解説

Kinect によるユーザインタフェース

The User Interface by Kinect

中 村 薫* 株式会社ゲッシュ
Kaoru Nakamura* *GEIS Co. Ltd.

1. Kinect とは

Kinect は米 Microsoft 社が発売している Xbox 360 ゲーム機のセンサ・コントローラである。Kinect の最大の特長は、ゲームをプレイする際にコントローラを持たないことである。Kinect には RGB センサ、距離を計測できる Depth センサ、四つのマイクから構成されるマイクアレイが搭載されている (図 1)。この中の Depth センサを利用して、プレイヤーを認識し、さらにプレイヤーの骨格までも認識することで、プレイヤーの動きをもとにゲームをコントロールすることができる。

この特長を生かして Kinect を様々なソフトウェア・アプリケーションに応用できないか、という探求が世界各地でなされている。Kinect は変換ケーブルによって USB ポートへの接続も可能であるため、発売当初からコンピュータへ接続して利用する試みが行われていた。

なお、本稿の情報は 2012 年 7 月現在の情報であり、発刊時の情報とは差異がある可能性があることに注意されたい。

2. Kinect for Windows および Kinect for Windows SDK

Xbox 360 用の Kinect は正式にはコンピュータへの接続は認められておらず、商用利用も不可能である。しかし、2012 年 2 月に Windows PC での商用利用が認められた「Kinect for Windows」が発売された (Mac OS や Linux での利用は認められていない)。同時に Kinect for Windows で利用可能な Kinect for Windows SDK もリリースされた。これによって、Kinect を利用した商業利用が可能になったのである。以降「Kinect」と表現した場合には「Kinect for Windows」を指すこととし、同様に「SDK」と表現した場合には「Kinect for Windows SDK v1.5」を指すこととする。Kinect for Windows SDK v1.5 は、5 月にリリースされた最新の SDK である (図 2)。

図 1 Kinect Sensor [1]

図 2 Kinect for Windows SDK

3. Kinect の機能

Kinect ができることは以下の項目である。SDK の基本機能も、以下の Kinect の機能に準じている。

- ・ RGB センサを利用したカラー画像の取得
- ・ Depth センサを利用した距離データの取得
- ・ プレイヤの認識および追跡 (6 人まで)
- ・ プレイヤの骨格の認識および追跡
- ・ 音声の入力および音声方向の認識

特徴的な距離データは通常モードで 0.8~4 [m]、Kinect for Windows 専用の近接 (Near) モードで 0.4~3 [m] の間を認識することができる (図 3)。プレイヤの認識は 6 人まで、そのうち骨格を追跡できるのは二人までである (図 4)。

骨格の追跡について、認識できる関節は全体で 20 カ所。Seated Mode と呼ばれる上半身のみの骨格追跡は 10 カ所となっている (図 5)。

原稿受付 2012 年 7 月 31 日

キーワード: Kinect, User Interface, Kinect for Windows SDK

*〒 107-0052 港区赤坂 4-3-1

*Minato-ku, Tokyo

図 3 Kinect for Windows の距離認識範囲 [2]

図 4 Kinect for Windows が認識できるプレイヤー [1]

図 5 全身の骨格追跡と上半身のみの追跡 (Seated Mode) [1]

4. Kinect の応用範囲

Kinect および SDK の機能は非常に小さい。しかし、Kinect を使った様々なアウトプットが出ていることから分かります。応用範囲は非常に広い。SDK 同様に配布されているサンプルや拡張ライブラリ集である「Kinect for Windows Toolkit」(図 6)を見ると分かるように、音声入力を利用した「音声認識」や、骨格の認識を利用した「顔検出」や「顔の向きの検出」(図 7)、XNA と呼ばれるゲーム用のライブラリと骨格の認識を組み合わせた「アバター」(図 8)など、既存の技術と組み合わせることで様々な使い方ができる。

5. Kinect を利用するメリット

興味深い点は「既存技術と組み合わせる」というところだ。先ほど上げた技術は、すでに今ある技術でも解決できることだ。しかし、Kinect を利用することでそれらをより

図 6 Kinect for Windows Developer Toolkit v1.5.0

図 7 Kinect を利用した顔検出

図 8 3D アバターサンプル

簡単に行うことができる。そのため開発者は「さらにその先」を考えることに注力できるだろう。顔の検出を例に挙げると、従来であればカメラ画像の更新ごとにすべての画素に対して顔検出を行う必要があった。Kinect を利用すると顔の位置が特定できるので、そこから一定の範囲内のみを検出の対象とすることで、より CPU のコストを下げることができる。

音声認識についても、従来であれば音声を認識するだけにとどまっていたが、Kinect を使うことで音声方向を利用することができ、さらに骨格追跡と組み合わせることで「誰が」「何を」話しているかを認識できるようになった。

もちろん今までにない、まったく新しいことが実現可能

にもなっている。Kinect にはカラー画像、距離データ、音声データを一つのデバイスで入力できる。それぞれのデータを連携させることで、今までは実現が難しかったソフトウェアも実現可能になるだろう。

6. Kinect の利用事例

2012 年 2 月の Kinect for Windows の発売（すなわち商用利用の解禁）以降、日本国内での Kinect 事例は非常に多いと感じている。すでにいくつかの Kinect アプリケーションが世の中で稼働されており、拙著も含めて書籍も多く出版されている。展示会に足を運ぶと、多くの企業が Kinect を使った製品の展示を行っており、様々な企業が Kinect に興味を持ち実現の可能性を模索しているように感じられる。

それと同時に Kinect で実現可能なことが多く、また今までに類似のデバイスが少ないことから、Kinect を持て余してしまっている現状もあるとセミナー開催などを通して感じている。

Kinect の利用方法は様々で、主にユーザインタフェースまたはセンサとしての利用になるが、ユーザインタフェース一つをとってみても、骨格を認識させてのコンピューターを操作することから始まり、外部のデバイスと連携させるためのリモートコントローラや、距離データのみを利用したインタフェースまで様々だ。センサとして利用する場合は、距離データと RGB データを組み合わせ、物体や人の認識を行っていることが多い。

これらの特性から、ロボットの目としての役割を担わせることもでき、今までよりも革新的な成果が出ている例も聞く。

7. Kinect のユーザインタフェースへの利用

Kinect が Xbox 360 のコントローラであることから、まず始めに思いつくのはユーザインタフェースへの利用だろう。ただし（筆者の感覚であることを前置きして）Kinect のユーザインタフェースへの利用は非常に難しい。

要因としては以下の理由が考えられる。これらは Xbox 360 のゲームや世の中に公開されているアプリケーションを実際に体験することで理解することができるので、未体験ということであればいち早く体験して欲しい。

- ・従来のコントローラやマウスのように細かい制御が難しい
- ・非常に疲れる
- ・入力に対するフィードバックがない

Kinect をユーザインタフェースとして利用する際には、これらに気を配る必要があると感じている。

(1) 従来のコントローラやマウスのように細かい制御が難しい

手を使って従来のマウスのように入力してもらおうとしても、従来と同じようにはいかない。完全に手を固定する

ことはできない、ポインタがぶれてしまい、細かい動きをすることが難しい。これによって従来のマウスを扱うような細かい操作はユーザにとって難しい作業となっている。

解決策としては、選択肢を大きくすることや、「スナップ」と呼ばれる選択肢に近づくと磁石のように張り付くようにするなどが考えられる。

(2) 非常に疲れる

これも手を使ってインタフェースさせるときに気をつけることになるが、ユーザに手を前に出している状態で操作させる状態になると、従来の入力方法と比べて体にかかる負担が大きくなる。

そのため、手順の単純化や操作のメリハリなど、負担を和らげることを考慮する必要がある。

(3) 入力に対するフィードバックがない

空間を使ったインタフェースでは、意図的にフィードバックを作らない限り、操作に対するフィードバックがかからない。マウスであれば腕と机のこすれ具合や、クリックしたときの感触、タッチパネルであれば画面との摩擦といったフィードバックがある。人間が操作する上でフィードバックが非常に大切であることは、Kinect を利用したインタフェースを使ってみるとよく分かる。

実際に Kinect を利用したインタフェースを実装する際は、より多くのフィードバックを与えること、意識することが大切になってくる（図 9）。

これらを解決する一番簡単な方法は「スケールの大きな物を作る」ことだと私は考えている。例えば図 10 はある遊園地のゲームイベントとして稼働している Kinect のコンテンツだ。これは巨大なスクリーンにコンテンツを映し出し、プレイヤーは Kinect でゲームをプレイする。このスケールだと臨場感や没入感から、自然な入力を行うことができた。また、ゲーム中の技の入力も緩く設定してあり、操作に対するストレスもほとんど感じられなかった。

8. ユーザインタフェースのパターン

Kinect を使ったユーザインタフェースには Kinect と合

図 9 Kinect を利用したユーザインタフェース風景 [3]

図 10 Kinect を利用したイベント

図 11 空中立体ディスプレイを操作している様子

う組み合わせと難しい組み合わせがあると感じている。

(1) 合う組み合わせ 1: 3D 出力装置との組み合わせ

3D テレビや立体出力装置と組み合わせ、3D の入力装置として Kinect を利用する方法である。Kinect の骨格座標は三次元であるため、出力装置も三次元であるほうが操作しやすく、また Kinect を入力装置として利用する意義が出てくる。

具体的には弊社の事例として、空中立体ディスプレイの入力装置として Kinect を利用したコンテンツを開発した。空中立体ディスプレイは、空間上に画像を結像するため、空間で操作できる Kinect と非常に相性が良い (図 11)。

(2) 合う組み合わせ 2: プロジェクタ

Kinect で入力したデータを、コンピュータを通してプロジェクタに出力し、現実世界へアウトプットする方法である。

具体的には、図 12 にあるような何もない壁に対してインタフェースするという方法がある。壁に対してプロジェクタと Kinect を照射し、壁との距離をもとに間に入った手をペンに見立てる。ペンに見立てた手の軌跡によって描かれた絵はプロジェクタを通して現実世界に戻り、その組み合わせからあたかも壁に絵を書いているように見せることができる。

これらの二つに共通していることは「現実を拡張している」ということだ。両方とも空間を拡張し、その拡張された空間に対してインタフェースさせるための道具として Kinect を利用している。

図 12 Kinect とプロジェクタを利用して、何もない壁に絵を描く [4]

(3) 難しい組み合わせ

では Kinect との組み合わせが難しいと考えているものは何かというと、一般的な 2D のテレビやディスプレイである。前述までのとおり、Kinect は現実を拡張したものに対して操作を行うことに向いている。そのため、2D の出力装置や仮想空間内に表現されたコンテンツに対する操作は難しいのではないかと感じている。これは Xbox のゲームをプレイしていても同様に、Kinect のハードウェアの特性を知り尽くしている Microsoft 社のゲームと、それ以外の会社のゲームで Kinect の操作性について大きな差異があるように感じている。実は従来のコンピューティング環境で Kinect を利用したインタフェースを実現することが難しいのではないだろうか。

9. 補助としての音声認識

このように Kinect を利用したユーザインタフェースにはまだまだ改善の余地がある。改善のための一つの補助ツールが音声認識だ。Kinect のマイクと Microsoft 社が元々持っている音声認識ライブラリを利用することで、音声コマンドを簡単に実装することができる。これによって、ジェスチャーインタフェースを補助する形で音声コマンドを追加することで、ユーザビリティを上げることができる。最新の SDK では日本語の音声認識もサポートされている。

幸い最近のスマートフォンでは音声認識を利用したインタフェースが実装されていることが多く、ユーザにとっても身近な存在となっているため、Kinect のインタフェースに音声認識を加えることはユーザにとっても心理的抵抗が少ないのではないだろうか。

10. 心理的な抵抗について

先ほど述べた心理的抵抗について補足したい。Kinect のジェスチャーコントロールや音声認識は、今までのコンピュータへの操作とは大きく異なり、自分自身の動作、音声を用いてコンピュータを操作する。そこには「恥ずかしさ」という気持ちが存在し、操作に対するモチベーションを低下させる可能性がある。実際に筆者も Xbox 360 のゲーム

を一人でプレイしていると、横にいる家族の目が気になり、なかなかゲームに集中できないといった経験がある。

Kinect を用いたインタフェースを実装する際は、このような「恥ずかしさ」を少しでも和らげる、または恥ずかしさを上回るメリットをユーザに対して提示する必要があるだろう。

11. Kinect のこれから

現在、Kinect には多くの方が注目している。SDK も少しずつバージョンアップされ、新しい機能が数多く追加されている。今年は Windows 8 も登場し、Windows 環境も大きく変わるだろう。Kinect はスマートフォンやタブレットのようなタッチインタフェースとの相性が良いため、Windows 8 との相性も良いと考えている。現状では SDK の制限により、Windows 8 の最大の特長である Windows ストアアプリ（旧メトロ UI）に対応した Kinect アプリケーションを実装することはできないが、近いうちに対応されアプリケーションも数多く発表されることだろう。ソフトウェア面だけではなく、ハードウェア面での進化も期待したい。外部電源が不要になればウェアラブルの道も開

け、様々な機器への搭載も期待できるだろう。ソフトウェアとハードウェア両面の進化によって、Kinect が様々なデバイスに搭載され、生活の一部として有効な位置を占めることを願っている。また、本稿を読まれた方たちが Kinect に興味を持ち、将来素晴らしい製品を開発するための参考になると幸いだ。

参 考 文 献

- [1] <http://msdn.microsoft.com/en-us/library/hh855347>
- [2] Kinect for Windows SDK v1.0 付属ドキュメント
- [3] 中村薫, 齋藤俊太, 宮城英人: KINECT for Windows SDK プログラミング C++編. 秀和システム, 2012.
- [4] 中村薫, 田中和希, 宮城英人: KINECT for Windows SDK プログラミング C#編. 秀和システム, 2012.

中村 薫 (Kaoru Nakamura)

2011 年より株式会社ゲッシュに勤務し Kinect 関連の開発、執筆、講演などを多数行う。Microsoft MVP for Visual Studio ALM を受賞し、開発のみならず開発マネジメントも得意とする。