

時空間 MRF モデルに基づく車両と歩行者の統一的トラッキング

上 條 俊 介[†] 坂 内 正 夫[†]

交差点において効率的な交通流を実現するためには、車両交通流だけでなく歩行者流をも観測して信号制御を行うことが重要である。一方、交差点において歩行者の安全を確保することを考慮すると、歩行者と車両の挙動の相互関係を解析し、事故を誘発しやすい状態を解明することが重要となる。このように、交差点における交通の安全と効率を両立するという目的のため、時空間 MRF モデルを改良し、これを適用したトラッキングアルゴリズムを開発した。時空間 MRF モデルは著者等が 2000 年より提唱している時空間画像における移動物体の領域分割を行うための確率モデルであるが、オクルージョンに強いという特性を持っている。本論文では、従来から提案している時空間 MRF モデルを拡張し、車両のような剛体だけでなく、歩行者のように部分ごとに異なる動きをする非剛体をも精度良くトラッキングできるように改良した。このように交差点画像に単一のアルゴリズムを適用した実験において、車両と歩行者を統一的にトラッキングするのに有効であることが確認された。

Simultaneous Tracking of Vehicles and Pedestrians by a Unified Algorithm of the Spatio-temporal MRF Model

SHUNSUKE KAMIJO[†] and MASAO SAKAUCHI[†]

To achieve efficient traffic flow, it is important to base control of traffic signals on observation of pedestrian flow as well as vehicle flow. In consideration of safety, it is also important to analyze behavioral relationship between pedestrians and vehicles, which tends to lead accidents at intersections. Toward the goals of efficiency and safety, we developed a precise tracking algorithm based on the Spatio-Temporal MRF model which is able to track both pedestrians and vehicles simultaneously against occlusions in the images. During the past few years, this model has been practically applied to acquire traffic flow statistics. However, in this paper, we present an improvement of the S-T MRF model so as to deal with flexible objects such as pedestrians as well as rigid objects such as vehicles. Based on experimental results, this model was able to simultaneously track pedestrians and vehicles against occlusion even in very cluttered situations. Consequently, the improved S-T MRF model was proven to be effective for traffic monitoring at urban intersections.

1. はじめに

トラッキングアルゴリズムに関しては、いうまでもなく従来から様々な研究が行われている。特に、車両のトラッキングに関しては、広く様々な研究がなされてきた。たとえば、Kolling ら¹⁾の研究は、その成功例としてよく知られている。Peterfreund ら³⁾は、‘Snakes’⁴⁾の手法をトラッキングに適用し、車両の輪郭線の抽出を行っている。Smith ら⁵⁾と Grimson ら⁶⁾は、オプティカルフローの統計解析やクラスタリングおよびベクトル量子化を用いている。しかし、これらの研究は、いずれもオクルージョンの問題を扱っていなかった。これに対し、Leuck ら⁷⁾と Gardner ら⁸⁾

は、車両形状の 3 次元モデルを仮定してパターンマッチングによりオクルージョンの問題を扱うことを試みている。しかし、この方法は車両形状の 3 次元を必要とすることから、様々な種類の車両や移動物体が混在する画像には適用が難しい。また、Malik ら⁹⁾は、オクルージョンの問題を reasoning 手法により解決しようと試みている。これらの方法は、いずれも移動物体の個数密度が低い画像では良いトラッキング結果を示しているが、一般の交通映像のように移動物体の個数密度が高い混雑した画像でランダムに動く移動物体どうしのオクルージョンの問題をロバストに扱うことは困難である。

一方、時系列画像から人物を抽出するトラッキングアルゴリズムに関しても、従来から様々な研究が行われてきた。羽下ら¹³⁾は、時間平均シルエットを用いて非剛体の移動物体をトラッキングしている。Li ら¹⁴⁾

[†] 東京大学生産技術研究所

Institute of Industrial Science, The University of Tokyo

は、画像中の車両や顔領域を事後確率推定を用いて検出している。Kornprobst¹⁵⁾ は、テンソル投票空間を用いてトラッキングを行っている。Haritaoglu¹⁶⁾ は、店舗内の人物追跡のために 2 値画像のシルエットおよび動きの物理的な拘束条件に注目した研究を行っている。Sminchisescu¹⁷⁾ は、人間の体の 3 次元モデルに基づいている。さらに、Hidalgo¹⁸⁾ は、時系列画像中の前景をキー領域としてトラッキングを行い、Izquierdo¹⁹⁾ は、動きに基づく画像領域分割を試みている。これらの方法は、いずれも有効性が示されているが、基本的に単体の人物のみを扱っており、オクルージョンの問題を扱っていない。

2001 年に発表された Paragios¹⁰⁾ および Tsaig¹¹⁾ による関連研究は、それぞれ歩行者や室内で机の前に座っている人物に関し、連続画像中の動きを MRF モデルを用いて領域分割することを試みている。しかし、文献 10) ではオクルージョンの問題を扱っていない。また、文献 11) ではオクルージョンの問題に言及されているが、当該論文で用いられている手法では、領域分割アルゴリズムにおいて時間軸方向の相関に関する定量的な定義が明確に述べられておらず、そのオクルージョンへのロバスト性を評価するための実験結果も示されていない。

これに対し、本論文の著者らが 2000 年に提唱した時空間 MRF モデル^{24)~26)} は、時間軸方向のテクスチャおよび領域についての相関性に着目し、これを評価関数の中で定量化することで時空間画像の領域分割を行っている。これは、2 次元すなわち空間軸方向のテクスチャ相関性はその存在が推定されるものの保証の限りでないのに対し、移動物体の時間軸方向へはテクスチャ相関性が必ず存在することが論拠となっている。

2. 本論文の目的

市街地の交差点における車両と歩行者の両者の状態やこれらの相互関係を詳細に把握することが、より効率が良く安全な交通を実現するために重要である。著者らが 2000 年より提唱してきた時空間 MRF モデル^{24)~26)} は、それ以来、様々な車両交通画像を用いた評価において有用であることが確認されたが、本アルゴリズムは主に車両のような剛体に対して適用され、歩行者のような非剛体については適用されていなかった。

時空間 MRF モデルは、移動物体の動きベクトルに基づいた時間軸方向のテクスチャ相関や時間軸および空間軸方向の領域相関をもとに、時空間画像の領域分割を最適化するアルゴリズムである(図 2)。しかし、



図 1 交差点における横断歩道の画像
Fig. 1 An image of crosswalk at an intersection.

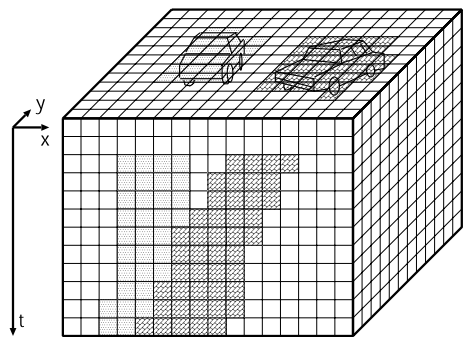


図 2 時空間 MRF モデルによる領域分割の概念
Fig. 2 Segmentation of spatio-temporal images.

従来のアルゴリズムでは剛体性を仮定したため、ある剛体領域に属するブロックに関して最適化の際に用いる動きベクトルとしては、剛体に関する代表値を適用していた。このため、図 1 における人物のような非剛体においてオクルージョンが生じている場合、非剛体どうしの境界で領域分割の最適化が正しく行われなかったという問題があった。

そこで、本論文では、従来の時空間 MRF モデルを改良し、車両等の剛体と歩行者等の非剛体の両者を統一的に扱い、精度の良いトラッキングを行うことを可能とすることを目的とし、図 1 に見られるような非剛体である歩行者同士が重なる画像を用いて領域分割のロバスト性を検証した。ここで、本論文での改良を行った時空間 MRF モデルは従来の拡張的な位置づけであり、車両等剛体の領域分割に関しても従来どおり行えるものである。したがって、本論文の実験においては、車両と歩行者の混在する画像を用いた実験を行っている。

3. 時空間 MRF モデルによる領域分割

3.1 最適化アルゴリズムの条件

一般にピクセル単位での領域分割は困難である．特に画像のフレームレートが低い場合には，隣接画像フレーム間で移動物体が 10～20 ピクセル程度も移動する場合があります，回転や遠近により，2 次元画像面に投影された見かけの形状も変化する．そこで，時空間 MRF モデルでは，従来より 8 ピクセル×8 ピクセルの集合を 1 ブロックと定義し，このブロックを単位とした領域分割を行っている．従来の領域分割の最適化においては，ブロック間のテクスチャ情報および領域ラベル情報の相関が考慮されている^{(24),(25)}が，本論文における改良では，さらにブロック間の動きベクトルの相関が考慮される．

本論文で，時空間 MRF モデルにおけるブロックとは，空間 MRF モデルにおけるサイトと同義として最適化の単位を指し，背景差分により移動物体が存在すると判定されたブロックについての領域分割を行う．本論文では，640 ピクセル×480 ピクセルの画像を用いているため，80 ブロック×60 ブロックに領域ラベルを付与することで領域分割が行われ，これを Object-map と呼ぶ．すなわち，時空間 MRF モデルは，前フレームに関する画素値 $G(t-1; i, j) = g(i, j)$ および Object-map $X(t-1) = x$ ならびに現画像フレームに関する画素値 $G(t, i, j) = h(i, j)$ が与えられたときに，現画像フレームに関する Object-map $X(t) = y$ を求める問題である．ここで， t は画像のフレーム番号すなわち時間座標を表し， i, j は画像の空間座標を表す．

著者らの従来研究^{(24),(25)}においては，最適化に際し 3 つのパラメータ M_{xy_k} ， D_{xy_k} ， N_{y_k} を用いた相関エネルギーを評価していたが，本論文では新しく動きベクトル $V_{C_k(t)}$ ， $V_{B_k(t)}$ をパラメータに導入し，これらを同時に評価するためのエネルギー関数を定義した．また， M_{xy_k} ， D_{xy_k} ， N_{y_k} を用いた評価に際して参照する動きベクトルが，従来研究においては移動物体を代表する動きベクトルであったのに対し，本論文ではブロックに属する局所的な動きベクトルとなる点が大きく異なる．

3.2 時間軸方向の相関パラメータ

以下のエネルギー関数は，パラメータ M_{xy_k} および D_{xy_k} を用いた Boltzmann 分布の指数部分から導かれる^{(24),(25)}．

まず最適化の前に各ブロックごとに，前フレームの画像 $G(t-1)$ から現フレームの画像 $G(t)$ への動き

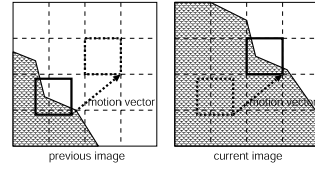


図 3 連続画像フレーム間のテクスチャ相関

Fig. 3 Texture correlation between consecutive image frames.

ベクトルをブロックマッチング手法を用いて求める．このように得られた動きベクトルの分布を用いてエネルギー関数 (1) が評価される．

$$U_{pre}(D_{xy_k}, M_{xy_k}) = b(M_{xy_k} - \mu_{M_{xy}})^2 + c(D_{xy_k} - \mu_{D_{xy}})^2 \quad (1)$$

$$D_{xy_k} = \sum_{0 \leq di < 8, 0 \leq dj < 8} |G(t; i + di, j + dj) - G(t-1; i + di - v_{mi}, j + dj - v_{mj})| \quad (2)$$

ここで，3.1 節で定義した記号に基づき， M_{xy_k} ， D_{xy_k} の添え字 x は時刻 $t-1$ における画像フレームに対して求められた Object-map が $X(t-1) = x$ であったことを表す．同様に添え字 y は，求めようとしている時刻 t における画像フレームに対する Object-map が $X(t) = y$ と推定されていることを示し， y の添え字 k は決定しようとしている Object-map $X(t) = y$ のうち k 番目のブロック $C_k(t)$ を対象とした最適化評価を行っていることを示す．また， $\mu_{M_{xy}}$ および $\mu_{D_{xy}}$ は，エネルギー関数が最小化されるために M_{xy_k} および D_{xy_k} がとるべき値を示しており，具体的な値は後述する．これらの評価関数は，著者らの論文^{(24),(25)}において導出されており，導出過程については当該論文を参照していただきたい．

第 1 のパラメータ D_{xy_k} は，式 (2) に示すように画素値の絶対差分和による画像 $G(t-1)$ と画像 $G(t)$ との間のテクスチャ相関を表す．いま，現フレームにおいて最適化の対象となっているブロック $C_k(t)$ に関する評価を行うことを考える．このテクスチャ相関は，付与する候補となるラベル O_m で表される移動物体に属する対象ブロック $C_k(t)$ に関して，像 $G(t)$ から画像 $G(t-1)$ への局所的な動きベクトル $-\vec{V}_{O_{mk}} = (-v_{mki}, -v_{mkj})$ を基に算出される (図 3)． $U_D(D_{xy_k})$ は $D_{xy_k} = u_{D_{xy}} (= 0)$ において最小値をとる．これは， D_{xy_k} が小さいほど， C_k が当該ラベルで表される移動物体に属する確率が高いことを表している．

なお，従来の時空間 MRF モデルでは， D_{xy_k} の評価において参照されるのは，いま決定しようとし

ている Object-map $X(t) = y$ の評価対象ブロック $C_k(t)$ に付与されている候補ラベルで表される移動物体 O_m の時刻 $t-1$ から t への動きベクトル $\vec{V}_{O_m(t-1;t)} = (v_{mi}, v_{mj})$ であった。当該従来手法では $\vec{V}_{O_m(t-1;t)}$ は、Object-map $X(t-1) = x$ において移動物体 ID O_m が付与されているブロックに関する時刻 $t-1$ から t への動きベクトルの最頻出値で代表されていたため、局所的な動きベクトル情報は参照されていなかった。

これに対し、本論文で改良されたモデルでは、 D_{xy_k} の評価において局所的な動きベクトル $\vec{V}_{O_{mk}(t-1;t)} = (v_{mki}, v_{mkj})$ が参照される点が異なる。ここで、いま最適化対象が Object-map $X(t) = y$ のブロック $C_k(t)$ である場合、 D_{xy_k} の評価に際して、すでに最適化された Object-map $X(t-1) = x$ のうちのどのブロックに関する動きベクトルを用いるかが問題となる。これには、3.5 節で述べる最適化プロセスの Step.2 において $X(t-1) = x$ から $X(t) = y$ へのマッピングを行う際に対応するブロックの動きベクトルを用いる。

第 2 のパラメータ M_{xy_k} は、最適化の対象である現フレームの Object-map $X(t) = y$ と決定済みである前フレームの Object-map $X(t-1) = x$ との相関を定量化するものである。いま、現フレームにおいて最適化の対象となっているブロック $C_k(t)$ に関する評価を行うことを考える。すると同様に、対象ブロック $C_k(t)$ における局所的な動きベクトル $-\vec{V}_{O_m} = (-v_{mi}, -v_{mj})$ を基に $C_k(t)$ を移動させたとき、前フレームの Object-map $X(t-1) = x$ においてラベル O_m が付与されているブロックとの重なっているピクセルの数として M_{xy_k} が求められる (図 4)。したがって、1 ブロックは 64 ピクセルで構成されることから、 $U_M(M_{xy_k})$ は $M_{xy_k} = \mu_{M_{xy}} (= 64)$ で最小値を、 $M_{xy_k} = 0$ で最大値をとる。すなわち、この重なりが大きいほど、当該ラベルの付与がより確からしいことを表す。

同様に、本論文で改良された時空間 MRF モデルでは、局所的な動きベクトルを参照する点が従来のモデルと異なる。

3.3 空間軸方向の相関パラメータ

第 3 のパラメータ N_{y_k} は、最適化の対象となっている現フレームの Object-map $X(t) = y$ 中におけるラベル分布の隣接相関を表すものである。

$$U_N(N_{y_k}) = a(N_{y_k} - \mu_{N_{y_k}})^2 \quad (3)$$

ここで、3.1 節で定義した記号に基づき、 N_{y_k} の添え字 y は、求めようとしている時刻 t における画像フレームに対する Object-map が $X(t) = y$ と推定されていることを示し、 y の添え字 k は決定しようとし

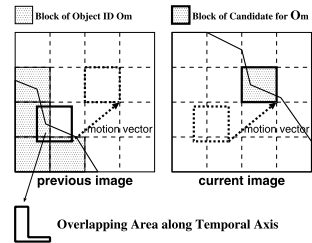


図 4 連続画像フレーム間の隣接条件

Fig. 4 Neighboring condition between consecutive image frames.

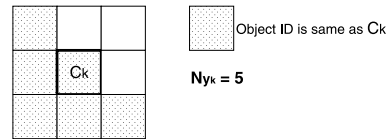


図 5 空間内のブロック隣接条件

Fig. 5 Neighboring condition in a spatial plane.

ている Object-map $X(t) = y$ のうち k 番目のブロック $C_k(t)$ を対象とした最適化評価を行っていることを示す。また、 $\mu_{N_{y_k}}$ は、エネルギー関数が最小化されるために N_{y_k} がとるべき値を示しており、具体的な値は後述する。この評価関数についても、著者らの論文^{(24), (25)} において導出されており、導出過程については当該論文を参照していただきたい。

N_{y_k} は最適化対象ブロック C_k にラベル O_m が付与されているとき、 C_k の隣接ブロック (図 5) のうち、同一のラベル O_m が付与されているブロックの数を表す。したがって、隣接ブロックは 8 個あることから、 $U_N(N_{y_k})$ は $N_{y_k} = \mu_{N_{y_k}} (= 8)$ において最小値をとる。すなわち、同一のラベルが付与されている隣接ブロックの数が多いほど、当該ブロックのラベル付与がより確からしいことを表す。

3.4 空間軸方向の相関パラメータの追加項

しかしながら、ブロックマッチング法を適用しただけでは、テクスチャの類似性や対応テクスチャの非存在等の問題から、動きベクトルが正しく求まらないことがある。特に移動物体の領域境界においては、オクルージョンによって対応するテクスチャが隠れることもあり、対応テクスチャの非存在により動きベクトルのエラーが顕著に生じる。こうしたエラーは、領域最適化時に当該ブロックが本来属すべき移動物体に属さないとする誤判定を生じる原因となる。特に領域面積の小さい移動物体では、このような誤判定により当該移動物体の ID ラベルが本来付与されるべきブロックが Object-map からすべて消滅してしまうことが生じうる。この問題は移動物体のトラッキングに失敗する

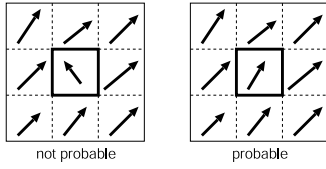


図 6 空間隣接ブロック間の動きベクトル相関

Fig. 6 Motion vector correlation among neighboring blocks in a spatial plane.

という結果を導くため、最適化プロセスにおいて当該動きベクトルのエラーを修正することが必要となる。このことは、人物のような非剛体の領域分割においてはさらに重要である。

この目的のためには、各ブロックごとに求める動きベクトル自体を、空間隣接ブロックの動きベクトルとの類似性を考慮したうえで（図 6）、先述の時間軸方向の最適化条件 M_{xy_k} および D_{xy_k} をも同時に満たすことを条件として確率緩和を行う必要がある。

この第 4 のエネルギー項として、著者らの論文^{24),25)} にはない以下の項を追加する。

$$U_{mv}(C_k(t-1)) = f \sum_{B_k} |\overrightarrow{V_{C_k(t-1)}} - \overrightarrow{V_{B_k(t-1)}}|^2 / N_{x_k} \quad (4)$$

ここで、 $B_k(t-1)$ は時刻 $t-1$ におけるブロック $C_k(t-1)$ の隣接ブロックを表し、図 5 のような隣接ブロックとして定義される。また、 $\overrightarrow{V_{C_k(t-1)}}$ 、 $\overrightarrow{V_{B_k(t-1)}}$ はそれぞれ、ブロック $C_k(t-1)$ 、 $B_k(t-1)$ に関する時刻 $t-1$ の画像フレームから時刻 t の画像フレームへの動きベクトルを表す。

この式自体は、動きベクトルの空間類似性のみを表しており、ブロック $C_k(t-1)$ における画像 $G(t-1)$ から $G(t)$ への動きベクトルと隣接ブロック $B_k(t-1)$ における動きベクトルとの類似性が高いほど、ブロック $C_k(t-1)$ において求められた動きベクトルが確からしいことを表す。人物等の非剛体であっても、隣接ブロック間では動きベクトルは連続的に変化するため、この条件が有効となる。

3.5 最適化プロセス

結果として、この最適化問題は、下記のエネルギー関数 (5) を最小化するための $X(t) = y$ を求める問題に帰着する。

$$\begin{aligned} U_{stmr} &\equiv U_{pre} + U_N + U_{mv} \\ &= a(N_{y_k} - \mu_{N_y})^2 + b(M_{xy_k} - \mu_{M_{xy}})^2 \\ &\quad + cD_{xy_k}^2 \\ &\quad + f \sum_{B_k} |\overrightarrow{V_{C_k(t-1)}} - \overrightarrow{V_{B_k(t-1)}}|^2 / N_{x_k} \end{aligned} \quad (5)$$

ここで、 $U(y_k)$ におけるエネルギー項 $U_M(M_{xy_k})$ と $U_D(D_{xy_k})$ は、上述のように候補となる移動物体の対応ブロックにおける局所的な動きベクトルを基に評価される。また、 $U_{mv}(C_k(t-1))$ は、画像 $G(t-1)$ から $G(t)$ への動きベクトルに対して評価される。

このように、空間隣接ブロックのラベルの類似性および動きベクトルとの類似性、時間軸方向のテクスチャおよび領域相関が同時に評価されることにより、従来では最適化が困難であった非剛体や面積の小さい領域の境界がより正確に決定されることが期待される。

最適化プロセスは、前フレームの画像、Object-map、Motion-Vector-map および現フレームの画像が与えられたときに、現フレームに対する Object-map $X(t) = y$ と前フレームから現フレームへの Motion-Vector-map $V(t-1; t) = v$ を求めることである。ここで、Motion-Vector-map $V(t-1; t) = v$ とは、すべてのブロック $C_k(t-1)$ に対して求めた動きベクトル $\overrightarrow{V_{C_k(t-1)}}$ の集合をいう。

最適化のアルゴリズムを、以下に示す。

[最適化プロセス]

- Step.1: $i = 0$ とする。前フレームから現フレームへの各ブロックの動きベクトルをブロックマッチング法により求めることで、Motion-Vector-map $V(t-1; t) = v^i$ を初期化する。
- Step.2: Step.1 で求めた $v^i (i = 0)$ に基づいて、前フレームで決定された Object-map $X(t-1) = x$ における各ブロックを動かし、現フレームに対する Object-map $X(t) = y^i (i = 0)$ を初期化する。その際、複数の異なる候補 ID のラベルが同一のブロックへ付与されることがあるが、当該ブロックを「最適化対象ブロック」と呼び、Step.3 以下での最適化の対象とする。
- Step.3: 個々の最適化対象ブロックに対し、異なる候補 ID ごとに当該 ID が付与されたかと仮定して、 $X(t) = y^i$ および $V(t-1; t) = v^i$ の情報から式 (5) のエネルギー U_{stmr} をそれぞれ計算する。
- Step.4: Step.3 で計算されたエネルギーに基づき、現ループの状態 $X(t) = y^i$ 、 $V(t-1; t) = v^i$ から次ループの状態 $X(t) = y^{i+1}$ 、 $V(t-1; t) = v^{i+1}$ へ確率緩和手法による状態遷移を行う。
- Step.5: 最適化終了判定。 $X(t) = y^{i+1}$ 、 $V(t-1; t) = v^{i+1}$ が終了条件を満たしていれば終了する。満たしていなければ $i = i + 1$ とし、Step.3 へ戻る。

[最適化プロセス終了]

3.2 節でも述べたように、従来の時空間 MRF モデルと異なる点は、改良されたモデルでは領域境界と動きベクトルをそれぞれブロック単位で最適化することである。このとき、領域最適化で参照する動きベクトルは、移動物体を代表する値ではなく、ブロックに属する固有の値である点が異なるということを再度強調しておく。すなわち、局所動きベクトルには、式 (4) のエネルギー項 U_{mv} による空間軸方向の束縛条件だけでなく、式 (1) のエネルギー項 U_{pre} により時間軸方向の束縛条件も同時に考慮されることとなる。これは、本論文の改良アルゴリズムでは、式 (1) のエネルギー項 U_{pre} において参照される動きベクトルが局所的値となるからである。このように、改良されたモデルは、アルゴリズムの内部構造が従来のモデルとはまったく異なるものである。

Step.1 において、時刻 $t-1$ から t への動きベクトル $\vec{V}_{C_k(t-1)} = (v_{k1}, v_{k2})$ はブロックマッチング法により求められる。その探索範囲は隣接画像フレーム間での物体移動距離等を参考にユーザにより設定され、たとえば本論文の実験では $-24 \leq v_{k1} \leq 24$, $-24 \leq v_{k2} \leq 24$ のように設定された。

Step.2 において Object-map $X(t) = y^0$ の初期状態は、Step.1 で求めた Motion-Vector-map $V(t-1; t) = v^0$ および前フレームの Object-map $X(t-1) = x$ から推定される。より具体的には、前フレームにおいて決定された Object-map $X(t-1) = x$ について、各ブロックを Motion-Vector-map $V(t-1; t) = v^0$ に基づいた量だけ仮に移動させる。たとえば、あるブロックに対する動きベクトルがピクセル単位で $(V_x, V_y) = (17, -4)$ のように求まっている場合、移動量がブロック単位に切り上げまたは切り捨てられ、当該ブロックが x 軸の正方向へ 2 ブロックおよび y 軸の負方向へ 1 ブロック移動したとして Object-map $X(t) = y^0$ へ登録する。その際、複数の移動物体の境界付近では、1 つのブロックに複数の候補 ID が付与されることが起こりうる。この場合には、その複数候補 ID が記憶され、いずれの ID が尤もらしいかについて Step.3 以下のループにおける最適化の対象となる。

Step.3 においては、Step.2 において複数の移動物体 ID が付与されたブロック（最適化対象ブロック）に関して、当該候補 ID が付与されると仮定した場合のエネルギー U_{stmrf} がそれぞれ計算される。

Step.4 においては、最適化対象となるブロックに付与された複数候補 ID に関するそれぞれのエネルギー U_{stmrf} を比較し、尤もらしい $X(t) = y^i$ が確率緩和

過程を用いて決定される。確率緩和過程としては、著者らの論文^{(24),(25)} に記述されていると同様に Gibbs Sampler を用いる。Gibbs Sampler は、次ステップで遷移可能な状態の候補のそれぞれについて遷移確率を求め、乱数を発生させる等して、その遷移確率に比例した確率で状態遷移を行うというサンプリングモデルである⁽²⁰⁾。ここで、動きベクトルの遷移できる範囲は Step.1 における探索範囲と同値に設定され、本論文では $-24 \leq v_{k1} \leq 24$, $-24 \leq v_{k2} \leq 24$ である。

Step.5 においては、たとえば $X(t) = y_i$ と $X(t) = y_{i+1}$ とで ID 値の異なるブロック数が閾値 n_{th_id} より小さいかどうかの判定が行われる。また、 $V(t-1; t) = v^i = (v_1^i, v_2^i)$ と $V(t-1; t) = v^{i+1} = (v_1^{i+1}, v_2^{i+1})$ とで各ブロック k における動きベクトルの差 $|v_{k1}^{i+1} - v_{k1}^i| + |v_{k2}^{i+1} - v_{k2}^i|$ が閾値 v_{diff} 以下であるブロックの数が閾値 n_{th_mv} より小さいかどうかの判定が行われる。そして、これら両方の指標が閾値以下になったときに最適化ループを終了する。実際には、解が振動して収束しない場合にループを強制終了するための制限回数を設ける等も行われる。

このように、最適化の各ループにおいて、各ブロックに対する解は ID と動きベクトルで構成される位相空間内をエネルギーの低い方向へ確率的に探索される。その結果、当該決定すべき対象ブロックの隣接ブロックの ID や動きベクトルの分布が、ループごとに絶えず変化する。したがって、当然のことながら、各ブロックについて起こる ID や動きベクトルの変化が、最適化の同一ループ内で処理の時系列順に隣接ブロックへ伝播しないようにする必要がある。すなわち、 $X(t) = y^i$ から $X(t) = y^{i+1}$ への書き換えは、コンピュータメモリ上の同一領域を時系列順に書き換えるのではなく、別の領域に書き換えておいたものを一時に入れ替える等の処理が必要である。

4. 実験結果

4.1 実験のセットアップ

交差点の画像を用いて、歩行者と車両を同時にトラッキングした結果を示す。本実験に使用した画像は、晴れの日に都内某所ビルの 3 階バルコニーから撮影した、車両と歩行者が混在する合計約 50 分の画像を使用した。撮影には Sony Handycam DCR-VX1000 を使い、画像解析はアナログ再生した映像を 640 ピクセル × 480 ピクセルの画像としてキャプチャして行った。パラメータは試行錯誤の結果、 $a = 1/2$, $b = 1/256$, $c = 32/1000000$, $f = 1/4$ とした。

なお、本論文で行った改良により、3.5 節の最適化



図 7 トラッキング結果の例
Fig.7 Example of tracking result.

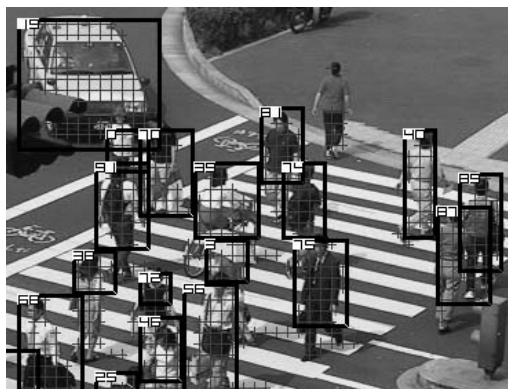


図 9 混雑した歩行者のトラッキング結果
Fig.9 Tracking result of clutter pedestrians.

ループ (Step.3, 4, 5) の処理時間は、著者らが開発した従来のアルゴリズムに比べて約 2 倍になった。このことにより、画像の読み込み・背景差分・3.5 節の最適化プロセス (Step1, 2, 3, 4, 5) を含むプログラム全体の処理時間は約 30% 増加した。その結果、改良されたプログラム全体の処理は平均で、実験で使用した PentiumIV 2.4 GHz CPU の処理能力の 65% を占めることとなった。

4.2 歩行者および車両の統一的トラッキング

図 7 は、交差点付近において歩行者帯および車道を約 8 m の高さから撮影した画像に対して、本論文のアルゴリズムを適用した結果である。当画像において典型的には、歩行者は数～十数ブロック、車両は数十～百ブロック程度で構成されている。このように、車両と歩行者の両者を単一のアルゴリズムでトラッキングすることができることが確認された。なお、画面端の方で外接長方形が移動物体を完全に覆っていないのは、本実験において、画面の外周から 3 ブロック程度をトラッキング処理対象外の領域としているためである。

次に、図 8 (a)(b)(c) は、様々なパターンで生じるオクルージョンに対しトラッキングを行った結果の画像を時系列に示している。図 8 (a) は車両どうしのオクルージョンに対するトラッキング結果を、図 8 (b) は歩行者と自転車との間のオクルージョンに対するトラッキング結果を、図 8 (c) は、歩行者と車両との間のオクルージョンに対するトラッキング結果を、それぞれ時系列に示している。特に、frame c-4 に見られるように、タクシーに乗ろうとしている人物とタクシーとの間のオクルージョンも正確に分離できている。結果として、改良された時空間 MRF モデルのこのようなシーンでの有効性が示されたといえる。

最後に、図 9 は非常に混雑した横断歩道において歩

行者同士・車両・自転車の分離を行った結果の画像を示している。このような激しいオクルージョンの状況においても、時空間 MRF モデルは異なる種類の移動物体間の分離に成功している。

3.5 節で述べたように、本論文のアルゴリズムによれば、非剛体と剛体のオクルージョンの問題を統一的に扱えるだけでなく、オクルージョンによって面積の減少した小領域の境界がより正確に決定されることが期待される。そこで本節では、オクルージョンしている物体の領域境界判定に関し、画面上に現れている領域面積に対する領域分割の頑健性を定量的に評価した (図 10)。

4.3 領域サイズとトラッキング精度の相関に関する実験

図の横軸は、オクルージョンが生じている移動物体領域に関し、画面に現れている領域面積をブロック数で表している。図の縦軸は、当該領域が隣接領域から独立した領域として分離できた成功率を表している。当該ブロック数には、領域分割の結果として得られるブロック数ではなく、本来画像中に見えている領域のブロック数を用いている。これにともない、たとえば本来 10 ブロックで構成されるはずの領域が 5 ブロックの領域として出力されている場合でも、領域が存在するように出力されていれば、領域分割は成功したものととして計数している。逆に失敗していると計数される場合は、オクルージョンによって隣接している移動物体領域の中に吸収され、独立した領域としては出力されない場合である。なお、計数に際して、オクルージョンされている側の物体かオクルージョンしている側の物体かは区別していない。

図 10 における 2 本の折線グラフのうち、下に位置するものが従来の時空間 MRF モデルによる実験結果

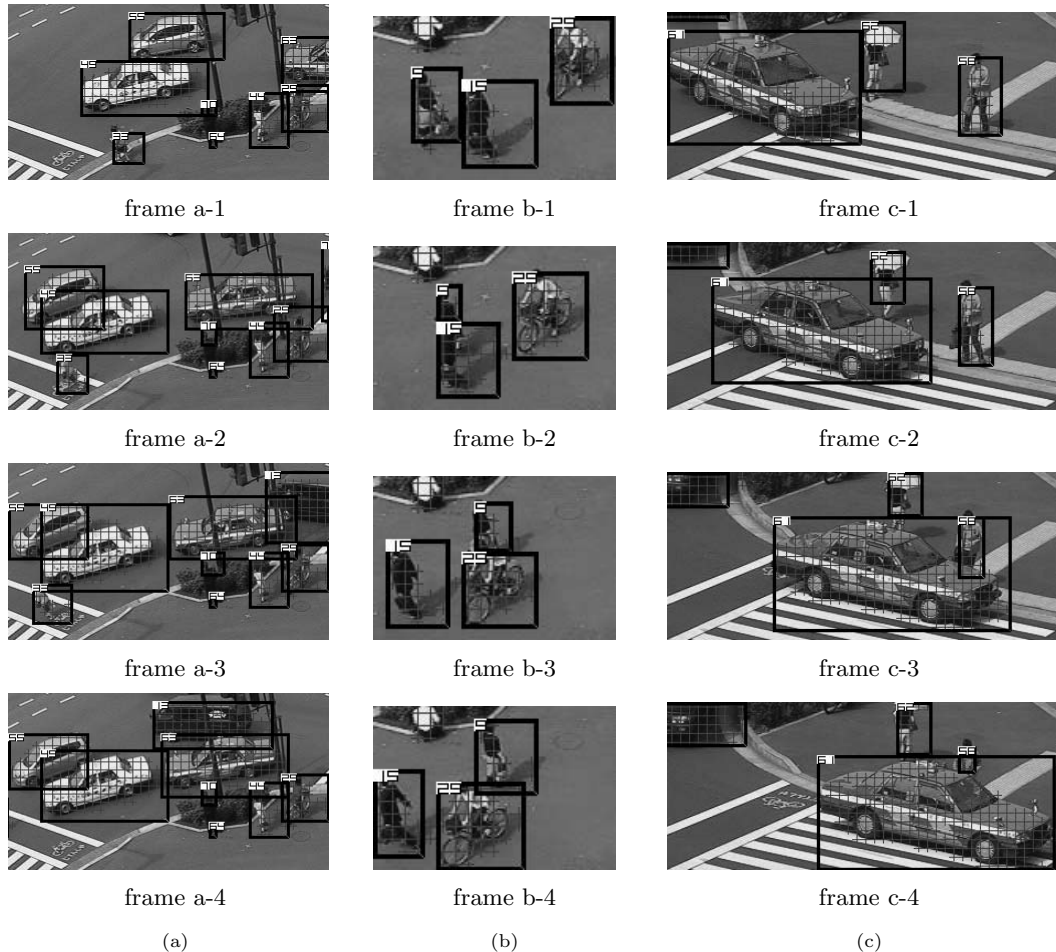


図 8 オクルージョン状況におけるトラッキング結果の時系列画像
 Fig. 8 Tracking sequence in occlusion situation.

を、上に位置するものが本論文で改良したモデルによる実験結果を示している。一見して分かる通り、本論文で改良した時空間 MRF モデルでは、面積が小さい領域における領域分割精度が向上し、特に面積が 3 ないし 5 ブロックの領域において改良の成果が最も顕著に現れている。しかし、面積が 1 ないし 2 ブロックの領域では、最適化プロセスで動きベクトルのエラーを修正するための隣接情報が少なすぎるため、大きな改良の成果は得られなかったと考えられる。

4.4 その他の考察

まず、本手法では画像上に見えている情報のみから領域の最適化を行っている。すなわち、移動物体が重なることによって見えなくなっている領域は情報がないため、アルゴリズム上は何ら考慮されず、推定等も行われない。したがって、領域分割が正しく行われているかどうかの精度は 4.3 節で示したように、画像上

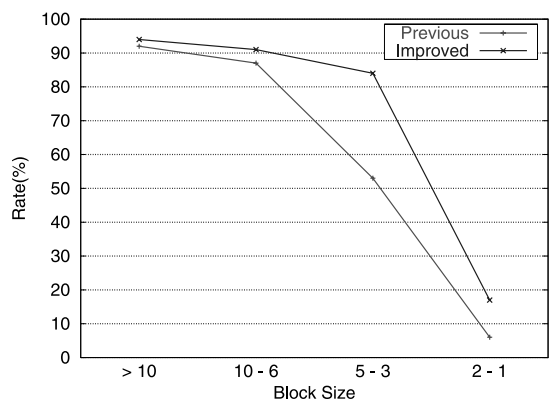


図 10 移動物体面積とトラッキング精度の関係
 Fig. 10 Object area vs. tracking resolution.

に見えている領域の面積には依存するが、重なる領域の面積には依存しない。

次に、本手法ではアルゴリズム上、重なる移動物体の個数は限定していない。しかし、実際の実験では、移動物体の領域境界において接する異なる移動物体は3つまでにとどまり、4つ以上となる場合は無視できるほど起こらない。このとき、実験上、同時に接する移動物体が3つになることで同時に接する移動物体が2つの場合より境界分割精度が低下することはなかった。ただし、「領域分割精度が低下する」とは、同時に接する複数の移動物体のうちある移動物体領域が領域として出力されなくなることを意味する。ただし、移動物体領域が出力されているものの領域の境界がずれる場合については本論文での評価対象としておらず、今後の課題としたい。

最後に、複数の移動物体が重なりながら長い時間並行して移動した場合に、当該複数の物体の領域が1つの領域として合わせるような領域分割ミスが起こりうる。本論文の実験では、物体領域面積が5ブロック以下の場合にこのような領域分割の失敗が現れ、50画像フレーム並行して移動する間に5ブロック以下の領域が並行する隣接物体領域に吸収される確率はおよそ20%であった。これが、図10における領域分割精度の低下の要因の1つとして現れている。逆に物体領域面積が10ブロック以上では、このような現象は無視できる程度にしか起こらない。

5. おわりに

本論文では、2000年に発表された時空間MRFモデルを改良し、車両のような剛体だけでなく歩行者のような非剛体に対してもオクルージョンに対して頑健なトラッキングを行えるようにした。当該改良の効果として、激しいオクルージョンによって小さい面積領域しか画面上に現れなくなった移動物体のトラッキングに関する頑健さも改善された。2000年に発表された時空間MRFモデルは、その後数年にわたり、様々な交通画像に適用した実環境評価実験により実用性が確認されてきている。本論文では、改良後のアルゴリズムに対して車両や歩行者を含む約70分程度の画像を用いて評価を行い、その効果が確認された。今後は、当該改良アルゴリズムに対しても数年にわたる実環境評価実験を行い、その有用性を確認していきたい。

参考文献

1) Kolling, H. and Nagel, H.: 3D Pose Estimation by Directly Matching Polyhedral Models to Gray Value Gradients, *IJCV*, Vol.23, No.3, pp.283-302 (1997).

2) Kolling, H., Nagel, H. and Otte, M.: Association of motion verbs with vehicle movements extracted from dense optical flow fields, *Proc. ECCV'94*, Vol.II, Lecture Notes in Computer Science 801, pp.388-347, Springer-Verlag (1994).

3) Peterfreund, N.: Robust Tracking of Position and Velocity With Kalman Snakes, *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.21, No.6, pp.564-569 (1999).

4) Kass, M., Witkin, A. and Terzopoulos, D.: Snakes: Active contour models, *Int'l J. Computer Vision*, Vol.1, pp.321-331 (1988).

5) Smith, S.M. and Brady, J.M.: ASSET-2: Real-Time Motion Segmentation and Shape Tracking, *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol.17, No.8, pp.814-820 (1995).

6) Stauffer, C. and Grimson, W.E.L.: Adaptive background mixture models for real-time tracking, *Proc. CVPR, 1999*, Vol.II, pp.246-252 (June 1999).

7) Leuck, H. and Nagel, H.-H.: Automatic Differentiation Facilitates OF-Integration into Steering-Angle-Based Road Vehicle Tracking, *IEEE CVPR'99*, pp.360-365 (1999).

8) Gardner, W.F. and Lawton, D.T.: Interactive Model-Based Vehicle Tracking, *IEEE Trans. PAMI*, Vol.18, No.11, pp.1115-1121 (1996).

9) Weber, K. and Malik, J.: Robust Multiple Car Tracking with Occlusion Reasoning, *Proc. ECCV'94*, Vol.I, Lecture Notes in Computer Science 800, Springer-Verlag, pp.189-196 (1994).

10) Paragios, N. and Ramesh, V.: A MRF-based Approach for Real-Time Subway Monitoring, *IEEE CVPR 2001*, pp.I1034-I1040 (2001).

11) Tsaig, Y. and Averbuch, A.: A Region-based MRF Model for Unsupervised Segmentation of Moving Objects in Image Sequences, *Proc. IEEE CVPR 2001*, Vol.I, pp.889-896 (2001).

12) Debrunner, C. and Ahuja, N.: Segmentation and Factorization-Based Motion and Structure Estimation for Long Image Sequence, *IEEE Trans. PAMI*, Vol.20, No.2, pp.206-211 (1998).

13) 羽下哲司, 鷲見和彦: 時間平均シルエットを用いた非剛体移動対象の追跡, 第9回画像センシングシンポジウム講演論文集, pp.265-270.

14) Li, B. and Chellapa, R.: Simultaneous Tracking and Verification via Sequential Posterior Estimation, *Proc. IEEE CVPR 2000*, Vol.II, pp.110-117 (2000).

15) Kornprobst, P. and Medioni, G.: Tracking

- Segmented Objects using Tensor Voting, *Proc. IEEE CVPR 2000*, Vol.II, pp.118–125 (2000).
- 16) Haritaoglu, I. and Flickner, M.: Detection and Tracking of Shopping Groups in Stores, *Proc. IEEE CVPR 2001*, Vol.I, pp.431–438 (2001).
- 17) Sminchisescu, C. and Triggs, B.: Covariance Scaled Sampling for Monocular 3D Body Tracking, *Proc. IEEE CVPR 2001*, Vol.I, pp.447–454 (2001).
- 18) Hidalgo, J.R. and Salembier, P.: Robust Segmentation and Representation of Foreground Key-Regions in Video Sequences, *IEEE ICASSP'01*, Vol.III, pp.1565–1568 (2001).
- 19) Izquierdo, E. and Ghanbari, M.: Motion-Based Object Segmentation in Scale-Space, *IEEE ICASSP'99*, Vol.VI, pp.3473–3476 (1999).
- 20) Geman, S. and Geman, D.: Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of images, *IEEE Trans. PAMI*, Vol.6, No.6, pp.721–741 (1984).
- 21) Chellappa, R. and Jain, A.: *Markov Random Fields: Theory and Application*, Academic Press (1993).
- 22) Majunath, B.S. and Chellappa, R.: Unsupervised Texture Segmentation Using Markov Random Field Models, *IEEE Trans. PAMI*, Vol.13, No.5, pp.478–482 (1991).
- 23) Andrey, P. and Tarroux, P.: Unsupervised Segmentation of Markov Radom Field Modeled Textured Images Using Selectionist Relaxation, *IEEE Trans. PAMI*, Vol.20, No.3 (1998).
- 24) 上條俊介, 松下康之, 池内克史, 坂内正夫: 時空間 Markov Random Filed モデルによる隠れにロバストなトラッキングアルゴリズム, 電子情報通信学会論文誌 D-II, Vol.J83-D-II, No.12, pp.2597–2609 (2000).
- 25) Kamijo, S., Matsushita, Y., Ikeuchi, K. and Sakauchi, M.: Traffic Monitoring and Accident Detection at Intersections, *IEEE Trans. ITS*, Vol.1, No.2, pp.108–118 (2000).
- 26) Kamijo, S. and Sakauchi, M.: Illumination Invariant and Occlusion Robust Vehicle Tracking

by Spatio-Temporal MRF Model, *9th World Congress on ITS*, Chicago (Oct. 2002). CD-ROM.

(平成 16 年 3 月 29 日受付)

(平成 16 年 9 月 11 日採録)

(担当編集委員 長尾 健司)



上條 俊介 (正会員)

1990 年東京大学理学部物理学科卒業。1992 年同大学大学院理学系研究科修士課程物理学専攻修了。1992 年 4 月より富士通株式会社にてプロセッサの開発に従事。1998 年 4 月東京大学大学院工学系研究科博士課程入学, 情報工学を専攻し 2001 年 3 月同修了。博士(工学)。2001 年 4 月東京大学生産技術研究所講師, 2002 年 12 月同助教授。主に ITS における画像認識の応用の研究に従事。IEEE, ACM, 電子情報通信学会等会員。



坂内 正夫 (正会員)

1969 年東京大学工学部電気工学科卒業。1975 年同大学大学院工学系研究科博士課程修了。工学博士。1975 年同大学工学部電気工学科専任講師。その後, 横浜国立大学工学部情報工学科助教授, 東京大学生産技術研究所助教授を経て, 1988 年同大学生産技術研究所教授, 1994 年同大学概念情報処理工学研究センター長, 1998 年 4 月より同大学生産技術研究所所長に就任(2002 年 3 月まで)。2002 年 7 月から国立情報学研究所副所長, 東京大学生産技術研究所教授併任。専門, マルチメディア情報処理。日本工学会アカデミー, IEEE, 電子情報通信学会, GIS 学会各会員。