

Kinect センサーを用いたプレゼン評価システム

Presentation evaluation system using the Kinect sensor

河上 奏太^{*1}, 伊藤 貴大^{*2}, 清水 翔平^{*2}, 松本 豊司^{*3}
Sota Kawakami^{*1}, Itoh Takahiro^{*2}, Simizu Shohei^{*2}, Matsumoto Toyoji^{*3}

*1 金沢大学 理工学域 機械工学類

*1 Mechanical Engineering, College of Science and Engineering, Kanazawa University

*2 金沢大学大学院 自然科学研究科 数物科学専攻

*2 Division of Mathematical and Physical Sciences,
Graduate School of Natural Science & Technology, Kanazawa University

*3 金沢大学総合メディア基盤センター情報教育部門

*3 Information Media Center, Kanazawa University

Email: gymnopedian@gmail.com

あらまし：プレゼンの経験が少ない学生でも、容易に学習を行えるプレゼン評価システムを開発している。システムは、Kinect センサーを用いて姿勢情報を取得し、模範例を参考にプレゼンテーションを行っている学生の動作を解析し、模範的なプレゼンテーション例と比較評価を行っている。

キーワード：Kinect、モーションキャプチャ、プレゼンテーション

1. はじめに

魅力的なプレゼンテーションを行うためにプレゼンテーションの練習をするというのは非常に難しい。原稿を覚えたり、スライドに工夫を凝らしたりすることは一人でも出来る。けれども、実際に話す自分の姿を確認して改善するには聞き手が必要である。さらに、どのようなプレゼンテーションを行えば身内以外の聞き手に興味を持ち続けてもらえるか、といったことは判断が難しい。

そこで、我々は Kinect センサーを用いてプレゼンテーションを行っている人の姿勢情報を解析することでプレゼンテーションを定量的に評価し、効果的なプレゼンテーションの練習を行うことが出来るシステムの開発を行った。

2014年3月の段階では評価する対象の腕の動きをカウントし、手を使ったジェスチャーをカウントすることで評価を行うシステムの開発を行った。しかし、単純に腕の動きをカウントするだけでは、マイクの有無・パソコンの操作・意味のない動き等が評価する際のノイズとなるといった課題があった。そこで、我々は評価方法を次のように変更した。模範となるプレゼンテーションの骨格情報と評価対象の骨格情報の類似度を数値化することで評価を行う方法に変更した。我々は、この評価方法でシステムを開発した。さらに、インターフェイスの改善も行った。

2. Kinect センサーについて

Kinect センサーとは、本来 2010 年にマイクロソフトより発売されたジェスチャーや音声認識によって直感的に操作が可能なゲームデバイスであった。と

ころが、センサーとして非常に優秀であったため発売当初から有志による解析が行われた。その後、2012年に公式に Windows 対応の Kinect センサー及び開発キットが公開された。

Kinect には、RGB カメラ・赤外線カメラ・マイクが搭載されており、今回は主に赤外線カメラを用いて深度情報から骨格情報を獲得し、利用する。

3. 評価方法

模範となるプレゼンテーションの動画を用意し、評価対象となる人に Kinect の前で模範のプレゼンテーションを模倣してプレゼンテーションを行ってもらおう。このときの評価対象のジェスチャーと、模範となるプレゼンテーションのジェスチャーを比較し、どの程度模範に近いプレゼンテーションを行えているかを数値化することで評価を行う。

3.1 ジェスチャーの類似度を評価する方法

「ジェスチャー＝時間的に対応のあるポーズの列」である。

ジェスチャーの類似度の評価：

- ①空間的な類似度・・・ポーズの類似度
- ②時間的な類似度・・・ポーズの時間的な変化の類似度

本来、ジェスチャーの類似度を評価する場合には、この2つを評価する必要がある。

私たちは、ジェスチャーを簡単に評価するために、①のみを評価することにする。

つまり、「動画に描画されたジェスチャーを評価する」という問題を次のように置き換える。

「動画を構成する 1 枚 1 枚の画像に描画されたポーズを評価する」

したがって、私たちは、画像に描画されたポーズ

の類似度を評価する。

これを動画を構成する画像のほとんど全てで行うことにより、ジェスチャーの類似度を評価できると考えている。

ここで、「ほとんど全て」とは、1 秒間における数十枚の画像を全て評価しないで、適当に何枚か選んで評価することにする。これは計算量を削減するためである。

3.2 見本と被験者の骨格情報を利用して、ポーズの類似度を評価する方法

■骨格情報の取得

私たちが使用する Kinect は、20 か所の骨格の 3 次元の位置情報、座標 (x, y, z) を取得することが出来る。

これらのうち、体の先端に位置する、頭・手(右・左)・足(右・左)の 5 か所を除く、15 か所の骨格の座標を使う。

つまり、見本と被験者の、この 15 か所の骨格の座標の比較を行う。

■ポーズを測る数学的な量

① 比較したい骨格を先述の 15 か所の中から 1 つ、この骨格 A を決める。また、この骨格 A の座標を a とする。

② この骨格 A に隣接する 2 つの骨格 B, C の座標をそれぞれ b, c とする。

③ a を始点とし、 b を終点とするベクトルを u 、 a を始点とし、 c を終点とするベクトルを v とする。

④ ベクトル u, v が成す角度を $\theta = (\theta_x, \theta_y, \theta_z)$ とする。

この角度 θ は、骨格 A の周囲における局所的なポーズの情報を表す量である。

したがって、各骨格におけるこの角度を、見本と被験者について求める。図 1 に使用する骨格の概要を示す。

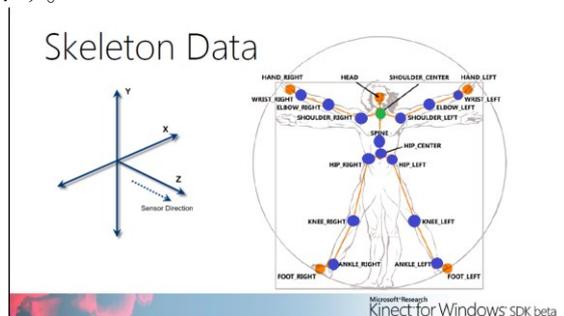


図 1 評価に使用する骨格

そして、それらを比較することによりポーズの類似度を評価する。

■評価方法

① 見本の、先述の 15 か所全ての骨格における角度、 $\theta_1, \theta_2, \dots, \theta_{15}$ を求める。

② 被験者の、先述の 15 か所全ての骨格における角度、 $\theta_1, \theta_2, \dots, \theta_{15}$ を求める。

③ 15 か所全体の二乗平均誤差 E は、

$$E = \frac{1}{15} \sqrt{\sum_{i=1}^{15} \{(\theta_{x_i} - \theta_{x_i}')^2 + \theta_{y_i} - \theta_{y_i}'^2 + (\theta_{z_i} - \theta_{z_i}')^2\}}$$

この誤差 E を用いて、被験者が、その時、見本のポーズと、どのくらい似たポーズを取っているか測ることが出来る。

図 2 に開発中の評価画面を示す。自分の骨格情報(画面左側)と見本の骨格情報(画面右側)が並んで表示される。左上に総合評価が表示され、右上のボタンを押すことで見本の動画が全画面表示で再生される。

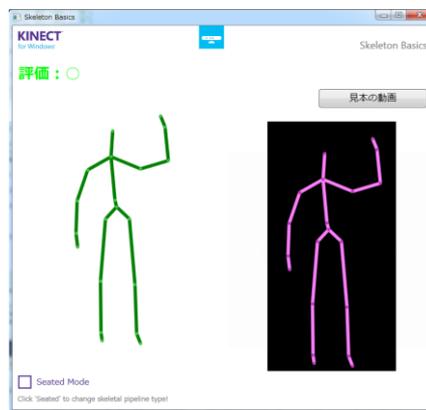


図 2 : 開発中の評価画面の様子

4. 今後の課題と展望

現在はまだシステム開発を行っている段階だが、今後完成したシステムを実際に使用してもらい、フィードバックを得ることでより良い評価方法の検討、インターフェースの改良を行っていく。精度と取得できる情報量が向上した次世代 Kinect も今後販売される予定なので、センサーの改善による評価精度の向上も期待できる。最終的にはゲーム感覚で楽しみながら模範的なプレゼンテーションを追体験することができるシステムを目指して開発を行っていく。

5. 参考文献・謝辞

参考文献

- (1) payashim : “Kinect のトラッキング原理「部位認識に基づく 3D 姿勢推定」”
- (2) Wall Street Journal: “Key Kinect Technology Devised in Cambridge Lab”, TECH EUROPE
- (3) 中村 薫: “Kinect for Windows SDK プログラミング C # 編”
- (4) 杉浦 司、中村 薫: 「Kinect for Windows SDK」実践プログラミング

本研究は、科学研究費補助金挑戦的萌芽研究(課題番号 25540161)の支援の下に行われた。ここに記して謝辞とする。