

3次元物体の形状と運動認識のための尤度付き 物理シミュレータの提案

ダニアル・ケオキ^{*1} 青木孝文^{*2} 戸田真志^{*3} 長谷川晶一^{*1}
^{*1}:電気通信大学 ^{*2}:東京工業大学 ^{*3}:公立はこだて未来大学

概要：カメラ画像から実世界の3次元情報を復元することが求められている。本研究では、実世界とCGを重畳させるために必要な3次元形状と運動認識をより効果的に行うために尤度付き物理シミュレータの提案をする。

Probability considered physic simulator for increasing recognition of three-dimensional shapes and movements of objects

Daniel Keoki^{*1} Takafumi Aoki^{*2} Masashi Toda^{*3} Shoichi Hasegawa^{*1}

^{*1}:The University of Electro-Communications ^{*3}:Future University-Hakodate

^{*2}:Tokyo Institute of Technology Precision and Intelligence Laboratory

Abstract: A research to increase recognition of shape and movement for three-dimensional objects using physic simulator with probability consideration. By this, object recognition via camera is hoped to increase.

1. 研究背景

近年、コンピュータの普及によりコンピュータ上で3DCG環境を構築することができるようになった。また、コンピュータの周辺機器であるカメラの普及とともに、複合現実感環境の構築までできるようになった。複合現実感環境とは、人が直接知覚できる環境に対して、コンピュータグラフィックスによって更に情報を追加し、実世界を拡張するという概念である。例えば、カメラ画像の一部にコンピュータグラフィックスを重畳させることで複合現実環境を構築することができる。

しかし、複合現実感環境を違和感なく実現するには実世界の情報をコンピュータに教える

必要がある。例えば、実世界の物体の形状情報を把握することで、その物体の真上にCGを重畳することが可能になる。

CyberCode[1]はカメラ画像を二値化し、事前に登録したパターンとビット列を比較することで2次元コードの認識をしている。この認識と同時に2次元コードの空間情報を得ている。これにより、コンピュータグラフィックスを正しく実世界に表示できるようになった。

ActiveCyberCode[4]はこれに加えて、コードの領域を拡張することで、この拡張領域に対する変化で複合現実感環境内のコンピュータグラフィックスオブジェクトの向きを変化させたり、オブジェクトを変えたりすることなど、

コンピュータグラフィックスの操作ができるようになった。

複合現実感環境でインタラクションをさらに可能にしたのは ARToolkit[2] という複合現実感環境の構築ツールである。ARToolkit の開発でコンピュータとカメラを持っていれば、手軽に複合現実感環境の構築ができるようになった。ARToolkit もカメラを使って、あらかじめ登録したマーカをカメラ画像から抽出し、そのマーカの位置や姿勢などを得ている。そしてカメラ画像上にコンピュータグラフィックスを重畳し、複合現実感環境を構築している。

また CyberCode を利用して複合現実感環境を用いた THE EYE OF JUDGMENT というトレーディングカードゲームが近年発売された。トレーディングゲームに刻まれた CyberCode をリアルタイムに読み取り、そしてバーチャルなモンスターをカード上に表示し、他プレイヤーのモンスターと戦わせて、決められた 3×3 マスを奪い合うゲームである。

しかし CyberCode や ARToolkit では、カメラでコードやマーカを認識する必要がある。そのため、マーカをカメラで認識できなかった際、コンピュータグラフィックスをカメラ画像上に重畳することができない。例えば、マーカの一部でも手などで隠してしまうとコンピュータグラフィックスが表示できなくなってしまう。その結果、コンピュータグラフィックスとインタラクションする際に、マーカの認識を邪魔しないように注意をする必要があるので、インタラクションの自由度が減ってしまう問題がある。また、マーカを設置する必要があることから複合現実感環境の構築にコストがかかり、実世界の景観を損ねる可能性もある。

一方で、マーカ認識を用いない方法として SLAM という方法が存在します。SLAM とは Simultaneous localization and Mapping と

う略で、未知の環境での地図の作成を行うと同時に自分の存在する位置を把握できるような探索用ロボットや自律ロボットに多く使われる技法である。主にロボットの周りに LRF (Laser Range Finder) などの外部センサを使って未知の環境を作成する手法である。そして、エンコーダーを使用して、作成した環境内での位置を推定する。SLAM の良い点はもちろん同時に未知の環境を作成することと同時に自分の位置を把握できる点だが、センサで検出した情報に誤差が含まれる場合に、誤差が蓄積する問題があり、正確な地図の作成あるいは、自分の位置を把握することができなくなる。また外部情報を知るためにたくさんのセンサを用いることが必要になり、膨大な計算量になることが多いことから、オンラインよりもオフラインでデータの処理を行う場合が多い。

しかし、近年カメラ画像を使った SLAM ができるようになりつつあり、Visual SLAM (vSLAM) と呼ばれる。vSLAM はカメラを用いて、カメラ画像内の特徴点を 3 次元的にマッピングとトラッキングすることで、フレーム間のカメラの相対位置を推定することができる。その中でリアルタイムに SLAM を可能にしたのは PTAM[3] という画像ベース SLAM である。この研究ではリアルタイムにカメラ画像中の特徴点を追跡とマッピングできることから複合現実感提示に利用できる。しかし、SLAM 特有の問題点として運動している特徴点に対して弱いことがあげられる。またカメラを用いることで、速い動きに対してカメラ画像中の特徴点を抽出できない問題が存在する。

カメラ画像からの物体形状と運動の認識が難しい。これは、フレーム間での複数特徴点集合の対応が困難であることがあげられる。具体的には、特徴点の位置に誤差が含まれることと特徴点が現れたり消えたりすることから不安

定である。解決方法として、特徴点の位置を何度も測定し、その平均をとる方法である。このことにより、特徴点の誤差と不安定性が減らせる。これは、物体が静止している場合に、カメラ画像中の特徴点の平均位置で物体の位置を認識できるが、物体が運動している場合には、単なるカメラ画像中の特徴点の平均位置とすることができない。また、運動している物体に対して特徴点が不安定なため、フレーム間での幾何的拘束対応や速度制約判定などといった方法がやぶれ、剛体が認識しにくい。

カメラ画像中の特徴点の誤差と不安定性を減らすには、高速なカメラで特徴点のある程度安定にすることができる。しかし、そのような方法では非常にコストが高い。これに対して一般的なカメラは安いですが、測定の際 60fps 程度の計測しか行えない。そこで、運動している物体を一般的なカメラで長時間の計測結果に渡り、平均を取る手法で物体認識が可能なシステムを提案する。

2. 研究目的

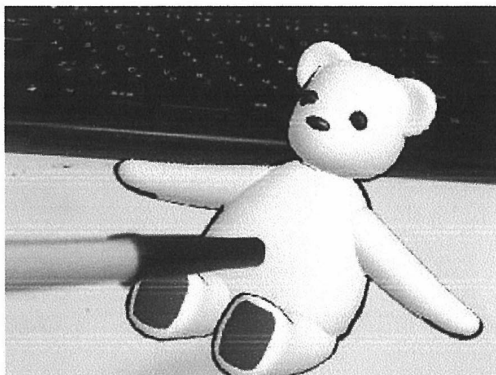


図1：インタラクシオンイメージ図

本研究で実現したいことはマーカを設置することなく複合現実感環境を構築し、動いている物体に対しても正確に CG を重畳すること

である。これによって例えば移動中の車の上を舞台にキャラクタを登場させ格闘技やシューティングゲームといった実世界指向のコンピュータエンタテインメントの実現が期待できる。

しかし、カメラ画像から動いている物体の形状と運動の認識が難しい。これは、フレーム間での複数特徴点集合の対応が困難であることがあげられる。特徴点の位置には観測ノイズが含まれることと特徴点が見えたり消えたりすることから不安定である。そのため、フレーム間での幾何的拘束対応や速度制約判定などといった方法がやぶれ、剛体の認識がしにくい。また、フレーム間での特徴点をマッチングは、特徴点の数が増えれば増えるほど計算時間がかかってしまい、リアルタイムで動作しなくなってしまふ。そのために何らかの動き予測を導入する必要があると考えられる。

本研究では、カメラ画像からの3次元物体の形状と運動認識を行うために、実世界とCGを重畳させるために必要な3次元形状と運動認識をより効果的に行えるために尤度付き物理シミュレータの提案をする。

3. 提案手法

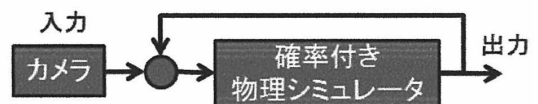


図2：システム構成

図2はシステムの構成を示し、カメラと確率付き物理シミュレータで構成されている。カメラで得られた画像を入力とし、カメラ画像の特徴点をシミュレータに渡す。そして、シミュレータ側に特徴点情報に基づいて、形状・物体と速度の認識を抽出する。これらの情報をシミュレーションすることにより、次の物体の位置や

速度などを予測する。

予測した結果を次にカメラで検出した物体の位置や速度などと比較し、その違いを評価した上で、現在検出した測定値の尤度をつける。予測結果と検出した情報が同じであればあるほど、その検出した情報と予測に使われる情報の尤度が高くなっていく。このことにより、シミュレーションを行うために渡されるデータの信頼が増し、認識精度が向上すると期待される。また、物理シミュレーションによって、次のステップでの特徴点の位置を覚剛体ごとに予測することが可能になるため、特徴点のマッピングの計算コストを削減することもできる。

その他にも、オクルージョン問題やカルマンフィルタで扱えない予測状況が解決できると期待できる。カメラで検出した物体の速度に尤度をつけることで、オクルージョンが発生した場合に、尤度の高い速度値で予測シミュレーションが行えるので、物体の認識が可能と考えている。また、運動物体が別の運動物体とぶつかるような場合に、カルマンフィルタで予測困難な場合でも、物理シミュレータを用いることで、物体同士がぶつかった後の各物体の跳ね返りが認識できると試みている。

3.1. 形状・物体認識

形状と物体認識はカメラ画像上の特徴点のフレーム間での対応で行う。物体認識に関しては特徴点のフレーム間の幾何的拘束条件に基づく手法で行う。これはフレーム間の特徴点でアフィンを求めて、その行列式の値が1に近いかどうかでフレーム間で同じ物体に載っている特徴点かどうか確認している。その他に、特徴点で構成されるベクトルで内積と外積を評価フレーム間で比較する方法をとっている。これをすべてのフレーム間の特徴点に行い、ある特徴点の集合が同じ物体に属していたら、一つ

の物体として認識する。

物体形状に関しては認識した物体の特徴点集合の Convex Hull から構成する。

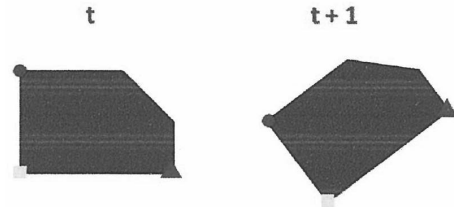


図3：物体拘束条件

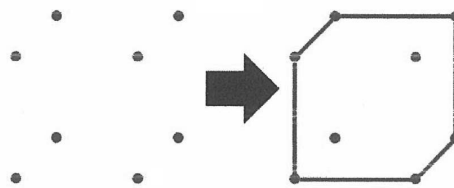


図4：形状認識 (Convex Hull)

3.2. 速度認識

物体速度は物体認識の履歴をみて、物体の速度を推定する。

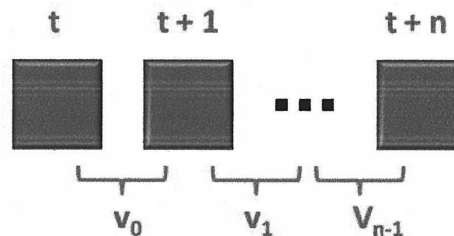


図5：速度認識

3.3. 確率付き物理シミュレータ

物体の運動を推定・予想するために物理シミュレータを利用する。しかし、物体が運動する

場合、カメラで検出した速度と位置に、センシング誤差が含まれるために物体の運動推定や予測を行うことが難しい。そこで、これらを可能にするための確率付き物理シミュレータを提案する。これにより、認識精度の向上が期待される。以下は、具体例を一次元の運動の認識を例に説明する。

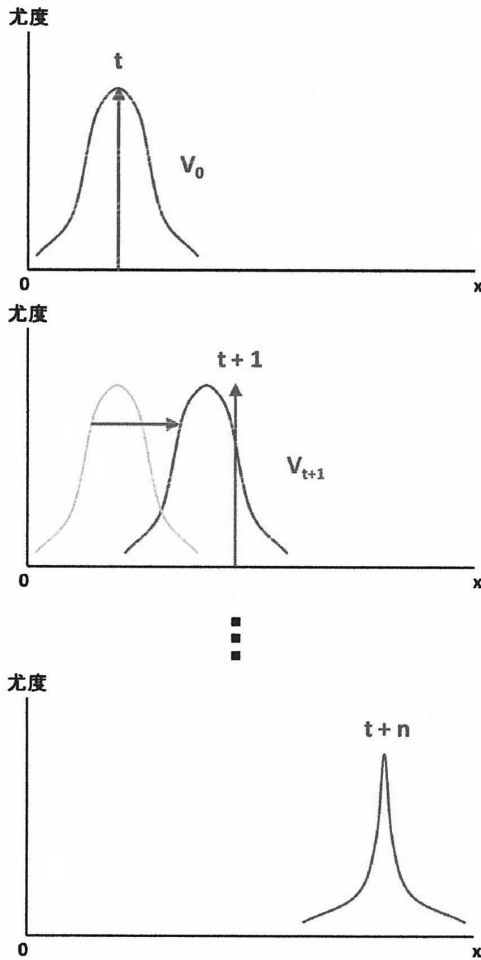


図 6：一次元運動尤度分布

最初はある時刻 t にシミュレータ内の物体の存在分布が図 6 に示しているようになっているとする。そして、その時刻 t に、カメラで

検出した速度がシミュレータに渡されて、シミュレーションによりこの存在分布が移動する。次に時刻 $t+1$ に検出した物体の位置がシミュレータにフィードバックされて尤度分布に投票され尤度分布の処理を行う。ここでの処理例として、平均または重みづけを考えている。入力とシミュレーションを繰り返した結果、最終的に時刻 $t+n$ に確立分布が狭くなり尤度分布の高い速度と位置が出ると考えている。

図 7 は 3 次元グラフで尤度分布、速度と位置の関係を表しているイメージ図である。尤度軸と位置軸から観察すると、時刻 t に物体の位置の尤度分布が広がっている。そしてシミュレーションを重ねた後、この尤度分布が右側へ移動し、時刻 $t+n$ の時に位置に対する尤度分布狭くなり、よりもっともらしい物体の位置が推測できる。

また図 7 に示したように、尤度軸と速度軸から観察すると、物体が等速の場合、速度の尤度分布が移動はしないが、尤度分布が最初は広いが、物体の速度を検出するたびに、尤度分布が狭くなり、もっともらしい物体の運動速度が推測できる。

4. 実験

4.1. 2次元運動と物体認識

物体の拘束条件でカメラ画像中の特徴点を評価して、物体認識を行った。最初にフレーム間のカメラ画像中の特徴点でアフィン変化を求め、そのアフィンの行列式の値を評価してフレーム間の対応を確認した。またフレーム間の特徴点がどの特徴点の集合に属するかを調べ、物体の認識を行った。そして、物体認識歴で物体の運動速度を割り出した。

4.2. 1次元運動：分布作成

簡単なマウスコンソールの位置を検出して、

コンソールの存在位置の尤度分布を作成した。この時、コンソールの位置を検出すると同時に検出ノイズがのるように正規乱数を付加した。

この実験から、検出したマウスコンソールの存在分布が最初は広いが、検出を繰り返した結果、分布が狭くなった。

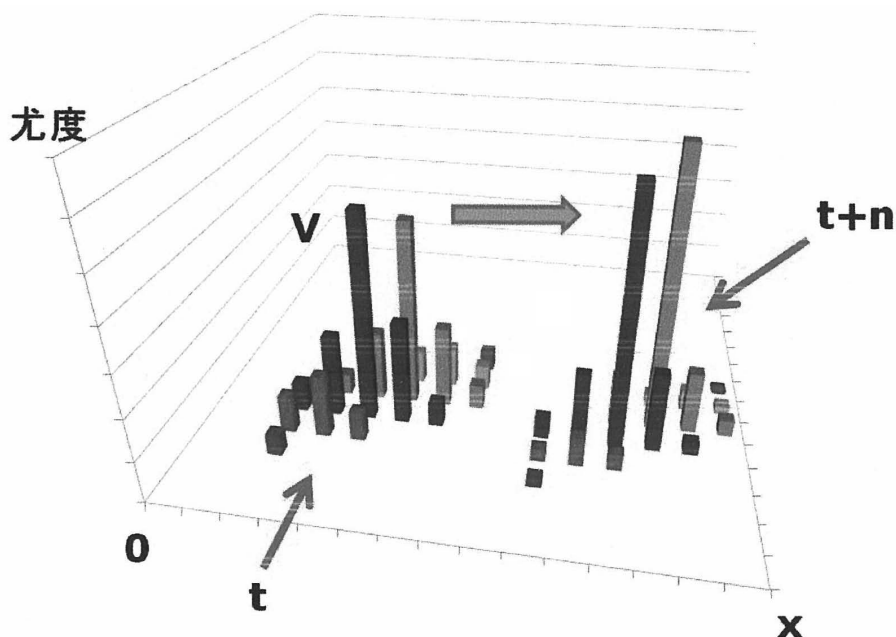


図7：尤度・速度・位置の関係グラフ

5. まとめと今後の課題

本論文では、3次元物体の形状と運動認識のための尤度付き物理シミュレータの提案を行い、その原理および応用性について説明を行った。また、1次元運動確率付きシミュレーションを構築し、提案手法の実現性を確認した。現在は物理付きシミュレータの振舞いやカメラで検出した情報の扱い方などを整えるために1次元等速運動で構想をしている段階である。今後は1次元運動についてのオクルージョン問題について評価実験を行い、

その後、2次元運動認識への拡張を行い、簡単な小さな複合現実感環境でのシステム動作を確認する。最終的には、3次元の運動認識を行い、実世界を舞台とした複合現実感環境を用

いて、ゲームキャラクターを提示し、そのゲームキャラクターとインタラクションを行えるようなシステムの構築を目指す予定である。

【参考文献】

- [1] Jun Rekimoto, "Augmented Reality using the 2D matrix code" WISS'96
- [2] Kato, H., Billinghurst, M. Marker Tracking and HMD Calibration for a video-based Augmented Reality Conferencing System. IWAR99.
- [3] Georg Klein and David Murray Parallel Tracking and Mapping for Small AR Workspaces. ISMAR'07
- [4] 綾塚祐二, 暦本純一, Active CyberCode: 直接操作できる二次元コード, WISS 2005