3

ゲームのインタフェース技術の最前線

白鳥 貴亮[†]

キーワード ジェスチャセンサ, モバイルプロジェクト

1 ま え が き

ゲームで使われているインタフェース技術の概要について解説をする。数年前までは十字キーやボタンを用いたゲーム操作が一般的であったが、任天堂社のWiiリモコンやマイクロソフト社のキネクトといった入力デバイスの登場に伴い、プレーヤ自身が身体を動かすことで操作するゲームが多く登場してきている。こういった入力デバイスについてまとめ、これからのゲーム向けインタフェースについて議論を行う。

2 ジェスチャセンサ

現在市販されているゲーム用入力デバイスは大きく分けて、慣性センサを用いるものとカメラなどの光学センサを用いるものの二つに分けられる。

2.1 慣性センサ

慣性センサ(Inertia Measurement Unit: IMU)とは、物理の慣性を利用してジェスチャを検知するセンサのことで、加速度センサやジャイロスコープがこれに相当する。

加速度センサは加速度を計測する慣性センサであり、任天堂社のWiiリモコン(図1)や近年の多くの携帯電話に搭載されている。ここで注意しなければならないのは、加速度センサが計測する加速度は、実際の加速度とは異なる点である。最も単純な加速度センサの原理では、おもりがぶら下がっているバネの伸縮距離を測定している。そのため、加速度センサが静止している状態、すなわち加速度がゼロであっても、重力の影響によってバネがおもりの重さに応じて自然長より伸びた状態にあり、センサは非ゼロの一定値を返し、またセンサが自由落下しているときにゼロの値を返す。この重力の影響はセンサの傾き(センサから見た時に重力がどの方向にかかっているか)によって変わるため、センサが静止、もしくはゆっくり動いている(理論的には等速で動いている)時の重力方向の検出に非常に有用であり、玉ころがしなどのゲームや携帯電話を持ち替えた

この図の著作権は著者に帰属し、学会への譲渡・転載許諾を得ることができませんでしたので削除いたしました。

時の表示変更などによく使われている。一方、プレーヤがセンサを激しく動かした場合には、センサがプレーヤのジェスチャに伴った加速度の影響も受けるため、重力方向の検出が難しくなり、センサ値の振幅の大きさなどジェスチャの激しさを表す指標が有用になる¹⁾。

ジャイロスコープはセンサの向きを計測する慣性センサであり、任天堂社のWiiモーションプラスや近年の携帯電話に搭載されている。ジャイロスコープは物理の角運動量保存則を利用したもので、回転している物体に外力が加わった時に元に戻ろうとする慣性力が生じる。この慣性力を計測することで、センサの角速度を計測することができる。また、加速度センサと組合せることでより精度の高い向きの計測が可能となる。理論的に計測できるのは角速度であるため、センサの向きを計算するには積分をしなければならず、使用前に初期化を行わなければならない。

慣性センサを用いたゲームは、任天堂社のWiiや携帯、タブレット端末を中心として多く開発されている。ゲームの操作としては大きく分けて二つあり、端末をゆっくりと傾けることで重力方向を検知してゲーム内の物体やキャラクターを操作するものと、ユーザが行った複雑なジェスチャをゲーム端末側で認識してゲーム内のキャラクターを操作す

[†] マイクロソフトリサーチアジア

"Recent Advances in User Interfaces for Video Games" by Takaaki Shiratori (Microsoft Research Asia)

るものに分けられる。前者は、前述のとおり容易に算出できるが、後者に関しては、慣性センサはノイズの影響が大きく、センサ位置や向きの精度の高い算出が困難であることから、ジェスチャのパターンをゲームデザインの時点で制限し、サンプルデータを用いて予め学習された認識器を使ってユーザの姿勢やジェスチャを認識するものが多い。慣性センサを用いて複雑なジェスチャを行っているユーザの姿勢を精度よく復元するのは、実用上非常に難しく、現在のゲームで姿勢の復元をしているものは見受けられない。

こういった慣性センサは、近年では基盤としても安価に入手することができ、地磁気センサと組合せたものはサンプルプログラムがついて100ドル前後で購入が可能である。

2.2 光学センサ

光学センサとは、CCDやCMOSなどの撮像素子を利用したセンサのことで、カメラなどがこれにあたる。

カメラを用いたゲーム用入力デバイスとしては、ソニー社のPlayStation Eyeが古くからある。PlayStation Eyeでは、各ピクセルの動き(オプティカルフロー)をゲーム画面上に当てはめることで、ゲーム画面内のキャラクタを操作したり、間接的にインタラクションを行ったりする。また、近年の顔領域の検出や顔認識の技術を応用することで、画面内のプレーヤの位置を検出することができ、より高度なインタラクションを可能としている。また、PlayStation Moveでも、PlayStation Eyeを用いてコントローラ先端部の球を検出し、3次元位置推定に用いられている。

マイクロソフト社のキネクトも、光学センサを用いた入力デバイスであるが、通常のカラーカメラだけでなく赤外カメラと赤外線プロジェクタも搭載しており、赤外カメラとプロジェクタのデータを複合してデプス画像(被写体の形状)を得られるのが特徴的である(図2)。赤外線プロジェクタからは予め用意されたパターン光が投影され、そのパターン光が周辺環境で反射したものをカメラで計測し、

3次元環境を復元する。復元方法は原理的にステレオカメラによる三角測量と同じであるが、ステレオカメラを用いる場合は、周辺環境のテクスチャを基にマッチングを行うのに対し、プロジェクタを用いる場合は、既知のパターン光をプロジェクタから照射し、その計測画像との間でマッチングを行うために、非常に効率良くかつ頑健に3次元復元を行え、最近の流行であるComputational Photographyの一手法である。さらに、3次元復元の後にデプス画像からプレーヤのスケルトンを推定している。キネクトの中には、さまざまな体型のスケルトン付きキャラクタモデルに、さまざまなモーションキャプチャデータを適用して合成したメッシュデータのデータベースがある。センサから得られたデプス画像をクエリーとして、このデータベースから類似した姿勢を検索することで、プレーヤのスケルトンを推定し、キャラクタの操作を可能としている²⁾。さらに、カラー画像から顔検出、認識をする技術も持ちあわせており、PlayStation Eyeと同様により高度なインタラクションも可能となっている³⁾。

キネクトは、ゲームにおいてはモーションキャプチャシステムとして主に使われており、慣性センサを用いた場合と比べて、精度の高いジェスチャの復元が可能である。そのためユーザのジェスチャをそのままキャラクタに反映させたり、ユーザの手足の動きをコマンドとして用いるなどの方法が実際のゲームに用いられている。しかし、現状のキネクトで得られるデプス画像のノイズは大きく、腕組みなど手足が身体、または他の四肢に近接しているような姿勢は計測が難しい。また一方で、キネクトはリアルタイムで計測できる安価なデプスセンサとしても注目を浴びており、主に拡張現実感(Augmented Reality: AR)の分野でもキネクトを応用したものが学会で多数発表されている⁴⁾。この場合は、複数のデプス画像を統合する処理が行われているため、ノイズの影響が軽減され、誤認識の確率を低く抑えることが可能となっている。

3 最近の研究動向

ゲーム向けのインタフェースは、主にヒューマンコンピュータインタラクション(HCI)の分野でも注目を浴びており、次世代に向けたシステムのデザインなどに関する研究は数多く行われている。前述の加速度センサやデプスセンサを用いた入力用インタフェースの研究は現在でも盛んに行われているが、新しい傾向として小型プロジェクタなどの出力デバイスとの融合を考慮したものも現れてきた。

既存の出力デバイスとしては、テレビや、携帯電話、タブレットなどの小型ディスプレイがあり、現在のゲームはこれらを前提として開発されている。一方で携帯可能な小型プロジェクタは、ユーザがそれを持ちながら動くことによって、どこにでも映像を投影することができ、センサで検知したユーザのジェスチャと組合せることで、新しいゲー

この図の著作権は著者に帰属し、学会への譲渡・転載許諾を得ることができませんでしたので削除いたしました。

この図の著作権は著者に帰属し、学会への譲渡・転載許諾を得ることができませんでしたので削除いたしました。

この図の著作権は著者に帰属し、学会への譲渡・転載許諾を得ることができませんでしたので削除いたしました。

ムなどのインタラクシオン用途が可能になると考えられる。また、小型プロジェクタは、携帯電話やタブレットへの内蔵といった形で、2014年までには3,900万台の普及が見込まれており、主に携帯ゲームを中心として新しい可能性を広げることができると考えられている。

プロジェクタを用いた場合のインタラクシオンでは、ゲームのキャラクターや映像を実世界に投影できる点が、既存のゲームと大きく異なる。昨今提案されている研究では、この点を利用し、ゲーム世界の中でのキャラクターや物体とのインタラクシオンだけでなく、仮想キャラクターと現実世界の物体との間でのインタラクシオンを目指したものが多い。現実世界と仮想世界との融合を目指したものとしてARが有名であるが、ARでは、ヘッドマウントディスプレイや携帯端末の画面を通して仮想物体やキャラクターが現実世界上に表示されるのに対し、プロジェクタを用いたシステムでは、現実世界に直接仮想物体を投影することで、デバイスを持ち合わせていないユーザにも閲覧が可能となる。

MotionBeamでは、小型プロジェクタを用いたインタラクシオンシステムのデザインにおける指針を示してあり、仮想キャラクター同士のインタラクシオンだけでなく、仮想キャラクターと現実世界の物体とのインタラクシオンについても述べている(図3)⁵⁾。MotionBeamでのプロトタイプデバイスは、携帯型プロジェクタ、IMU、赤外線検知器、ソナーを搭載し、基本的にはユーザのジェスチャをIMUで推定し、それに合わせてキャラクターの動きや視点変更を行い、プロジェクタから映像を投影する(図4)。また、現実世界とのインタラクシオンには、環境中に赤外LEDと物体の画像を隣接して配置し、デバイスが赤外LEDを検知した時に、キャラクターがそれに応じた動きを始める。例として、トランポリンのシールとそのシールの直ぐ側に赤外LEDを設置し、デバイスがその赤外LEDを検知した時に、キャラクターがジャンプを始める、といった現実世界の物体

が仮想キャラクターに影響を与えるデモを挙げている。同様に、写真フレームに同様の装置とフレームのフックが外れる機構を取り付け、キャラクターがフレームの上に来た時にフレームが外れて落ちる、といった仮想キャラクターが現実世界の物体に影響を与える、というプロトタイプシステムも提示している。こういった現実世界とのインタラクシオンには、環境側に大なり小なり特殊な機構を設置する必要があり、現状では家でのゲームプレーにあまり向いていないかもしれないが、遊園地などのアトラクションには前もってそういった機構を取り付けておくことが可能であり、新しいエンタテインメントを提供できると思われる。

PicoPetは、デジタルペットを飼うシステムで、携帯型プロジェクタを用いてペットを現実世界に投影して操作を行う⁶⁾。本システムは、仮想キャラクター同士のインタラクシオン、および、現実世界から仮想キャラクターへのインタラクシオンを主として設計されている。PicoPetでのデバイスには通常のカメラが搭載されており、カメラで投影されているペットの周辺のテクスチャや色を解析して、ペットの動作や状態に影響を与える機能を実装している。例として、りんごにペットを投影した場合はペットがりんごを食べ始める、花柄のある服や壁に投影した場合はペットが花で遊びだす、といったインタラクシオンを実装している。またペットとともにマーカを投影し、そのマーカを他のユーザが持つデバイスのカメラで検知することで、ペット間のインタラクシオンも可能となっている。

SideBySideでは、携帯型プロジェクタを含むデバイスを用いて複数のユーザがインタラクシオンを行えるシステムについて提案している⁷⁾。本システムは、主に仮想キャラクター同士のインタラクシオン、特に複数のユーザが投影する仮想キャラクターや仮想物体間でのインタラクシオンを主眼においたものである。SideBySideでは、プロジェクタから赤外光と可視光が放出され、赤外光のチャンネルではAR

マークが、可視光のマークではキャラクターのアニメーション等が投影される。この赤外光のチャンネルで投影されるARマークは、相手ユーザ側の赤外カメラで検出、キャラクターの位置や状態などを認識することで、自分のキャラクターや物体がそれに応じた反応を示す、というインタラクションが可能となっている。例として、ボクシングゲームを実装しており、自分のキャラクターがパンチをしたときに、相手のキャラクターが一定距離内にいる場合のみにダメージを与える、などの機能を挙げている。マークを使うという点ではPicoPetと同じであるが、赤外光を使うことで、ユーザからはマークが見えずキャラクターのみが見えるため、マークが操作やインタラクションの没入感を妨げない、という点で優れている。

一方で、プロジェクタを通して現実世界の物体へのインタラクションを目指したシステムとして、石井らの研究では、ユーザがレーザポインタを用いて現実世界に線や円などを描くことで、ロボットにコマンドを送るシステムを提案している⁸⁾。天井にカメラを設置し、レーザポインタの軌跡を追従、認識し、コマンドに変換してロボットに送る。例として、掃除機の現在の場所と移動先の場所を指定すると、天井に設置されたプロジェクタによってレーザポインタの軌跡がユーザへのフィードバックとして表示され、ロボットが移動先へと押していくデモが提示されている。なお、本研究では携帯型プロジェクタを用いていないが、ユーザがレーザポインタで空間中に簡単な絵を描く手法をとっており、これをプロジェクタで代替するのは容易であるため、この紹介の中に含めた。

同様のシステムで携帯型プロジェクタを用いたものとしては、PICOntrolが挙げられる⁹⁾。PICOntrolにおけるプロジェクタは、グラフィカルユーザインタフェースをユーザに対して表示する機能と、プロジェクタからコードパターンを投影することで現実世界の物体にコマンドを送る機能の二つを持ち合わせている。例として、ラジコンの操作を上げており、ラジコンに直接プロジェクタの映像を当てることで、ラジコンの移動や回転などを指示できる"Direct-manipulation"のモードを実装する一方、ラジコンに移動する場所を指定する"Follow-the-center"のモードを実装している。また、照明を少し改造し、PICOntrolのプロジェクタの光を検知してオン・オフを切替えたり、表示される矢印にそってデバイスを動かすことで明るさを調整する、などのデモも示している。

最後の二つの論文でのデモは、ロボットや家具などが人間の日常生活をより快適なものにするヒューマンロボットインタラクションを主眼においたものであるが、現実世界を舞台にしたゲームへの適用ももちろん可能である。現状ではまだあまり家庭用のロボットやプロジェクタやセンサとの連携が想定された電化製品は普及していないが、プロジェクタ内蔵型のデバイスが標準となった場合は、対応製

品が多く市販されることが予想される。そうなった場合は、日常生活が快適になるだけでなく、日常生活の中でゲームを繰り返されることが多い可能性がある。

4 むすびと今後の展開

以上、本稿では現在市販されているゲーム用インタフェースの紹介と、最近の研究動向についてまとめた。近年はセンサ自身の開発だけでなく、センサデータを利用するためのアルゴリズムやデータベース技術の進展もめまぐるしく、数年前には不可能、もしくは実用面で意味がないとまで考えられていた研究分野にも再び注目が集まってきている。そういった動向を踏まえて、近年では企業だけでなく大学でも実用化を踏まえた研究が行われている。常に最先端の情報を収集、取捨選択して行かなければならなくなっており、斬新な研究成果、製品を作っていくには大変な分野であると同時に、将来が非常に楽しみな分野でもある。

今後の展望としては、まず音声認識技術の適用が考えられる。前述のキネクトにはマイクロフォンアレイも搭載されており、プレーヤからの音声入力を認識できるゲームの登場が期待される。ここで注目したいのは、従来使われていたモノクロのマイクではなく、マイクロフォンアレイが搭載されている点である。マイクロフォンアレイは複数のマイクから構成される音声入力装置であり、ビームフォーミング (Beamforming) などの技術を用いることで、ノイズの除去や特定の話者の音声の分離、音源の位置推定などが可能となる。さらに、デプスカメラからのデプス画像、もしくはキネクトのようなスケルトン推定機構と組合せることによって、より高レベルなプレーヤの状態推定を行うことが可能となるであろう。筆者の主観であるが、音声はコントローラの操作やジェスチャなどにあまり影響を与えることなくプレーヤが発音できるため、新しいゲーム操作、または、より細かなゲーム操作を生み出すことができるのではないかと考えている。

また、ジェスチャによるインタフェースの最も大きな課題として、触感フィードバック (Tactile Feedback) が挙げられる。現状のゲームコントローラにはバイブレーション機能が搭載されており、打撃をくらうなどの感覚を振動のオン・オフで伝えるものは多数適用されているが、精度の高いジェスチャ入力が可能となった現状を見るに、バイブレーション機能だけでは不十分になってきている。われわれが現実世界で物体を動かすとき、その物体を今触っているかどうか、触っている部分が角なのか、球面なのか、それとも平面なのか、またはその接触部分が滑りやすいか、ざらざらしているのか、など触感を通して非常に多くの情報を得ており、その結果、物体を見なくても手のひらで転がしたり変形させたり投げたりすることができる。プレーヤのジェスチャを精度よく推定できるようになった現在、ゲームなど仮想空間と接するシステムで欠けているのはこ

の触感であり、そのためゲーム内の物体を触ろうとしたり操作しようとするときの難しさの一つはここにある。触感フィードバックの研究は、ジェスチャ認識と同様昔から盛んに行われており、テーマパーク向けのシステムはすでに実用化されており、また家庭用ゲーム向けのプロトタイプシステムとして、レーシングゲーム中の振動や車の動きなどをプレイヤーの背中に伝えるフィードバックシステムも開発されている¹⁰⁾。触感フィードバックが実用化された暁には、近年のゲーム用入力インタフェースの進展と同じような衝撃があるのではないかと考えている。

(2012年10月9日受付)

〔文 献〕

- 1) T. Shiratori, J.K. Hodgins: "Accelerometer-based User Interfaces for the Control of a Physically Simulated Character", ACM Transactions on Graphics, 27, 5 (2008)
- 2) J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, A. Blake: "Real-Time Human Pose Recognition in Parts from a Single Depth Image", in Proc. IEEE International Conference on Computer Vision and Pattern Recognition, 1297-1304 (2011)
- 3) T. Leyvand, C. Meekhof, Y.-C. Wei, J. Sun, B. Guo: "Kinect Identity: Technology and Experience", IEEE Computer, 44, 4, 94-96 (2011)
- 4) S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R.A. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A.J. Davison, A. Fitzgibbon: "KinectFusion: Real-Time 3D Reconstruction and Interaction Using a Moving Depth Camera", in Proc. ACM Symposium on User Interface Software and Technology, 559-568 (2011)
- 5) K.D.D. Willis, I. Poupyrev, T. Shiratori: "MotionBeam: A Metaphor for Character Interaction with Handheld Projectors", in Proc. ACM SIGCHI Conference on Human Factors in Computing Systems, 1031-1040 (2011)
- 6) Y. Zhao, C. Xue, X. Cao, Y. Shi: "PicoPet: "Real World" Digital Pet on a Handheld Projector", ACM Symposium on User Interface Software and Technology (demo) (2011)
- 7) K.D.D. Willis, I. Poupyrev, S.E. Hudson, M. Mahler: "SideBySide: Ad-Hoc Multi-User Interaction with Handheld Projectors", in Proc. ACM Symposium on User Interface Software and Technology, 431-440 (2011)
- 8) K. Ishii, S. Zhao, M. Inami, T. Igarashi, M. Imai: "Designing Laser Gesture Interface for Robot Control", in Proc. IFIP Conference on Human-Computer Interaction, 479-492 (2009)
- 9) D. Schmidt, D. Molyneaux, X. Cao: "PICOntrol: Using a Handheld Projector for Direct Control of Physical Devices through Visible Light", in Proc. ACM Symposium on User Interface Software and Technology (2012)
- 10) A. Israr, I. Poupyrev, C. Ioffreda, J. Cox, N. Gouveia, H. Bowles, A. Brakis, B. Knight, K. Mitchell, T. Williams: "Surround Haptics: Sending Shivers Down Your Spine", SIGGRAPH Emerging Technology (2011)

しらとり たかあき
白鳥 貴亮 2002年、東京大学工学部電子情報工学科卒業。2007年、同大学院情報理工学系研究科電子情報学専攻博士課程修了。同年より、Carnegie Mellon Universityにて、Postdoctoral FellowおよびProject Scientist。また、2010年より、Disney Research, Pittsburghにて、Postdoctoral Associateを経て、2012年より、Microsoft Research Asia研究員。キャラクターアニメーションの制作およびインタラクションのためのユーザインタフェースに関する研究に従事。