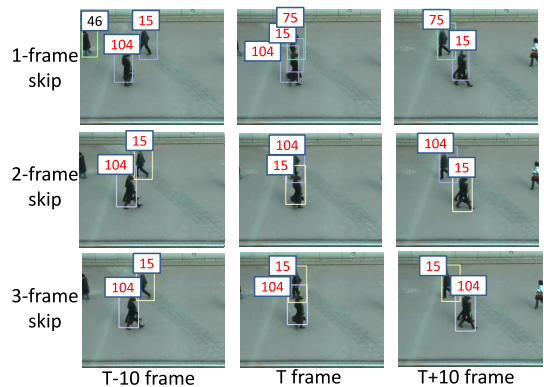


図 10 図 8 と同じシーケンスにおいて、画像をスキップしたときの追跡結果. スキップフレーム数が 3 の場合のみ、ID15 と ID104 の追跡結果は入れ替わらず追跡成功している。

Fig. 10 Tracking results with skipped frames. This sequence is equal to the one shown in Fig. 8. Tracking results of ID15 and ID104 were correct only when $N_S = 3$.

7. む す び

大域的最適追跡における人のすれ違いによる失敗を、グラフのノード分割による進行方向制約によって改善した。従来法と比較して、提案法により追跡成功率が向上していることを定量的に確認した。

グラフィカルモデルによる大域的最適追跡では、効率探索のために隣接フレーム間のノードのみを結ぶため、追跡中に1フレームでも検出失敗すると継続的な追跡ができない。この問題の解決のため、1) 人検出の成功率向上、2) フレームをまたいでエッジを結ぶ高次グラフにおける最適探索、3) 途中で切れてしまった追跡結果を繋ぐ後処理、などが課題となる。

文 献

- [1] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," CVPR, pp.203-208, 2000.
- [2] M. Isard and A. Blake, "Condensation - conditional density propagation for visual tracking," Int. J. Comput. Vis., vol.29, no.1, pp.5-28, 1998.
- [3] H. Pirsivash, D. Ramanan, and C.C. Fowlkes, "Globally-optimal greedy algorithms for tracking a variable number of objects," CVPR, pp.1201-1208, 2011.
- [4] T.E. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Sonar tracking of multiple targets using joint probabilistic data association," IEEE J. Ocean. Eng., vol.8, no.3, pp.173-184, 1983.
- [5] M. Isard and J. MacCormick, "Bramble: A bayesian multiple-blob tracker," ICCV, pp.34-41, 2001.
- [6] W. Choi and S. Savarese, "Multiple target tracking in world coordinate with single, minimally calibrated camera," ECCV, pp.553-567, 2010.
- [7] C. Stauffer, "Estimating tracking sources and sinks," IEEE Workshop on Event Mining, 2003.
- [8] L. Zhang, Y. Li, and R. Nevatia, "Global data association for multi-object tracking using network flows," CVPR, 2008.
- [9] H. Jiang, S. Fels, and J.J. Little, "A linear programming approach for multiple object tracking," CVPR, 2007.
- [10] A. Andriyenko and K. Schindler, "Globally optimal multi-target tracking on a hexagonal lattice," ECCV, pp.466-479, 2010.
- [11] Y. Ma, Q. Yu, and I. Cohen, "Target tracking with incomplete detection," CVIU, vol.113, no.4, pp.580-587, 2009.
- [12] S. Pellegrini, A. Ess, and L.J. Van Gool, "Improving data association by joint modeling of pedestrian trajectories and groupings," ECCV, pp.452-465, 2010.
- [13] A.A. Butt and R.T. Collins, "Multi-target tracking by lagrangian relaxation to min-cost network flow," CVPR, pp.1846-1853, 2013.
- [14] A. Ess, B. Leibe, K. Schindler, and L.J. Van Gool, "A mobile vision system for robust multi-person tracking," CVPR, 2008.
- [15] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," CVPR, pp.886-893, 2005.
- [16] S. Maji, A.C. Berg, and J. Malik, "Classification using intersection kernel support vector machines is efficient," CVPR, 2008.
- [17] M. Hirzer, P.M. Roth, and H. Bischof, "Person re-identification by efficient impostor-based metric learning," AVSS, pp.203-208, 2012.
- [18] R.K. Ahuja, T.L. Magnanti, and J.B. Orlin, Network flows - theory, algorithms and applications, Prentice Hall, 1993.

(平成 26 年 6 月 13 日受付, 8 月 8 日再受付,
9 月 17 日早期公開)



岡田 亜沙美

2014 年奈良先端科学技術大学院大学修士課程修了。現在、アルソックセキュリティ科学研究所勤務。在学中、人群解析及び広域対象追跡に関する研究に従事。



浮田 宗伯 (正員：シニア会員)

2001 年京都大学大学院博士後期課程修了。同年奈良先端科学技術大学院大学情報科学研究科助手。2007 年同准教授。2002 年～2006 年まで科学技術振興機構さきがけ(「情報基盤と利用環境」領域)研究員兼任。2007 年～2009 年までカーネギーメロン大学客員研究員兼任。博士(情報学)。対象追跡、人体運動解析・姿勢推定に関する研究に従事。



萩田 紀博 (正員：フェロー)

1978 年慶應義塾大学大学院工学研究科電気工学専攻修士課程修了。同年電電公社(現 NTT)武蔵野電気通信研究所入所。NTT 基礎研究所、ATR メディア情報科学研究所長などを経て、現在 ATR 知能ロボティクス研究所長。この間、文字認識、画像認識、コミュニケーション科学、コミュニケーションロボットなどの研究に従事。工学博士。IEEE、情報処理学会、日本ロボット学会、人工知能学会各会員。