

## Kinect センサーを用いたジェスチャー評価アルゴリズム

## Gesture evaluation algorithm using the Kinect sensor

清水 翔平<sup>\*1</sup>, 伊藤 貴大<sup>\*1</sup>, 河上 奏太<sup>\*2</sup>, 松本 豊司<sup>\*3</sup>  
 Shimizu Shohei<sup>\*1</sup>, Ito Takahiro<sup>\*1</sup>, Sota Kawakami<sup>\*2</sup>, Toyoji Matsumoto<sup>\*3</sup>

<sup>\*1</sup> 金沢大学大学院 自然科学研究科 数物科学専攻

<sup>\*1</sup> Division of Mathematical and Physical Sciences,  
 Graduate School of Natural Science & Technology, Kanazawa University

<sup>\*2</sup> 金沢大学 理工学域 機械工学類

<sup>\*2</sup> Mechanical Engineering, College of Science and Engineering, Kanazawa University

<sup>\*3</sup> 金沢大学 総合メディア基盤センター

<sup>\*3</sup> Information Media Center, Kanazawa University

Email: shishou@stu.kanazawa-u.ac.jp

あらまし : 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ジェスチャーを解析し、比較評価と採点を行うシステムを開発している。本アルゴリズムでは、より明確な比較評価と採点の方法を目指した。

キーワード : Kinect、ジェスチャー

## 1. はじめに

魅力的なプレゼンテーションを行うために一人で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。

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。ここでは、主にジェスチャーの評価方法について述べる。

## 2. Kinect センサーについて

Kinect には、RGB カメラ・赤外線カメラ・マイクが搭載されている。今回は RGB カメラを用いてスクリーンショットを撮り、赤外線カメラを用いて深度情報から骨格情報を取得する。ここで、骨格情報とは図1に示す20箇所の骨格一つ一つの3次元の位置座標 $(x, y, z)$ のことである。

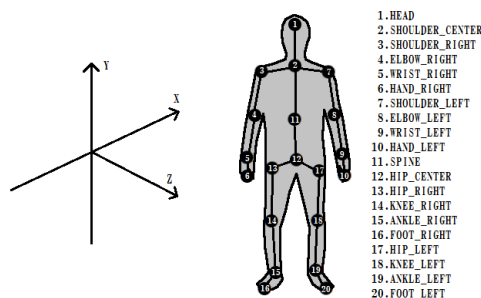


図1 位置情報を取得できる骨格

## 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姿勢情報と、模範となるプレゼンテーションの姿勢情報を比較し、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。また、スコアで評価することにより、競争意識を高めるねらいもある。なお、このアルゴリズムはジェスチャーごとに変更する必要がない点で優れている。

### 3.1 模範的なプレゼンテーション動画の作成

TED が公開している、教材として適切なプレゼンテーションの動画がある。筆者がそれを真似てプレゼンテーションを行う。

その様子を Kinect 用いて1フレーム(1秒=30フレーム)毎にスクリーンショットとして保存する。それを繋ぎ合わせ、模範のプレゼンテーション動画とする。同時に1枚のスクリーンショットに対応するすべての骨格情報を取得し、保存する。

### 3.2 誤差評価

二つのジェスチャーがどれくらい似ているか(誤差)を数値化する。いま、ジェスチャーの情報というのは、あるフレーム $a$ からあるフレーム $b$ までの骨格情報の組である。したがって、3.1で取得した骨格情報と、評価対象となる人が行ったジェスチャーの骨格情報とを比較した1フレーム毎の誤差を求めることが出来るはずである。いま、 $j$ 番目のフレーム誤差を $E_j(a \leq j \leq b)$ とおく。

例えば、図3のようにひじを曲げているポーズがあったとする。左ひじを評価の対象となる*i*番目の骨格とする。左ひじから左肩へのベクトルと、左ひじから左手首へのベクトルの成す角を $\theta_i$ とする。図で $\theta_i = \theta'_i$ ならば、角度は同じであるが違うポーズである。従って、ひじの角度とは異なる要素で比較する必要がある。

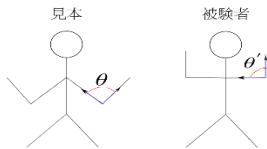


図3 ひじを曲げるポーズ

図4のように、2つのベクトルの組の始点をそろえて考える。このとき、角度 $\epsilon$ と角度 $\eta$ に注目する。これらが0になるときは、模範のベクトルの組が比較対象のベクトルの組と同じ向きを向いていることになる。つまり $\epsilon = \eta = 0$ のとき、ポーズが一致しているといえる。

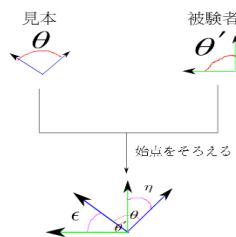


図4

ゆえに、*N*個の骨格を使用すると仮定すると、*j*番目のフレームにおける誤差 $E_j$ は次のように定義できる。

$$E_j = \sqrt{\frac{1}{N} \sum_{i=1}^N (\epsilon_i^2 + \eta_i^2)}$$

### 3.3 スコア算出法

3.2 で求めた誤差 $E_j (\geq 0)$ は、模範となるプレゼンテーションに近ければ近いほど小さい値となるので、 $E_j$ の逆数を取り、 $S_j$ とおく。 $E_j$ が0のとき、 $S_j$ は $\infty$ に発散してしまうため、

$$S_j = \begin{cases} S_j & (S_j \leq 1.5) \\ 1.5 & (S_j > 1.5) \end{cases}$$

と定めておく。 $S_j$ の総和を仮のスコア $S'$ とすると、

$$S' = \sum_{j=a}^b \frac{1}{E_j} = \sum_{j=a}^b S_j, 0 \leq S' \leq 1.5(b-a)$$

と表される。 $S_j$ の平均値が1以上、

すなわち $S' \geq b-a$ のとき、合格点を与えるものとする。ボーダーライン $S'_{border}$ は次のようになる。

$$S'_{border} = b-a$$

### 3.4 スケーリング

3.3 の評価方法で実際にジェスチャーを行う。すると、例えば棒立ちのまま動かずに評価を行ったときにも一定のスコアが出てしまうことが分かる。このスコアを $S_{stand}$ とする。 $S_{stand}$ をスコアの最低点とした新しいスコア $S''$ を以下のように定義する。

$$S'' = \begin{cases} 0 & (S' \leq S_{stand}) \\ S' - S_{stand} & (S' > S_{stand}) \end{cases}$$

新しいボーダーライン $S''_{border}$ も同様に、

$$S''_{border} = S'_{border} - S_{stand}$$

と定める。

最後に $S''$ の最高点 $S''_{max}$ が1000点となるようにスケーリングを行う。スケーリング後のスコアを真のスコア $S$ とすると、

$$S = \frac{1000}{S''_{max}} S''$$

真のボーダーライン $S_{border}$ も同様に、

$$S_{border} = \frac{1000}{S''_{max}} S''_{border}$$

となる。また、合格不合格の範囲をそれぞれ3等分し、順に Magnificent, Excellent, Good, Close, Bad, Very Bad と評価を与える。図6は講義で実際にシステムを使用してもらった学生のスコアである。

ボーダー (545点)	Aさん		Bさん	
	スコア	評価	スコア	評価
1回目	366	Bad	310	Bad
2回目	478	Close	164	Very Bad
3回目	344	Bad	843	Magnificent
4回目	353	Bad		
5回目	84	Very Bad		
6回目	557	Good		

図6 実際のスコア

## 4. 今後の課題と展望

学生にシステムを実際に使用してもらってスコアを出し、アンケートを行った。ボーダーが適切か分からないといった意見を得たので、今後より良い評価方法を検討していきたい。

## 5. 参考文献・謝辞

- (1) 河上 奏太、伊藤 貴大、清水 翔平、松本 豊司：“Kinect センサーを用いたプレゼン評価システム”，教育システム情報学会第39回全国大会

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。