

## 伝統芸能の学習時における注視行動の分析

神田直弥

### 1. はじめに

民俗芸能・伝統芸能は貴重な無形文化財である。少子高齢化や地方の過疎化に伴う後継者不足が進む中で、これらの文化財の伝承が課題となっている。モーションキャプチャは動きを3次元のデジタルデータとして記録することができる技術であるが、近年、これらの無形文化財の保存と解析を目的に利用されるようになってきている（八村, 2007）。

無形文化財の踊りを記録することで、踊りに見られる基本動作の分析が可能になる。これにより動作の構造の把握や、複数の踊りに見られる動作の比較が可能になる（白鳥ら, 2005; 三浦ら, 2008）。また、基本動作に符号を割り当ててデータベースに保存し、舞踊符を作成することで一連の動きを再現することにより、データの再利用を可能にする方法が提案されている（湯川ら, 2000）。符号化により、データを保存する際の圧縮も可能になる。

モーションキャプチャで記録できるのは、身体各部の位置や角度の3次元情報である。衣装やその動き、顔の表情、視線、舞台上における位置関係等を記録・保存できない点が限界である（八村, 2007）。これに対して、玉本（2019）は、衣装を着けた人のモデルを作成してCGアニメーションとし、舞台などの背景のCGモデルとあわせることで、舞台も含めて視点の変更や、ズームイン・ズームアウト等の操作ができるシステムを提案している。

モーションキャプチャ技術は、無形文化財の保存、伝承、鑑賞のみでなく、学習にも利用することができる。学習者が自らモーションキャプチャを装着することで、自らの動作の確認や練習を行った後での動作の熟達を確認することができる（佐藤ら, 2010）。モーションキャプチャを学習という観点で見た場合の特徴を、松本ら（2011）は以下のように整理している。

1. 3D-CGで動きを表示しズームで拡大したり様々な角度から動きを観察できる（マルチアングル）
2. 熟練者の動きと学習者の動きを重ねて比較できる

### 3. 動きの特徴や熟練者と学習者の動きのずれをグラフやアニメーション表示で可視化することができる

松本ら（2011）はこれらの特徴を生かし、民俗舞踊の学習場面を対象にしてモーションキャプチャを用いた「熟練者と学習者の比較ビューア」を開発した。これは、熟練者の動きと学習者の動きをCGで重ね合わせて表示し、動きの違いやタイミングの違いを示すものである。比較ビューアを用いることで、腕の開き具合や、手や脚の角度、動きの大きさ、スピード、タイミング等について熟練者とどのように異なるのかを観察可能であり、学習者の動作の改善に有効であることが指摘されている。松本ら（2011）の開発した比較ビューアは処理に時間を要し、リアルタイムでの確認が困難であったが、柴田ら（2012）によりリアルタイムで熟練者の動きと比較しながら学習をすることができるシステムが提案されている。

こうしたシステムの利用は、熟練者の動作の習得を可能にすると考えられる。ただし、学習者がモーションキャプチャを装着する必要があるため、誰もが気軽に利用することは難しい。一方で、熟練者がモーションキャプチャを装着して踊りを記録し、そのデータに基づいて作成したCGアニメーションを観察して学習するという環境は、熟練者のデータがあれば容易に設定できる。しかし、熟練者の踊りをビデオで撮影した映像と比較し、モーションキャプチャのデータに基づき作成したCGアニメーションを観察した場合に、学習が容易になるのかどうかは明らかになっていない。

## 2. 目的

映像を用いた動作の学習は、映像を観察するフェーズと、身体を動かしながら練習をするフェーズに分割できる。CGアニメーションとビデオ映像で学習成果に差が見られるとするならば、観察フェーズにおける観察方法の違いが影響すると考えられる。そこで本研究では、モーションキャプチャを用いて記録したデータを基に熟練者の動作をCGアニメーションで再現した映像を利用する場合と、ビデオで撮影した映像を用いる場合の注視行動を比較することで、両者の違いを明らかにすることを目的とする。

### 3. 方法

#### 3.1. 刺激映像

玉本（2019）が、慣性センサ式のモーショントラッキングを用いて記録をした黒川能の演目「大麻」のCGアニメーションと、記録時に撮影したビデオ映像をそれぞれ刺激として用いた。

ビデオ映像では、モーショントラッキングのセンサを装着するためボディースーツを着用している（図1）。CGアニメーションについても条件をそろえるため、着物を着用していないモデルを使用した（図2）。



図1 ビデオ映像

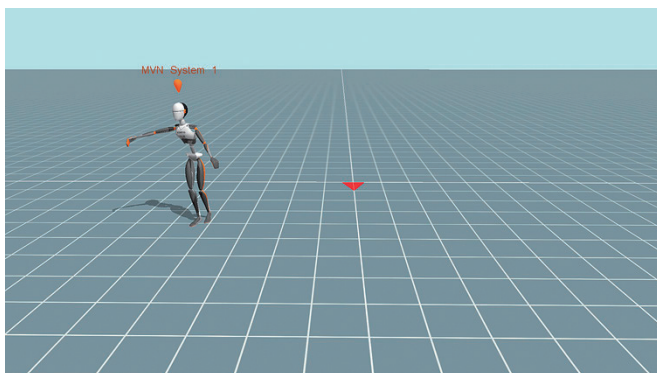


図2 CGアニメーション映像

これらの映像を用いて学習を行うにあたり、動作の量が似通った30秒間の2場面を抽出し、学習対象とした。2場面についてビデオ映像とCGアニメーションをそれぞれ準備し、合計4種類の映像を刺激映像として用いた。

### 3.2. 学習環境

消灯した実験室内において、超単焦点液晶プロジェクタ（RICOH製IPSiO PJ WX4130N）を用いて壁面に映像を提示した。学習者は映像の前に立ち、映像を観察しながら自由に身体を動かし、動作の練習を行った。

映像を投影した壁面から学習者までの距離は200cmとした。ただし、練習中は動作を伴うため多少の変動はありうる。投影された映像のサイズは縦210cm×横360cmであり、画角は水平方向で約103°である。映像内の演者のサイズは、映像内での奥行方向への移動があるため固定ではないが、最も大きいサイズの場合で、ビデオ映像の場合は高さ105cm、CGアニメーションでは60cmであり、視角ではそれぞれ30.1°、17.2°に相当する。

学習中の眼球運動測定は、瞳孔画像処理方式で眼球運動を検出する竹井機器工業株式会社製TalkEye Liteを使用し、30Hzで記録した。なお、実験室内の照度について、実験参加者の眼前照度は鉛直面で13.5lxであった。

CGアニメーションについては、ズームや視点の変更が可能であるが、映像の観察中にこれらの変更を行うためにPCでの操作を行うと、マウスやキーボードへの注視が生じ、映像に対する注視時間をビデオの場合と揃えることが難しくなる。このため、本研究ではズームや視点の変更は行わず、前方からの視点映像を活用して学習を実施した。

### 3.3. 手順

眼球運動計測装置を装着し、キャリブレーションを行った後に学習を行った。学習は、映像を「観察」するフェーズと、体を動かしながら学習をする「練習」フェーズに分割し、次の手順で実施した。

最初に30秒間映像を「観察」し、この際の注視行動を記録した。観察中は立位で静止した状態とし、身体をなるべく動かさないように求めた。身体動作により、注視点が注視対象からずれることを防ぐためである。

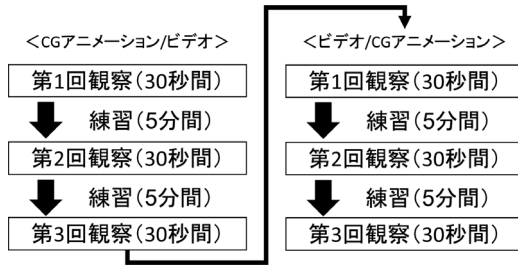


図3 実験の流れ

次に、5分間の「練習」時間を設けた。30秒の映像を10回繰り返して再生し、身体を自由に動かしながら演者の所作を習得するように求めた。その後、第2回目の「観察」を30秒間行い、注視行動を記録した上で、再度「練習」の時間を5分間設け、第3回目の「観察」を行った。

最後に、映像を再生し、映像内の演者に合わせて踊った後、出来栄えについて5段階の自己評価を行うことを求めた。

これらを1試行とし、ビデオ映像とCGアニメーションを用いた学習をそれぞれ行った。用いる2種類の映像は異なる場面とし、ビデオとCGアニメーションの実施順については偏りが生じないように配慮した。なお、実験参加者は学習する演目について事前知識がない男性7名であり、平均年齢は24歳であった。

#### 4. 結果

1秒間あたり30Hzで記録した眼球運動映像をコマ送りし、注視個所と注視時間を記録した。

注視対象については、使用した2つの場面において、左右の手や足を動かす回数が異なっており、四肢を左右に分けて分析すると場面間での偏りが大きくなることから、左右を統一し「頭」「上半身」「腕」「腰」「脚」の5種類とした。

注視時間については先行研究（神田・石田, 2000）を踏まえ、4コマ以上、すなわち133msec以上同一の領域内に注視点が位置した場合を注視と定義した。ただし、同一の領域内であっても、右腕から左腕に注視点が移動する等の、サッケードがある場合は連続した注視とはみなさず、サッケードの前後で別の

注視として、それぞれの注視時間を記録した。

図4は、CGアニメーション、ビデオについて観察の回数および注視対象別に注視頻度を図示したものである。映像と回数と部位を要因とした三要因分散分析を実施した。その結果、映像の主効果は有意ではなく( $F(1,6)=0.005$ ,  $p>0.05$ )、回数の主効果も有意ではなかった( $F(2,12)=0.403$ ,  $p>0.05$ )。部位の主効果は有意であった( $F(4,24)=14.034$ ,  $p<0.01$ )。映像×回数( $F(2,12)=0.282$ ,  $p<0.05$ )、映像×部位( $F(4,24)=1.804$ ,  $p>0.05$ )、回数×部位( $F(8,48)=2.070$ ,  $p>0.05$ )、映像×回数×部位( $F(8,48)=0.978$ ,  $p>0.05$ )の交互作用はいずれも有意でなかった。部位について、Bonferroni法を用いた多重比較を行った結果、腰に比べて腕及び頭の注視頻度が多く、脚に比べて腕の注視頻度が多かった。

図5は、各部位に対する1回あたりの注視時間の平均値(平均注視時間)である。映像と回数と部位を要因とした三要因分散分析を実施した。その結果、映像の主効果は有意ではなく( $F(1,6)=0.018$ ,  $p>0.05$ )、回数の主効果も有意ではなかった( $F(2,12)=0.395$ ,  $p>0.05$ )。部位の主効果は有意であった( $F(4,24)=10.606$ ,  $p<0.01$ )。映像×回数( $F(2,12)=1.715$ ,  $p<0.05$ )、回数×部位( $F(8,48)=0.579$ ,  $p>0.05$ )、映像×回数×部位( $F(8,48)=0.309$ ,  $p>0.05$ )の交互作用はいずれも有意でなかった。映像×部位( $F(4,24)=3.170$ ,  $p<0.05$ )の交互作用は有意であった。Bonferroni法を用いた多重比較の結果、腕、頭、上半身に比べて腰の

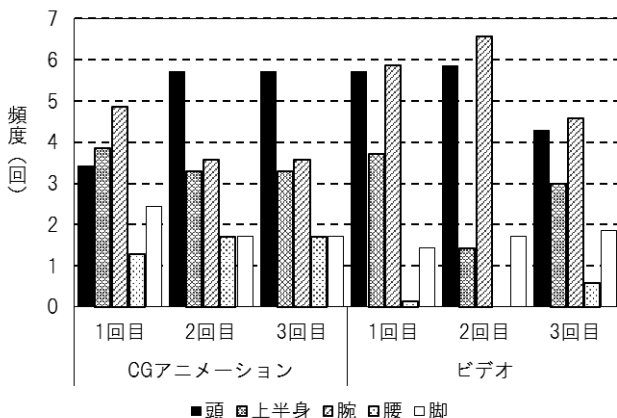


図4 注視頻度

注視時間は短かった。交互作用については、腰はCGの方が注視時間が長く、脚はビデオの方が注視時間が長かった。また、ビデオでは、腰に比べて腕と頭の注視時間が長かったが、CGではこうした差が見られなかった。

図6は各部位に対する各実験参加者の総注視時間の平均値を图示したものである。映像と回数と部位を要因とした三要因分散分析を実施した結果、映像の主効果 ( $F(1,6)=0.635, p>0.05$ )、回数の主効果 ( $F(2,12)=0.148, p>0.05$ )は有意

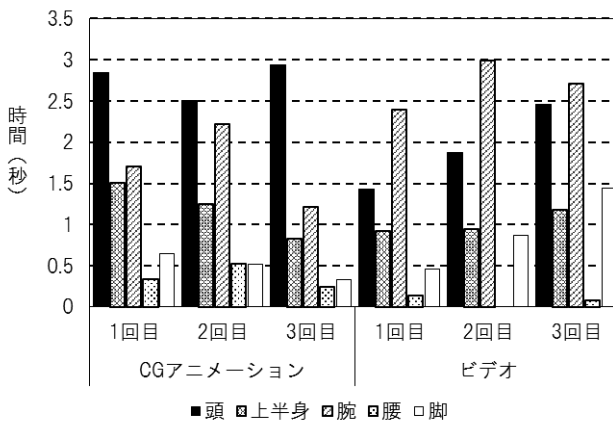


図5 平均注視時間

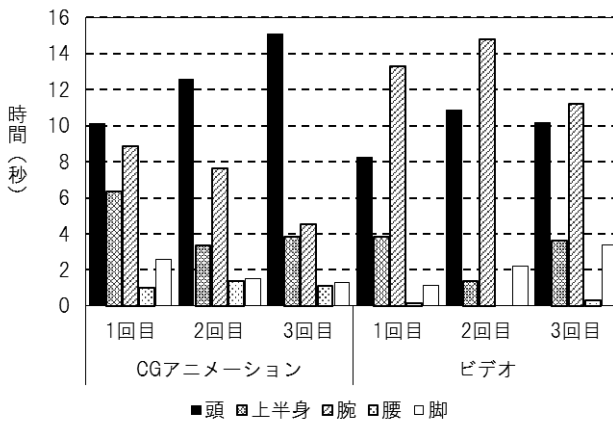


図6 総注視時間

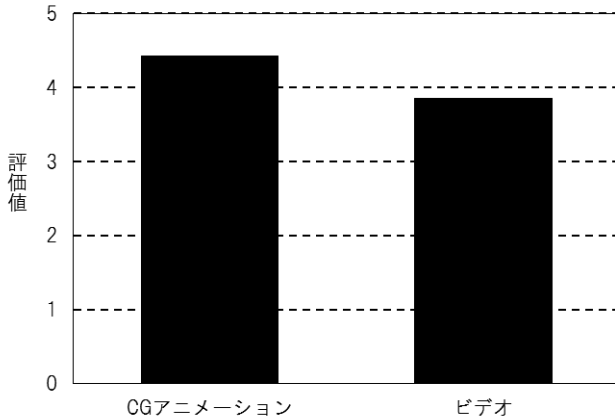


図7 踊りの自己評価

ではなかったが、部位の主効果は有意であった( $F(4,24)=10.210, p<0.01$ )。交互作用については映像×回数 ( $F(2,12)=4.656, p<0.05$ )、映像×部位 ( $F(4,24)=3.553, p<0.05$ )が有意であったが、回数×部位 ( $F(8,48)=1.702, p>0.05$ )、映像×回数×部位 ( $F(8,48)=0.485, p>0.05$ )は有意ではなかった。Bonferroni法を用いた多重比較の結果、腰、脚に比べて腕の総注視時間が長かった。交互作用については、CGアニメーションに比べてビデオの方が腕に対する総注視時間が長く、ビデオ映像では腰、脚、上半身に比べて腕の総注視時間が長かった。

学習後に映像を見ながら踊った際の自己評価について、映像を要因とした一要因分散分析を実施した。その結果、有意確率は10.3%となり、映像の主効果は有意ではなかった( $F(1,6)=3.692, p>0.05$ )。

## 5. 考察

同一の場面をCGアニメーション映像として提示した場合とビデオ映像として提示した場合の注視行動を調べた。CGアニメーションはモーションキャプチャによる記録に基づいて作成しており、本来は任意の方向から閲覧可能であり、ズームも可能であるが、今回はこれらの機能は使用せず前方から見た映像のみを用いた。したがって、CGかビデオかの違いはあるが、同様の映像を提示していたといえる。



注視頻度、平均注視時間、総注視時間について、映像の主効果は有意ではなかったが、平均注視時間、総注視時間では映像が関連する交互作用が有意であった。平均注視時間では、腰に対する1回あたりの注視時間がビデオよりもCGアニメーションの場合に長く、脚に対する1回あたりの注視時間はCGアニメーションよりもビデオの方が長かった。総注視時間では腕への注視がCGアニメーションに比べてビデオの方が長かった。また、ビデオ映像の場合は腕に対する総注視時間が、腰や脚、上半身よりも長かった。CGアニメーションの場合は部位別の平均注視時間や総注視時間に差が見られなかったが、図6を見る限りでは、頭部への総注視時間が最も長い傾向があった。

これらをまとめると、ビデオでは腕や脚に対する注視時間がCGアニメーションに比べて長く、CGアニメーションでは頭部や腰に対する注視時間が長い傾向があるといえる。今回利用した2場面は、頭部や腰について詳細に確認をする必要があるような細かな動きがあるわけではなかったことから、これらの部位への注視は、その部位の詳細な動きを確認しているというよりは、頭部や腰部を中心に上半身や下半身に幅広く注意を向けていると考えられる。したがって、CGアニメーションでは特定の部位を詳細に観察するのではなく、全体を見るような注視が比較的多かったといえる。一方、ビデオについては頭部への注視も多いが腕への注視時間が最も長く、腕の動きを詳細に確認していたと考えられる。

今回は、1回目、2回目、3回目の観察時において、注視行動は有意に変化しなかった。ただし、図6を見ると、CGアニメーションでは頭部への注視時間が回数を重ねるごとに長くなっており、他の部位への注視は全体的に減少傾向が見られた。一方で、ビデオの場合は回数に伴う注視行動の変化に一貫性が見出しにくく、3回目になって脚への注視時間が増加する一方、2回目に増加し、3回目に減少する部位や、2回目に減少し、3回目に増加する部位も見られた。

3回目の観察の後、実験参加者は映像に合わせて踊ることが求められた。出来栄への自己評価に有意な差はないものの、有意確率は10.3%であり、CGアニメーションの方が若干評価が高い傾向が見られた。CGアニメーションにおける頭部への注視が、頭部を中心に幅広く見る注視であるならば、観察の回数を重ねるごとに頭部への注視時間が増えたという結果を、踊りの自己評価が高

い傾向にあるという結果と組み合わせて考えると、CGアニメーションの方がより早く踊りを学習することができていたことを示唆する。

ただし、CGアニメーションとは異なり、ビデオ映像では演者が扇を所持していた。ビデオ映像では腕への総注視時間がすべての部位で最も長かったが、扇の動きを注視していた可能性もある。扇の動きを注視することで、手首の動きの理解にもつながり、より詳細に動作を再現することが可能になる。モーションキャプチャでは装着するセンサの数が限られるため、手指の詳細な動きを再現することは難しい。

まとめると、CGアニメーションでは踊りの全体像を早めに覚えることが可能であり、ビデオや手などの細かな動きの把握に向いていると考えられる。それゆえ、CGアニメーションを用いて全体像を把握し、その後にビデオを活用して細部を確認するというような、両者を併用する方法が望ましいと考えられる。

## 6. 結論と今後の課題

CGアニメーションとビデオ映像の観察時における注視行動を記録・分析した結果、注視行動には差が見られることが明らかになった。両者の注視行動の違いに着目すると、CGアニメーションは全体像の把握に有利であり、ビデオ映像は手指の動きのような細部の確認に優れている可能性があることが示唆された。

ただし、今回はモーションキャプチャの利点であるズームや視点の変更は実施していないことから、これらを行いながら学習を行う際の注視行動が記録できれば、CGアニメーションを用いた学習の利点がさらに見出される可能性がある。今後は上記を踏まえつつ、実験参加者数を増やししながらデータを収集することで結果の信頼性を高めるとともに、最終的な踊りの実施の際にモーションキャプチャを装着し、モデルの動作との差異を定量的に表現していく必要がある。

## 文献

- 神田直弥, 石田敏郎. 自転車乗用中における運転者の視認特性の検討. 日本交通心理学会平成12年度秋季大会(第62回)大会発表論文集. 2000, 79-82
- 佐藤克美, 海賀孝明, 渡部信一. 舞踊の熟達化を支援するためのモーションキャプチャ活用. 日本教育工学会論文誌. 2010, Vol.34(Suppl), 133-136
- 柴田傑, 玉本英夫, 海賀孝明, 横山洋之. 身体動作の3次元計測によるリアルタイム舞踊学習支援システム. 日本バーチャルリアリティ学会論文誌. 2012, Vol.17, No.4, 353-360
- 白鳥貴亮, 中澤篤志, 池内克史. モーションキャプチャと音楽情報を用いた舞踊動作解析手法. 電子情報通信学会論文誌. D-II, 情報・システム, II-パターン処理. 2005, J88-D-II(8), 1662-1671
- 玉本英夫. 民俗芸能・伝統芸能をモーションキャプチャで記録する. 文部科学省私立大学研究ブランディング事業日本遺産を誇る山形県庄内地方を基盤とした地域文化とIT技術の融合による伝承環境研究の展開. 2019, 23-29
- 八村広三郎. 高精度デジタル化のための最新技術. 伝統舞踊のデジタル化. 映像情報メディア学会誌. 2007, Vol.61, No.11, 1557-1561
- 松本奈緒, 三浦武, 海賀孝明, 柴田傑, 齋藤龍一, 桂博章, 玉本英夫. 秋田の盆踊りの学習におけるデジタルコンテンツを用いた学習支援の効果と限界—モーションキャプチャ技術を応用した学習支援装置作成の試み—. 舞踊学. 2011, Vol.34, 1-10
- 三浦武, 水戸部一孝, 湯川崇, 海賀孝明, 谷口敏幸, 吉村昇. 舞踊動作データの因子分析による基本動作の抽出. 情報処理学会論文誌. 2008, Vol.49, No.9, 3302-3315
- 湯川崇, 海賀孝明, 長瀬一男, 玉本英夫. 舞踊符による身体動作記述システム. 情報処理学会論文誌. 2000, Vol.41, No.10, 2873-2880