

背景情報を活用した フレームアウトにロバストな動画像追跡

縄田 竜志[†] 生駒 哲一[†] 河野 英昭[†] 前田 博[†]

ビジュアルトラッキングでは通常、カメラレンジからのフレームアウトが生じた場合、追跡を中断する。本研究では、背景から推定した地理情報を用いることで、一時的なフレームアウトであればフレーム外のターゲット位置を予測し追跡を継続する方法を提案する。シーンは、自律飛行無人ヘリによる自動車の追跡を想定する。動画像からターゲットの速度と姿勢及び背景の移動量をパーティクルフィルタにより逐次推定し、追跡を行う。ターゲットは色で尤度を求め、背景移動量は特徴点のオプティカルフローで尤度を求める。地理情報として道路領域と建物領域を利用し、背景移動量とエッジ情報と色情報より道路・建物領域を検出する。検出された道路領域から、未知部分の道路を3次スプラインで外挿する。ターゲットの観測モデルに自動車は道路上を走行している確率が高いという尤度を加えることで、フレームから消失したターゲットを予測する。シミュレーション実験で提案法の有用性を示す。

Robust visual tracking using background information in dynamic image including out-of-field-of-view

Ryuji Nawata[†] and Norikazu Ikoma[†]
and Hideaki Kawano[†] and Hiroshi Maeda[†]

We propose a method of robust visual tracking in dynamic image including out-of-field-of-view. Visual tracking is usually unable to track target coming out range of a camera. But proposed method is able to continue tracking target in such case by taking into account background information. A scene is Unmanned Aerial Vehicle is tracking a car on a road. State of the car is velocity and posture. And state of background is quantity of translation movement and Angle of rotation. The both state sequentially is estimated from an image sequence by particle filter. A likelihood of the car calculated based on a color information. And, a likelihood of background calculated based on an optical flow. The background information is road and building. There information is extracted based on a color and edge. Unknown road are extrapolated from known road with cubic splines. Target disappearing from the frame is predicted by observation model which imposes prior knowledge that the car normally moves on the road. Effectiveness of

proposed method has been shown through tracking experiments with simulation image sequences.

1. はじめに

ビジュアルトラッキングは、動画像中の移動対象物を追跡する技術であり、防犯カメラによる人物自動追跡・異常検知（システムプロダクト社）や顔認識・追跡機能を搭載したハンディカム（パナソニック株式会社）など様々な製品で利用されている。また、ビジュアルトラッキングはロボットや医学など様々な分野で利用可能であり研究が盛んに行われている。

動的状況で動画を撮影すると、前景物による一時的な「隠れ」や画像から一旦出てまた戻る「フレームアウト」が生じる。ビジュアルトラッキングにおいて、隠れ問題は様々なアプローチで盛んに研究が行われている。一方、フレームアウトとは撮影対象がカメラレンジ内にいない状況であるので、画像中にない物体は追跡できないのが普通である。しかし一時的なフレームアウトであれば、フレームアウト前の追跡対象の情報と時々刻々と画像から観測される背景情報を利用して追跡を継続できると考える。例えば、子供の運動会を撮影していて子供がフレームアウトした場合を考える。フレームアウト後も子供はトラック内を走っていると考えられ(背景情報)、子供が時計周りにトラックを走っていたならば、フレームアウト後も時計周りに走っていると考えられる(追跡対象の情報)。隠れやフレームアウトにロバストなビジュアルトラッキング技術は、対象がカメラに写っていないからという条件を受けないので、トラッキングの自由度が高くなる。

鈴木ら[1]は、人物頭部追跡においてレンジセンサから得られた3次元形状より環境モデルを構築し、室内空間における人物頭部の存在可能性を考え、安定した追跡を行っている。背景情報を利用した動画像追跡は、撮影対象が観測可能な状態での追跡精度の向上も見込める。また、背景情報を利用することは背景を推定する必要があるので、近年盛んに研究されている航空画像や衛星画像からの地形情報推定とも関連する研究であり応用可能であると考えられる。

本研究では、背景情報を積極的に活用したビジュアルトラッキング方法を提案し、一時的なフレームアウトを生じる動画像に対応する。

扱うシーンは、自律飛行無人ヘリコプター(以下、無人ヘリ)による地上を走行する自動車の追跡とする。これは、無人ヘリの姿勢が頻繁に変化するため、フレームアウ

[†]九州工業大学
Kyushu Institute of Technology

トを多々生じること、現存の無人ヘリが指定した飛行経路を自律飛行する物であり、指定しない自律飛行が可能ならば便利であり、それには、フレームアウトにロバストな追跡が必要だからである。

動画像において自動車を追跡するには、時々刻々と観測される画像 \mathbf{I} から自動車の状態 \mathbf{x} の事後分布 $p(\mathbf{x} | \mathbf{I})$ を逐次推定する必要がある。自動車の状態を状態空間モデルで表現し、パーティクルフィルタを用いて逐次状態推定して自動車を追跡する。パーティクルフィルタとは、状態の事後分布を多数の重み付粒子で近似し、観測値が得られる度に粒子を更新する逐次アルゴリズムであり、動的システムの状態推定に向いているフィルタである。

隠れやフレームアウトの問題は、センサから対象情報を観測できないことである。対象の状態を観測できなければ、追跡対象の状態の観測過程を表現した観測モデルから尤度を計算することができない。そこで、背景情報から間接的に追跡対象の状態を観測できる観測モデルを提案する。

まず、背景から道路形状に着目する。自動車は大抵道路上を走行していると想定できるならば、自動車は道路の続く方向へ姿勢を向けていると仮定できる。道路形状を利用することで、自動車の姿勢を間接的に観測することができる。ここで、道路形状は極端に変化しないと考えられるので、画像中の道路形状から画像外の道路形状を予測できる。これによりフレーム内外での自動車の姿勢を観測ことができ、自動車の姿勢に関する尤度がフレーム内外で計算できる。

次に、背景上での自動車の存在可能性を考える。自動車は大抵道路上を走行していると想定できるので、道路上に存在している可能性が高い。逆に建物などが建っている場所では存在可能性が低い。それ以外の場所は、存在する可能性もある。よって背景を道路、建物、それ以外の3項目に分類することで、画像上でターゲットの存在確率分布を生成できる。ここで、地形状況は極端に変化しないと仮定すると、画像中の背景から画像外の背景を補外できる。これにより、画像外へ存在確率分布を拡張でき、フレーム内外で存在確率を計算できる。自動車の存在確率は自動車の位置と関係するので、自動車の位置に関する尤度がフレーム内外で計算できる。

次の章から、まず大枠の追跡アルゴリズムであるパーティクルフィルタによる動画像追跡を説明する。その後、道路形状を考慮した状態空間モデルを提案し、シミュレーション実験で背景を利用することの有用性を示す。次に、道路形状と背景上の存在確率による観測モデルを提案し、シミュレーション実験で有用性を示す。

2. パーティクルフィルタによる動画像追跡

パーティクルフィルタは、状態空間中の多数の粒子を用い、対象の事後分布を粒子の密集の度合いによって近似的に推定する方法である。この章では、パーティクルフィルタのアルゴリズムを一般的なパーティクルフィルタによるターゲット追跡で示す。

時刻 k の自動車の状態は位置 x_k, y_k 、速度 V_k 、姿勢 θ_k とする。この状態をまとめて状態ベクトル $\mathbf{x}_k = \{x_k, y_k, V_k, \theta_k\}$ で表す。観測データは画像データ \mathbf{y}_k である。

追跡対象の状態の事後分布 $p(\mathbf{x}_k | \mathbf{y}_{1:k})$ を、 N 個の重み付き粒子 $\{\{\mathbf{x}_{k|k}^{(n)}, \pi_k^{(n)}\}\}_{n=1}^N$ で近似する。これを逐次的に更新することで、状態推定を行う。ここで $\pi_k^{(n)}$ は正規化重みであり、

$$\pi_i^{(n)} \geq 0, \sum_{n=1}^N \pi_k^{(n)} = 1 \quad (2.1)$$

を満たす。この逐次的更新は次の通り行う。

1 時刻前の重み付き粒子 $\{\{\mathbf{x}_{k|k-1}^{(n)}, \pi_{k-1}^{(n)}\}\}_{n=1}^N$ は分かっているものとする。まず、1 期先予測粒子 $\{\mathbf{x}_{k|k-1}^{(n)}\}_{n=1}^N$ を生成する。この新しい粒子は、状態の時間的変化を表現するシステムモデルを用いて

$$\mathbf{x}_{k|k-1}^{(n)} \sim f(\mathbf{x}_k | \mathbf{x}_{k-1|k-1}^{(n)}) \quad (2.2)$$

と生成する。自動車の動きを考えると、短時間では速度と姿勢がほぼ変化せず、一定と仮定できるので、速度と姿勢をランダムウォークモデルで表現する。自動車のシステムモデルは、

$$\begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} x_{k-1} \\ y_{k-1} \end{bmatrix} + \begin{bmatrix} V_{k-1} \cos \theta_{k-1} \\ V_{k-1} \sin \theta_{k-1} \end{bmatrix} \quad (2.3)$$

$$V_k = V_{k-1} + v_V \quad v_V \sim N(0, \tau_V^2) \quad (2.4)$$

$$\theta_k = \theta_{k-1} + v_\theta \quad v_\theta \sim N(0, \tau_\theta^2) \quad (2.5)$$

と表現する. v_V と v_θ は, システムノイズと呼ばれ, 位置と速度の不確か性を表現する. 生成した (重み付き) 粒子 $\{(\mathbf{x}_{k|k-1}^{(n)}, \pi_{k-1}^{(n)})\}_{n=1}^N$ が, 1期先予測の分布 $p(\mathbf{x}_k | \mathbf{x}_{k-1})$ を近似している. 次に粒子の重みを計算する. これは, 尤度に比例する値

$$\pi_k^{(n)} \propto h(\mathbf{y}_k | \mathbf{x}_{k|k-1}^{(n)}) \quad (2.6)$$

をとるようにする (適宜正規化を行う). 自動車の特徴量をHS色ヒストグラムとする. 観測画像 \mathbf{y}_k 中から位置 (x_k, y_k) を中心とする矩形領域の画像の特徴量 $q(\mathbf{x}_k)$ を得る. 予め作成しておいた自動車のテンプレート色ヒストグラム p とのヒストグラムの類似度を求める. 類似度の計算は Bhattacharyya 係数を使用して,

$$L = \sqrt{pq(\mathbf{x}_k)} \quad (2.7)$$

となり, 類似度より自動車の特徴量による尤度は,

$$p(\mathbf{y}_k | \mathbf{x}_k; p) \propto \exp\{k(1-L)\} \quad (2.8)$$

となる.

こうして得られた粒子と重みの組 $\{(\mathbf{x}_{k|k-1}^{(n)}, \pi_k^{(n)})\}_{n=1}^N$ が, ろ波分布 $p(\mathbf{x}_k | \mathbf{y}_{1:k})$ を近似している. リサンプリングでは, 重み $\pi_k^{(n)}$ に比例する確率で, 粒子を復元抽出する. すなわち,

$$\mathbf{x}_{k|k}^{(n)} = \begin{cases} \mathbf{x}_{k|k-1}^{(1)} & \text{with prop } \pi_k^{(1)} \\ \vdots & \vdots \\ \mathbf{x}_{k|k-1}^{(N)} & \text{with prop } \pi_k^{(N)} \end{cases} \quad (2.9)$$

とする. また重みは均一にする. すなわち $\pi_k^{(n)} = 1/N$. リサンプリング後の粒子の

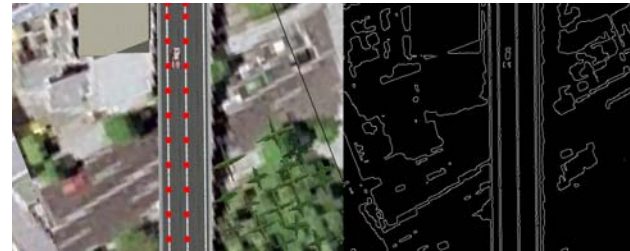
集合 $\{\mathbf{x}_{k|k}^{(n)}\}_{n=1}^N$ も, ろ波分布 $p(\mathbf{x}_k | \mathbf{y}_{1:k})$ を近似している[2].

3. 道路形状を考慮した動画像追跡

ここでは, まず背景の道路に着目し, 背景情報を活用した動画像追跡が有用であることを検証する. 画像中から道路形状を推定し, 道路形状を考慮した状態空間モデルを提案する. まず, 道路形状推定方法を示し, 次に道路形状を考慮した状態空間モデルを構築する.

3.1 道路形状認識

エッジ画像では図1のように自動車の進行方向に対して垂直方向近傍で白線が強いエッジとして検出される. 検出したエッジを追跡し, 白線を数個の座標として保持する.



これらの座標より道路形状を定式化することで形状を認識する. 道路の座標を $\{(rx^{(i)}, ry^{(i)})\}_{i=1}^n$ とすると区間 $[ry^{(i)}, ry^{(i+1)}]$ での3次スプライン関数の区分多項式は

$$S^{(i)}(y) = a^{(i)}(y - ry^{(i)})^3 + b^{(i)}(y - ry^{(i)})^2 + c^{(i)}(y - ry^{(i)}) + d^{(i)} \quad (3.1)$$

となる. 道路形状 \mathbf{R} は

$$\mathbf{R} = \left\{ \begin{array}{l} \text{左}(S^{(1)}, S^{(2)}, \dots, S^{(n-1)}) \\ \text{右}(S^{(1)}, S^{(2)}, \dots, S^{(n-1)}) \end{array} \right\} \quad (3.2)$$

となる. へりの操作量 $\mathbf{hx} = \{hx, hy, h\theta\}$ (並進運動と回転運動) が, センサより取得できるとすると, 次のフレームでの道路の位置は, 前フレームの道路の位置をへりの操作量より遷移させ求める. 新しく観測される道路領域は, パーティクルフィルタを用いて推定する. 道路の白線状態 $\mathbf{rx} = \{rx, ry, R\}$ を位置と曲率と取る. システム

モデルは、スプライン関数 \mathbf{R} より、

$$rx_k = S_k^{(n-1)}(ry_k) \quad (3.3)$$

$$ry_k = ry_{k-1} + hx_k \quad (3.4)$$

$$R_k = R_{k-1} + v_R \quad v_R \sim N(0, \tau_R^2) \quad (3.5)$$

と表現する。 v_R はシステムノイズであり、曲率の不確実性を表現する。システムモデルより予測した白線の位置と曲率より新しく観測される白線形状 $\tilde{S}_k^{(n)}$ を計算する。白線形状と画像中の近傍エッジとの距離の最小 2 乗誤差を計算する。

$$d = (\tilde{S}_k^{(n)} - \mathbf{G})^2 \quad (3.6)$$

観測モデルは、

$$h(\mathbf{G} | \mathbf{rx}) \propto \exp(-(d)^2) \quad (3.7)$$

と表現する。パーティクルの重み w_k は

$$w_k = h(\mathbf{G} | \mathbf{rx}) \quad (3.8)$$

より計算される。

3.2 道路形状を考慮した状態空間モデル

第 2 章で、自動車の動きを表現したシステムモデルと自動車の特徴量による観測モデルを示した。ここでは、その状態空間モデルを、道路形状を考慮した状態空間モデルへ拡張し、フレームアウトに対応できることを示す。

自動車とは、大抵道路上を走行していると仮定できるので、この事前知識をシステムモデルに考慮する。道路上を走行しているということは、自動車の姿勢は道路の進行方向とほぼ同じであると仮定できる。自動車の位置 (x_k, y_k) での道路の進行方向

$r\theta_k$ は、道路形状 \mathbf{R}_k より

$$r\theta_k = \tan^{-1} \frac{x_k}{S_k^{(i)}(y_k)} \quad (3.9)$$

と定義できる。自動車の姿勢を表現した(2.5)式を

$$\theta_k = r\theta_k + v'_\theta \quad (3.10)$$

と拡張する。これは、自動車の時刻 k での姿勢は、時刻 k での道路の進行方向と同じであるというモデルである。ただし、必ず同じであることはないので、姿勢のシステムノイズ $v'_\theta \sim N(0, \tau_\theta'^2)$ で、姿勢の不確実性を表現する。

次に観測モデルを拡張する。フレームアウト状況では、自動車の特徴量が観測できないため尤度は小さくなり、時間経過と共に粒子が拡散する。同様に自動車は大抵道路上を走行していると過程すると自動車の位置は道路に近い位置に観測される。道路形状 \mathbf{R} より道路形状による観測モデルは、

$$h(\mathbf{I}_k | \mathbf{x}_k) \propto \exp\{-(x_k - S_k^{(i)}(y_k))\} \quad (3.11)$$

と表現する。(3.11)式は、道路中心に近い位置にいるほど自動車であるというモデルである。(2.8)式と(3.11)式は、異なる評価関数であるので、観測モデルは、

$$p(\mathbf{I}_k | \mathbf{x}_k) \propto \alpha p(\mathbf{I}_k | \mathbf{x}_k; p) + (1-\alpha)p(\mathbf{I}_k | \mathbf{x}_k) \quad (3.12)$$

となり、混合分布モデルで統合する。

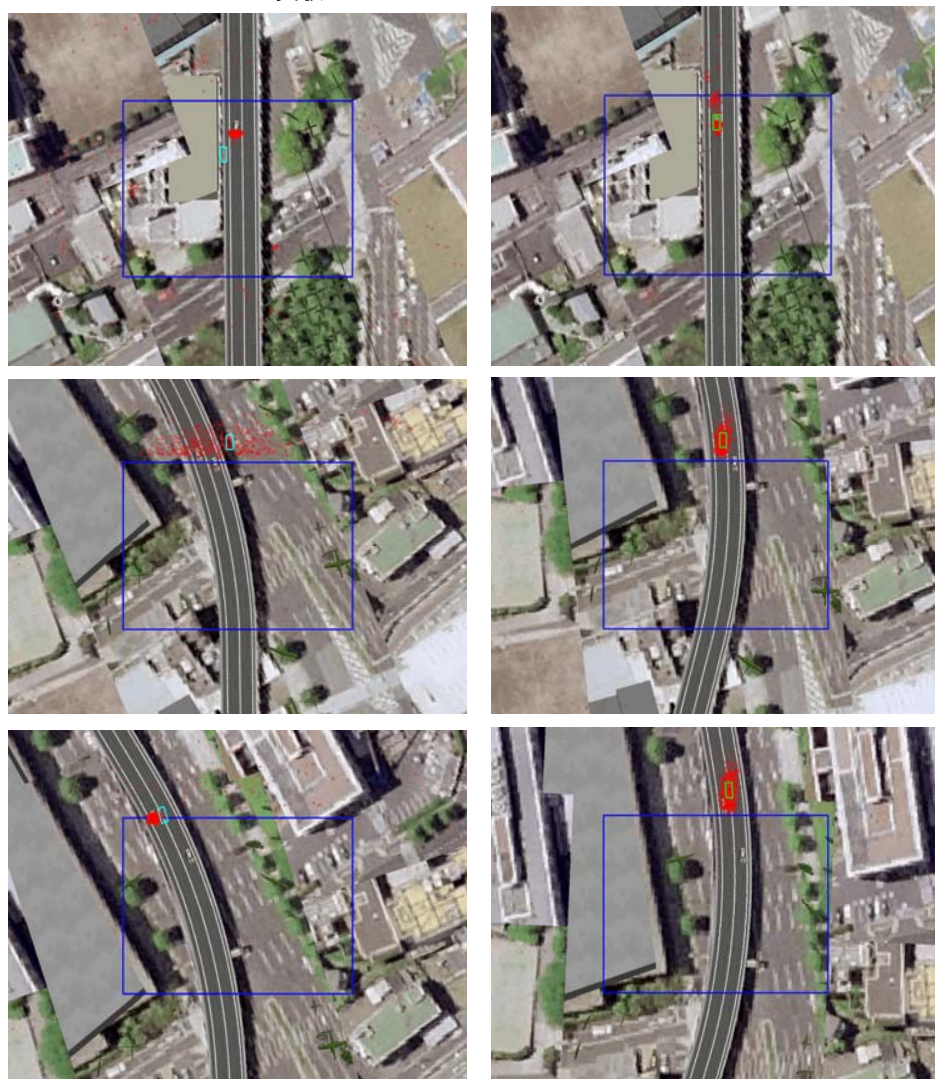
道路形状はフレーム内外で形状が大きく変化しないと考えられるので、推定した道路形状 \mathbf{R} より、フレーム外の道路形状を 3 次スプライン外挿で表現する。これにより、フレーム外の道路形状は観測できることになる。問題は、(2.8)式の自動車の観測モデルの扱い方である。画像の縁付近では、粒子は次の 3 種類の状況が生じる。①完全フレームアウト：自動車の特徴量観測不可 ②部分フレームアウト：自動車の特徴量部分観測 ③フレームアウトなし：自動車の特徴量完全観測。この 3 種類の状況の粒子を極力不公平なく扱うことが望ましい。そこで平均色ヒストグラム $q(\mathbf{I}_k)$ を導入する。

平均色ヒストグラムとは、時刻 k での観測画像 \mathbf{I}_k を色ヒストグラム化する。①の場合には、平均色ヒストグラムとの類似度

$$L = \sqrt{(pq(\mathbf{I}_k))} \quad (3.13)$$

を求める。②、③の場合は、(2.7)式で類似度を求める。ただし、②の場合は、観測されない領域のピクセルを観測されている領域から補間する。

3.3 シミュレーション実験



(a) 道路形状考慮しないモデル

(b)道路形状考慮したモデル

図1. シミュレーション実験結果

第2章の道路形状を考慮しない状態空間モデルと3.2節で提案した道路形状を考慮した状態空間モデルでシミュレーション実験を行った。OpenGLでシミュレーション空間を作製し、自動車は指定した走行経路を走行する。ヘリは、自動車の推定値から自動車が画像中心に写るような簡単な制御で飛行する。それぞれのモデルでの10, 65, 80フレームを図1に示す。青枠内が観測画像で、枠外はフレームアウト状況をみやすくするため表示しており、実際には観測されていない。

フレームアウト状況で道路形状を考慮しない場合は、粒子が拡散しており自動車を把握できていないが、提案法では粒子の近傍に多数の粒子が存在し自動車を把握できている。フレームアウトした後自動車が曲がっているが、提案法ではヘリが追尾していることから自動車の追跡を継続できていると考えられる。動画像において背景情報を活用することは有用であることが示された。

4. 背景情報を活用した動画追跡

前章までで、背景を活用する動画追跡の有用性を示した。ここでは、背景中の道路形状だけでなく、背景上での自動車の存在確率を新たに提案し、より現実に近い状況でのフレームアウトを生じる動画像に対応する。前章では、ヘリの移動量をただしく取得可能としたが、センサ値には普通ノイズが加わっており、そもそもセンサを搭載していない場合も考えられる。そこで、背景の移動量を画像から推定する。

まず、背景上の自動車の存在確率分布を求める方法を述べ、次に背景移動量の推定方法を述べる。最後に背景情報を活用した観測モデルを提案し、実験結果を示す。

4.1 背景上の自動車の存在確率分布

画像中の背景には、道路形状以外にももっと多くのパロキシカルな情報が含まれている。例えば、ターゲット以外の自動車はターゲットの移動経路に影響を及ぼす。道路領域は自動車の存在確率が高い領域であり、建物領域は自動車の存在できない領域であり、それ以外の領域は自動車が道路をはみ出して走行することもあるので、自動車が存在する可能性のある領域である。ここでは、道路と建物とそれ以外領域に着目し、背景を3項目に領域分類する方法を示す。

まず、観測画像 I_k を教師ありピクセル分類法により分類する。分類項目を道路領域と建物領域とそれ以外領域とし、道路領域とそれ以外領域についての典型的な統計量を求める。建物は、さまざまな色が存在するため典型的な統計量を求めない。バンドは各画素のR, G, Bの3バンドとするとそれぞれの統計量は、

$$r\mu = \{r\mu_1, r\mu_2, r\mu_3\} \quad (4.1)$$

$$rS = \{rS_1, rS_2, rS_3\} \quad (4.2)$$

$$\mathbf{o}\mu = \{o\mu_1, o\mu_2, o\mu_3\} \quad (4.3)$$

$$\mathbf{o}s = \{os_1, os_2, os_3\} \quad (4.4)$$

となる。これは、それぞれの分類項目の統計量を各バンドの平均 μ と分散 s で表している。分類したい画素の値をベクトル \mathbf{x} で $\mathbf{x} = \{x_r, x_g, x_b\}$ と表すと、分類したい画素 \mathbf{x} と道路項目の平均及びそれ以外項目の平均との標準ユークリッド距離は、

$$\mathbf{rd}^2 = (\mathbf{x} - \mathbf{r}\mu)^t \cdot (\mathbf{x} - \mathbf{r}\mu) \quad (4.5)$$

$$\mathbf{od}^2 = (\mathbf{x} - \mathbf{o}\mu)^t \cdot (\mathbf{x} - \mathbf{o}\mu) \quad (4.6)$$

となる。どの項目に分類するかは、標準ユークリッド距離と統計量の分散から

$$\text{if}(\mathbf{rd}^2 < \mathbf{rs}) \quad \mathbf{x} = \{0,0,255\} \quad (4.7)$$

$$\text{if}(\mathbf{od}^2 < \mathbf{os}) \quad \mathbf{x} = \{0,255,0\} \quad (4.8)$$

$$\text{if}(\mathbf{rd}^2 > \mathbf{rs} \cup \mathbf{od}^2 > \mathbf{os} \cup (\mathbf{rd}^2 < \mathbf{rs} \cap \mathbf{od}^2 > \mathbf{os})) \quad \mathbf{x} = \{255,0,0\} \quad (4.9)$$

より決定する。(4.7)式が道路領域、(4.8)式がそれ以外領域、(4.9)式が建物領域である。ピクセル分類をフレームが取得される度に行う。分類する前の画像を図2に示し、分類した結果画像を図3に示す。図3では、道路の中央分離帯や歩道、自動車が建物領域として誤検出されている。ただし、誤検出が多く発生しているのは、道路領域近傍であるので、道路形状で補正する。道路形状 \mathbf{R} 領域内の画素は道路領域のピクセルであるとして、ピクセル分類した画像を補正する。また道路領域近傍の建物領域(歩道)も道路領域として、補正する。道路形状を検出した結果画像を図4に示し、補正した結果画像を図5に示す。

ピクセルの輝度値によって存在確率 p を、

$$\begin{aligned} \text{if}(R == 255) \quad p &= 0.1; \\ \text{if}(G == 255) \quad p &= 0.8; \\ \text{if}(B == 255) \quad p &= 1.0; \end{aligned} \quad (4.9)$$

と定義する。



図2 原画像

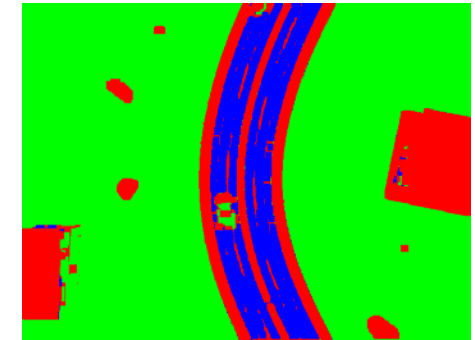


図3 ピクセル分類結果画像



図4 道路形状検出結果

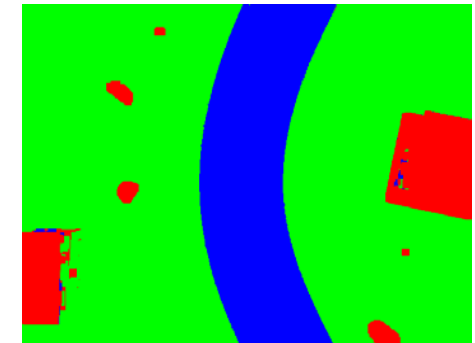


図5 補正した画像

4.2 背景状態推定

時刻 k での背景の状態 $\mathbf{hx}_k = \{hV_{xk}, hV_{yk}, hw_k\}$ を画像上での x 方向移動量, y 方向移動量, 回転角とする。背景の状態空間モデルを構築し、パーティクルフィルタを用いて背景状態推定を行う。推定した背景状態は、カメラの動きと等価である。

システムモデルでは、背景の移動量と回転角の時間変化を表す。システムモデルは、

$$hV_{xk} = hV_{xk-1} + v_{hVx} \quad (4.10)$$

$$hV_{yk} = hV_{yk-1} + v_{hVy} \quad (4.11)$$

$$hw_k = hw_{k-1} + v_{hw} \quad (4.12)$$

と表現する．移動量と回転角は，1時刻前とほぼ同じであるというランダムウォークモデルで表現する．移動量と回転角の不確実性を，移動量のシステムノイズ $v_{V_x} \sim N(0, \tau_{V_x}^2)$ ， $v_{V_y} \sim N(0, \tau_{V_y}^2)$ と回転角のシステムノイズ $v_{hw} \sim N(0, \tau_{hw}^2)$ で表す．

背景の状態 \mathbf{hx} は，画像中から直接には観測できないが，オプティカルフローとして間接的には観測が可能である．画像からのオプティカルフロー推定では，Kanade-Lucas-Tomasi(KLT)法[4]などの濃度勾配に基づく方法が，計算量の観点から有利で高速に動作する．安定なオプティカルフローを抽出するには，信頼できる画像特徴を抽出することも重要である．画像中の特徴点は harris corner detector[5]で検出する．ただし，画像の縁付近に検出された特徴点は，特徴点群から除く．ピラミッド型KLT法により，時刻 k での特徴点群 $\mathbf{x}^{(i)}$ のオプティカルフロー $\mathbf{V}^{(i)} = \{u, v\}$ が画像から観測される．オプティカルフローとカメラの動きの関係を表した方程式は，

$$u = -\frac{v_x}{Z} - w_y - w_z y + x\left(\frac{v_z}{Z} + w_x y - w_y x\right) \quad (4.13)$$

$$v = -\frac{v_y}{Z} - w_z x + w_x + y\left(\frac{v_z}{Z} + w_x y - w_y x\right) \quad (4.14)$$

である．本研究で考慮する回転角は w_z だけで， v_z は考慮せず，さらに特徴点は画像座標系であり，ターゲットの状態も画像座標系であるので(4.12)式，(4.13)式は，

$$u = -v_x + w_x Y \quad (4.17)$$

$$v = -v_y + w_x X \quad (4.18)$$

と書ける．(4.16)式，(4.17)式は，画像上のある特徴点のオプティカルフローが背景状態によって発生する過程を表現している．背景状態によって計算されたオプティカルフローを $\mathbf{V}(\mathbf{hx})$ とすると観測モデルは，

$$p(\mathbf{V} | \mathbf{hx}) \propto \exp(-\sqrt{(\mathbf{V} - \mathbf{V}(\mathbf{hx}))}) \quad (4.18)$$

と定義する．

4.3 背景情報を活用した状態空間モデル

ここでは，推定した背景状態と抽出した背景情報を活用した状態空間モデルを提案する．ターゲット自動車のシステムモデルは，

$$\begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} \cos(h\hat{\theta}_k) & -\sin(h\hat{\theta}_k) & hx_k \\ \sin(h\hat{\theta}_k) & \cos(h\hat{\theta}_k) & hy_k \end{bmatrix} \begin{bmatrix} x_{k-1} + V_{k-1} \cos(\theta_k) \\ y_{k-1} + V_{k-1} \sin(\theta_k) \\ 1 \end{bmatrix} \quad (4.19)$$

$$\theta_k = \theta_{k-1} - h\hat{\theta}_k + v^\theta \quad (4.20)$$

$$V_k = V_{k-1} + v^V \quad (4.21)$$

と表現する．(4.19)式は現時刻の位置は，1時刻前の位置から力学モデルにより遷移し，推定した背景の状態より現時刻の画像座標の位置へ遷移することを表現する．(4.20)式は，現時刻の姿勢は1時刻前の姿勢とほぼ同じであるランダムウォークで表現し，カメラの回転を考慮したモデルとなっている．(4.21)式は，現時刻の速度は，1時刻前の速度とほぼ同じであるランダムウォークで表現する．姿勢と速度にシステムノイズを加え姿勢と位置に関する不確実性を表現する．

次に，背景情報を活用した観測モデルについて述べる．色ヒストグラムによる観測モデルは，(2.8)式と同様である．4.1節で抽出した背景情報は，領域分類画像，道路形状の2つである．

まず，領域分類画像の活用方法を示す．背景は画像の縁とフレーム外とで情報が極端に変化するとは考えにくい．そこで領域分類した画像からフレーム外の建物領域，それ以外領域を画像の縁のピクセルより補外する．道路領域については，抽出した道路形状よりフレーム外の道路領域を3次スプライン外挿する．これにより，フレーム外の粒子についても道路情報を観測することができるので，フレームアウトに対応できる．粒子が予測した位置 $(\tilde{x}_k, \tilde{y}_k)$ を中心とする矩形を考える．矩形で考えるのは，点で考えた場合，本当は道路上の点であるのにノイズなどで偶然に建物領域と誤検出したピクセルから間違った存在確率を計算してしまうのを防ぐためである．矩形内の各ピクセルの存在確率を計算し，平均を取った値を存在確率 $p(\mathbf{x}_k)$ とする．

次に，道路形状 \mathbf{R} の活用方法を提案する．第3章では，道路形状をシステムモデル

に考慮していたが、ここでは観測モデルに考慮する。粒子が予測した位置 $(\tilde{x}_k, \tilde{y}_k)$ の道路進行方向 $r\theta$ は、(3.9)式より求まる。よってターゲットの姿勢に関する観測モデルは、

$$p(\mathbf{R} | \mathbf{x}) \propto \exp(-(\theta - r\theta)^2) \quad (4.22)$$

で定義される。よってターゲットの観測モデルは、

$$p(\mathbf{I} | \mathbf{x}) \propto p(\mathbf{I} | \mathbf{x}; p) * p(\mathbf{R} | \mathbf{x}) * p(\mathbf{x}) \quad (4.23)$$

と表現する。(4.23)式は、各観測モデルを同時分布モデルで統合している。

4.4 シミュレーション実験

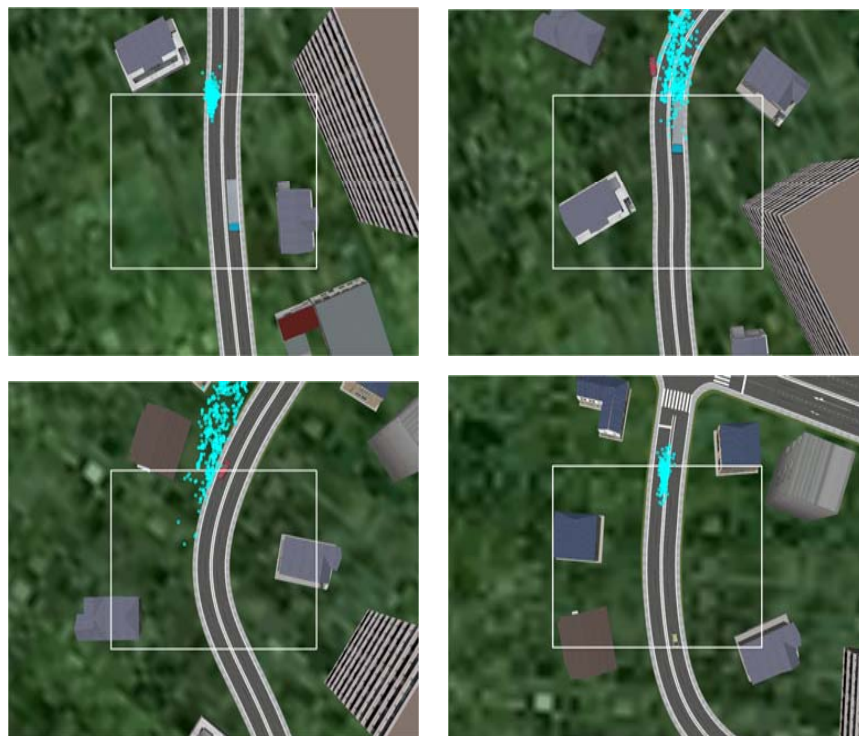


図7 シミュレーション実験結果

提案した状態空間モデルでフレームアウトが生じる動画像におけるターゲット追跡実験を行った。FORAM8社のドライビングシミュレータで、交差点のある道路や複数の自動車、複数の建物があるより現実に近い町を作製した(図7)。ターゲット自動車とヘリは、定義した移動経路を走行及び飛行する。ヘリの移動量は、取得できないとする。ターゲット以外の自動車はドライビングシミュレータのランダム走行機能でランダムに走行する。40, 80, 120, 160 フレームの結果を図6に示す。白枠内が実際の観測画像で、枠外はフレームアウトの状況を見やすくするため表示しているだけで実際は観測されていない。

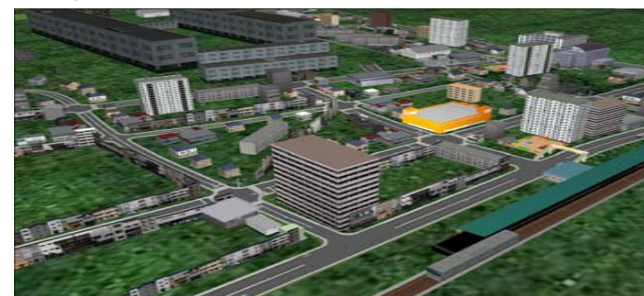


図6 シミュレーション空間

自動車がフレームアウトした後もある程度追跡を継続できているが、自動車が道路からはみ出た場合に追跡が失敗している。

5. おわりに

背景情報を活用した動画像追跡方法を提案し、フレームアウトが生じる動画像で自動車の追跡を行った。背景情報を積極的に活用することで、フレームアウトした自動車を追跡することができた。

今後の課題として、実画像での実験や様々なシーンでの提案手法の適用が挙げられる。

謝辞

本研究は、統計数理研究所共同利用研究(20-共研-4103)による

参考文献

- 1) Michael Isard and Andrew Blacke: "CONDENSATION-conditional density propagation for visual tracking" Int.J.Computer Vision, 29, 1, 5-28,(1998)
- 2) 鈴木達也, 岩崎慎介, 小林貴訓, 佐藤洋一, 杉本晃宏, 「環境モデルの導入による人物追跡の安定化」, 電子情報通信学会論文誌 DII, vol.J88-DII no.8, pp.1592-1600, 2005.