

平成 22 年 5 月 6 日現在

研究種目：若手研究（B）
 研究期間：2007～2009
 課題番号：19700174
 研究課題名（和文）シミュレーション型動画画像解析のためのモンテカルロモデリング法の構築
 研究課題名（英文）Monte Carlo Modeling for Simulation-Based Video Analysis
 研究代表者
 川本 一彦（KAWAMOTO KAZUHIKO）
 千葉大学・総合メディア基盤センター・准教授
 研究者番号：30345376

研究成果の概要（和文）：計算機シミュレーションは、計算機内で実世界の現象を模擬する技術であり、とくにシミュレーションに確率的な要素を組み込んで、試行のたびに異なる結果を出力するようにしたものをモンテカルロシミュレーションと呼ぶ。本研究では、この方法を活用して、実世界を映した動画画像とそれを模擬するシミュレーションモデルを比較する手法について研究し、物体トラッキング、カメラ運動の推定、および人の行動解析などに応用した。

研究成果の概要（英文）：Computer simulation is a method for modeling real-world phenomena on computers. In computer simulation, a stochastic modeling approach is called *Monte Carlo simulation*, which includes repeated random sampling to compute results. In this research project, we applied Monte Carlo simulation to computer vision applications. In particular we developed a method for combining Monte Carlo simulation and observations captured by a video camera and applied the method

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
平成 19 年度	800,000	0	800,000
平成 20 年度	1,500,000	450,000	1,950,000
平成 21 年度	800,000	240,000	1,040,000
年度			
年度			
総計	3,100,000	690,000	3,790,000

研究分野：コンピュータビジョン

科研費の分科・細目：情報学・知覚情報処理・知能ロボティクス

キーワード：逐次モンテカルロ法，粒子フィルタ，統計的時系列解析，数値シミュレーション

1. 研究開始当初の背景

原子物理学や地球物理学など自然科学分野では、シミュレーション技法を理論構築および実験に続く第3の解析手法として位置づけ、広く活用されている。このシミュレーション技法を動画画像解析でも同様に発展させ、この分野における強力な接近法を構築す

ることは、次世代の動画画像解析のための基盤技術として有効である。とくに、動画画像解析の諸課題の多くは、確率統計モデルで定式化されるため、確率統計モデルを組み込んだモンテカルロシミュレーション技法を応用することが有効であると期待できる。

そこで、モンテカルロシミュレーション技

法をもとに、横断的な分野からさまざまな手法を積極的に導入する。例えば、統計科学の分野では、地球物理学における複雑な相互作用を持つ現象（津波など）の計算機シミュレーションのために、データ同化へ逐次モンテカルロ法が適用され始めている。データ同化は、観測データをシミュレーション数値モデルに当てはめる方法であり、いわゆる(パラメトリックな)モデル当てはめの概念を数値モデルへ拡張した方法である。このような試みをさらに拡張し、現在までに得られているさまざまな物理・幾何モデルを包含するハイブリッドな状態空間モデルを構成し、そのシミュレーション型数値計算法を発展させる。数値シミュレーションという古典的な方法は、近年のパーソナルコンピュータの処理能力を最大限に生かすことで、個人ユーザーレベルではじめて可能になると考えられる。

2. 研究の目的

モンテカルロシミュレーション技法と確率統計的なモデリング法を組み合わせた方法をモンテカルロモデリング法と呼び、この方法論を動画像解析の諸問題へ応用することを目的とする。

動画像解析・認識の方法として、(1)物理モデル型、(2)事例・学習型、そして(3)ヒューリスティック型、等がある。(1)では、投影過程を精密な物理モデルで記述し、その逆問題を解くという定式化を基本とする。(2)では、それとは反対に、大量のデータから帰納的に推論規則を導く。(3)では、経験則等からアルゴリズムを設計する。モンテカルロモデリング法の学術的な特色は、これらの方法を柔軟に組み合わせることができるところにある。例えば、ある確率分布からのサンプリングという操作を手続き的な操作（例えば事例から導かれた推論規則）に置き換え、物理モデルと事例モデルを組み合わせたハイブリッドな状態空間モデルを構築することも可能になる。さらに、学習器を組み合わせることで適応的な学習型モデルも利用できる。さらに、このように構成した状態空間では、勾配型の数値最適化は適用することが困難になり、さらに発見的な方法でも定量的な信頼性評価が困難である。しかし、このような解析的な取り扱いが困難空間でも、モンテカルロモデリング法のシミュレーション型数値計算による組織的な方法によって計算評価が可能になる。

また、本研究によって、モデル構築そして検証というフィードバックサイクルを効果的に進められるようになる。モンテカルロモデリング法では、原理的には、ほとんど全てのモデルに対して計算を進めることができる。したがって、線形性・正規性といった制約を外すことができ、さまざまなモデルを利用で

き、「実際に計算できる」ため検証を効果的に進めることができる。このような方法論は、動画像解析分野における新しい計算パラダイムとして学術的な価値も高い。

3. 研究の方法

本研究では、モンテカルロモデリング法のための関連するソフトウェアライブラリの開発および基本アルゴリズムの設計のほか、以下の具体的な課題を取り上げる。

- (1) 適応的な可変モデルによる高精度な物体追跡技術の開発、
- (2) 超パラメータの自動決定による超解像処理技術の開発、
- (3) カメラ運動の逐次推定
- (4) 人の行動解析のための数値シミュレーションと動画像計測の融合

まず、主要設備として、高精細・高速カメラを導入し、パーソナルコンピュータ上で画像計測システムを構築する。このシステムで撮影した映像は、動画像の高解像度化に対するベンチマークデータとして利用する。さらに、高解像度化のための数値シミュレーションモデルをパーソナルコンピュータ上に実装し検証を進める。

研究代表者はすでに時系列フィルタリングのための統計的モデルに基づく逐次モンテカルロ法を実装している経験を有している。すでに開発したソースコードの一部などを基にして、数値シミュレーションモデルの実装を進める。平行して、高精細・高速カメラにより動画像を収集し、統計的モデル構築のための予備実験を繰り返す。さらに、統計的モデルでは捕らえきれない色パターン変化を他の非線形手法を使ってモデル化するための実験蓄積を行う。

4. 研究成果

主要な研究成果について個別に報告する。

- (1) 適応的な可変モデルによる高精度な物体追跡技術の開発

動画像の物体追跡において、追跡対象の運動のモデル化は、追跡アルゴリズムの性能を決める重要な作業である。もし運動に関して事前に知識がなければ、「滑らかに動く」という仮定を要請することが多い。この仮定は汎用的で多くの場合上手くいくが、急激な運動変化など想定外の運動に対しては適切にモデル化できず、追跡が失敗してしまいやすい。一般に、どのようなモデルを採用しても、そのモデルに対する想定外の運動が発生する可能性はあり、それが原因で追跡失敗が起こりうる。なぜならば、追跡対象を事前に設定した仮定だけに基づいて盲目的に探索し

ても、実際の対象をとらえられるかどうかは分からないからである。

そこで、適応的な運動モデルとして、オプティカルフロー駆動型運動モデルを提案し、粒子フィルタで実装する。このモデルでは、連続する2画像間の疎なオプティカルフローから運動変化を推定し、その推定値によって状態を「駆動」させる。オプティカルフローという付加情報を用いて状態推定するため、「滑らかな運動」という知識だけを用いた運動モデルに比べて、適応的に運動をモデル化できる。また、オプティカルフローは特定の物体に固有な知識ではなく局所的な濃淡パターンだけから計算できるため、広いクラスの追跡対象を扱うことができる。さらに、粒子フィルタの推定精度や効率性に影響を与えるパラメータである分散を自動的に調整する方法を導入する。分散は多くの粒子フィルタでは手動で経験的に設定されるが、この手間を自動化することができる。

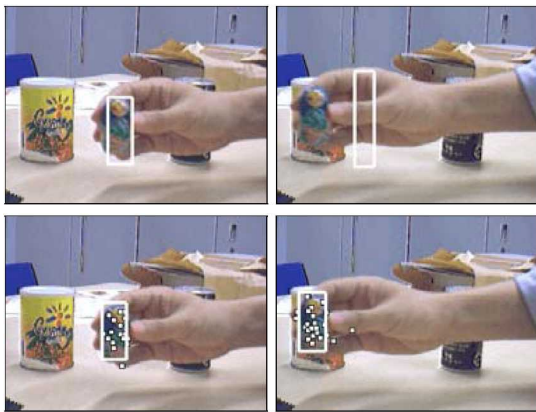


図 1

図 1 の上段は従来手法による物体追跡の一例で、下段は提案手法による結果である。この例で示すように、従来法では困難であった急激な運動変化に対して、提案手法は高精度に追跡が達成できている。

(2) 超パラメータの自動決定による超解像処理技術の開発

同一対象をずらしながら撮影した複数の低解像度画像から高解像度画像を推定する処理を一般に超解像処理と呼ぶ。カメラの高解像度化が進んでいる現在では、超解像処理は一見必要ない技術とも思えるが、例えば安価なカメラに対する画像認識の前処理、防犯カメラの高解像度化、画像位置合わせによる広範囲空間の高精細ディスプレイへの表示など広い応用がある。また、超解像処理を信号の補間というより一般的な立場で考えれば、普遍的な課題であり理論的にも興味深い。

超解像処理に対する手法は多く提案されているが、大きく分けると最尤推定(あるいはガウス雑音を仮定した最小二乗推定)と最

大事後確率推定の枠組みに分類できるものが多い。前者は非ベイズ的に問題を扱い、後者は事前知識を入れてベイズ的に問題を扱っていると考えられることができる。

ここでは、後者の最大事後確率推定の枠組みで超解像処理を定式化する。このときどのような事前分布を導入するかがまず問題となる。広く用いられている空間的な2階差分が0(空間的な滑らかさを要求することに相当)を期待する事前分布を用いることにする。本研究の貢献は、このような事前分布を決めるパラメータの自動決定法の提案にある。事前分布に含まれるパラメータは超パラメータと呼ばれる。超パラメータは、尤度(データへの当てはまり項)と事前分布(罰金項)をバランスする役割を持ち、その値は最終的な結果に大きく影響を与える。この正則化パラメータの確率モデルをさらに導入した階層ベイズで超解像処理をモデル化し、ABIC(赤池ベイズ情報量基準)でその最適値をデータから決定する。これにより、手動で試行錯誤的に決定する手間を省くことができるようになる。

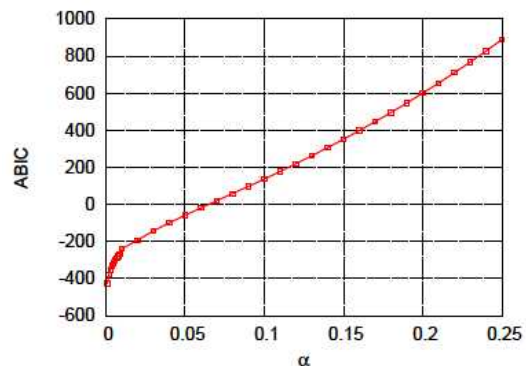


図 2

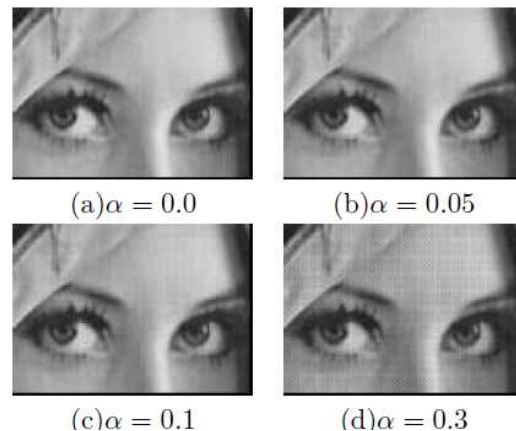


図 3

図 2 に正則化パラメータ α をグリッド探索した結果を示す。この例では、ABIC の値は単調増加する結果が得られている。図 3 に $\alpha = 0.0, 0.05, 0.1, 0.3$ に対応する復元画像を示す。

(3) カメラ運動の逐次推定

動画像からのカメラ運動の推定はコンピュータビジョンの最も基本的な課題であり、すでに多くの研究成果がある。これらの手法の多くは、基本的には、2 画像間の特徴点対応から基礎行列あるいは基本行列を推定し、それを回転行列と並進ベクトルに分解する手順を取る。

ここでは本稿では、このような従来の計算手順とは逆に、まず回転と並進を求め、それから基本行列を構成する手順を取る手法を提案する。提案手法では、基本行列の回転 3 自由度を単位回転軸の球面座標とその周りの回転角度および単位並進 2 自由度を球面座標でパラメータ化し、それらの時間変化を陽に表現する確率モデルを導入する。この確率モデルでは、連続する画像間での運動の時間的な滑らかさを仮定し、この仮定に基づき過去の運動履歴から現時刻での運動をベイズ的に推定する。このような滑らかさの仮定により、特徴点が少ない(極端には 5 点未満)ときや空間点とカメラの配置が臨界曲面上に乗っている(あるいはそれに近い)ときでも、推定値の計算が破綻することはない。なぜならば、過去の推定値から滑らかに変化したという知識がいわゆる正則化項の役割を果たし、パラメータが一意に決まらない不定の状況でも対処できるためである。さらに、このパラメータ化は幾何学的な意味が明確なため確率モデルをつくるうえで直観が働きやすく、そしてカメラ運動に対する事前知識も組み込みやすい利点もある。一方で、運動パラメータのゆる度をエピ極線と特徴点の間の幾何学的距離で評価するため、非線形な観測モデルが現れ解析的な取り扱いが難しくなる。そこで、数値近似アルゴリズムとして粒子フィルタを導入し、カメラ運動パラメータの事後分布をモンテカルロ近似する。さらに、実環境での利用を想定して、いくつかの工夫を導入している。

図 4 にシミュレーション動画像(3 次元空間中の適当な領域に 3 次元点をランダムに 200 点生成し、カメラが動きながらその投影像を取得する模擬環境を設定)を用いた結果を示す。図 4 上段は、推定した回転角度と真の回転角度の絶対角度誤差を示している。図 4 下段は、並進方向の角度誤差を示している。それぞれ、モデルに含まれる分散の値を変えて、3 つの結果を描画している。

図 4 上段の結果は、高精度に推定できているが、その理由は、回転角度は、 0° 付近に存在するはずだ、という強い事前知識を組み込むことができ、事前分布にしたがって真値の周辺に粒子を初期時刻から生成することができるからである。図 4 下段の結果は、とくに最初の 6 フレーム目程度までは、大きな推定誤差が出ている。これは、並進方向について

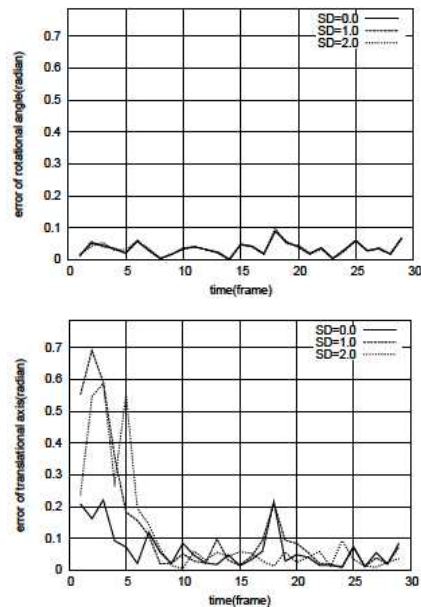


図 4



図 5

は、定義域全体の一様分布を初期分布としたため、強い事前知識を課さなかったためと考えられる。すなわち、真値周辺以外にも、エピ極線と特徴点の距離が比較的小さくなるような局所解などにも粒子が存在することになっているためと考えられる。しかし、時間の経過とともに収束していき、回転軸と同程度の推定誤差に落ち着いていく。

図 5 は、手持ちカメラから撮影した実動画像を用いた結果である。図中の直線は、カメラ運動パラメータの推定値から構成した基本行列をもとに前時刻の特徴点から引いたエピ極線である。提案手法の利点は、このようにエピ極線が正確に 1 点で交わるような解を出力できるところにある。この結果は、実際のカメラ運動と対応が取れていることを示している。

(4) 人の行動解析のための数値シミュレーションと動画像計測の融合

人の動きを予測し解析する技術は、歩行者

の行動意図推測、災害時での避難シミュレーション、交差点監視カメラや車載カメラによる事故分析・予防など非常に多くの応用分野を持つ技術課題である。

自由意志を持った人間の動きを予測することは非常に困難な課題である。1つのアプローチとして、自由意志を捨象した行動規範モデルを導入し、そのモデルに基づく数値シミュレーションモデルを構築することである。すなわち、計算機上で仮想世界を構築し、その中で未来の人の位置や速度を計算する発想である。このアプローチは大きく分けて2つある。1つはマクロな視点に立ったアプローチで、例えば、人の流れを流体や気体として扱いナビエ-ストークス方程式やボルツマン方程式で記述する方法である。しかし、このアプローチはあまりにも規範的すぎ、個々の人を没个性的に扱いきるきらいがある。効果を発揮するのは、夏のお祭りや初詣、または建物内での非常事態での非難時のように非常に混雑した雑踏の中で自由意志が実際ほぼ奪われているような状況に限られると思われる。一方、個々の人をもう少し个性的に扱うアプローチとして、各人にミクロ的な規範モデルを課す方法がある。このミクロモデルでは、個々の人はそれぞれ自身の目的地を持ち、他人や障害物との相互作用を受けながら進路が計算される。

本研究では、このようなミクロ型歩行数値シミュレーションモデルと画像計測技術を結合し、仮想世界と現実世界をつなぐ方法を提案する。このアイデアは、気象学や海洋学ではデータ同化と呼ばれている。データ同化では、仮想世界での出来事を現実世界のデータに基づいて評価する。すなわち、現実世界から乖離したシミュレーションパスを除き、現実世界をよく記述するシミュレーションパスを残す。データ同化の利点は、数値シミュレーションモデルに含まれる各種パラメータ（初期・境界条件など）や不確実性をデータに基づいて合理的に推定できることにある。人行動シミュレーションモデルでは、例えば、人の速度、目的地の位置、そして他人・障害物との反発や誘引の強さを決定するパラメータの設定に自由度がある。現実世界では個々の人で異なるのが通常であり、これらの値に応じてさまざまなシミュレーションパスが生成され、さまざまな未来が描かれる。平均的な歩行者パラメータというものを想定しても、必ずしも現実世界をよく記述できるとは限らない。提案モデリング法では、まず、これらのパラメータの時間発展を確率分布で表現し、それらの確率分布に従って、多数の順方向のシミュレーションパスを生成する。次に、それらのシミュレーションパスで計算された仮想的な人の位置と画像計測で求めた現実の人位置との差を評価する。

このとき、その差が小さいシミュレーションパスほど高評価を与える。最後に、この評価値に基づいて、シミュレーションパスを復元抽出し、高評価のパスを多数複製し、低評価のパスを消滅させていく。この手順は、一般状態空間モデルとその逐次モンテカルロ近似解法である粒子フィルタに基づく。さらに、このモデリングの応用として、人と人の間の関係を解析する課題を取り上げる。この課題では、人と人の間に仮想的な力を導入し、その力を隠れ変数として推定する。

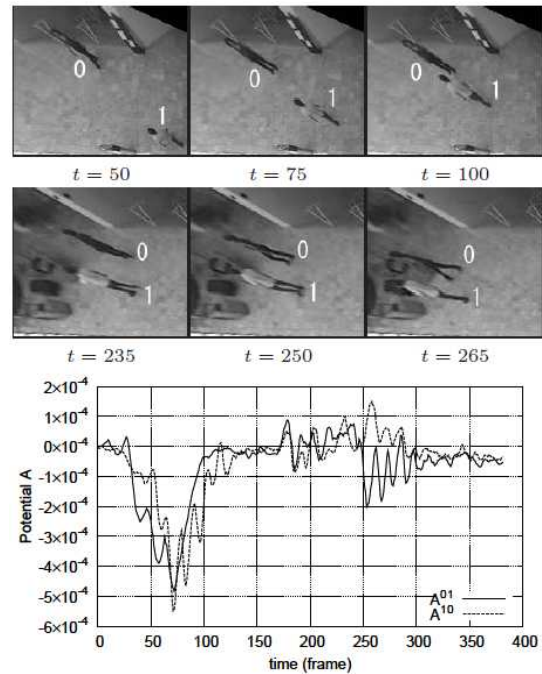


図 6

図 6 に、歩行者 2 人が登場する動画画像を用いて、歩行者間の反発・誘引効果を検証した結果を示している。図 6 上段は、動画画像の一部である。この動画画像では、2 人の歩行者がそれぞれ右下と左上から登場し、それぞれ歩み寄って行き、最終的には左方向へ一緒に歩いていく。歩行者 2 人を区別するために、画像中にそれぞれ番号(0 と 1) を割り当てている。以下、それぞれ 0 番と 1 番と呼ぶ。図 6 下段のグラフに、反発・誘引係数の平均推定値を時間方向にプロットしたものを示す。図 6 下段を見ると、 $t=0 \sim 100$ では誘引・反発係数はともに負の値を取っている。この結果は、両者がお互いに引きつけられていることを示している。図 6 上段の画像と比較すると、実際に両者が近づいていることが確認できる。また、 $t=250$ 付近から、0 番が 1 番から受ける係数値が負の値、その逆が正の値を取り、両者の符号が逆になる箇所が出てくる。実際の動きでは、0 番は左上に行こうとしていたが ($t=235$)、1 番が左下に向かおうとしたため、0 番は 1 番の後を追うように左下に方向転換している ($t=250, 265$)。すなわ

ち, 0 番が 1 番の後を追うという行動が誘引を引き起こしている. 実際には, 1 番が進路決定の主導権を握っているかどうかはこの実験では不明であるが, 少なくとも 1 番が先に歩き 0 番が追いかけるという状況が, 比較的早い段階で認識できることを示唆している.

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 6 件)

川本一彦, 運動パラメータの直接サンプリングによる逐次カメラ運動推定, 電子情報通信学会和文論文誌 D, Vol.J93-D, No.8, 2010 (採録決定)

Kazuhiko Kawamoto, Random Scanning Algorithm for Tracking Curves in Binary Image Sequences, International Journal of Intelligent Computing in Medical Sciences and Image Processing, Vol.2, No.2, 2009, 101-110

Kazuhiko Kawamoto, Hierarchical Bayesian Modeling for Estimating Shared Hidden States with Application to Tracking, Journal of Advanced Computational Intelligence & Intelligent Informatics, Vol. 13, No.2, 2009, pp.80-85

Kazuhiko Kawamoto, Optical Flow Driven Motion Model with Automatic Variance Adjustment for Adaptive Tracking, Lecture Notes in Computer Science, Vol.4843, 2007, 555-564

[学会発表](計 14 件)

川本一彦, 超解像処理のためのベイズ型情報量基準に基づく正則化パラメータの自動決定, ファジィシステムシンポジウム, 2009

Kazuhiko Kawamoto, Particle Filtering and its Applications to Computer Vision, International Seminar of Information Technology, 2009

川本一彦, 人流シミュレーションモデルと画像計測の融合, 電気学会一般産業研究会, 2008

川本一彦, オプティカルフロー駆動型運動モデルによる対象領域追跡, 画像の認識・理解シンポジウム, 2007

6. 研究組織

(1)研究代表者

川本 一彦 (KAWAMOTO KAZUHIKO)

千葉大学・総合メディア基盤センター・准教授

研究者番号: 30345376