

ダンスの質的評価に関する視線の分析

劔持 若菜¹ 河瀬 彰宏¹

概要: 本研究では、ダンス評価時における視線計測実験を実施し、AIST Dance Video Database の映像を用いてダンス経験者と未経験者のダンス評価時の注視傾向を比較した。注視部位の時間に対して統計的検定を実施した結果、注視部位とダンス経験に関連があること、経験者が頭部や手先を注視する傾向にあること、未経験者が胴体など身体を中心に近い部分を注視する傾向があることを明らかにした。

Analysis of Gaze Measurement for Qualitative Evaluation of Dance

1. はじめに

1.1 ダンスのあり方の変化

近年、私たちの生活の中でダンスがより身近なものになっている。2012 年に実施された中学校の学習指導要領の改訂に伴って、体育の授業でダンスが必修となったことで、中学生がダンスに触れる機会が増加した。学生のクラブ活動・部活動としてのダンスについては、2000 年代に数多くの大会や競技会が創設された。2008 年に開始した日本高校ダンス部選手権、2015 年に開始したマイナビ HIGH SCHOOL DANCE COMPETITION などがある。このような大会の中には、株式会社エイペックスや株式会社マイナビなど大手企業が主催しているものもあり、ダンスは表現活動としてだけでなく、競技としても広く周知されていることがわかる。

株式会社ヤマハミュージックジャパンによると、2015 年のダンスを習っている小学生の割合はその親世代と比較しておよそ 5 倍になっており、大幅に増加していることがわかる [6]。さらに、ダンスの普及は児童生徒世代だけでなく、様々な世代に広がっている。一般社団法人ストリートダンス協会は、ストリートダンス人口は 2020 年現在、600 万人を超えていると発表している [4]。政府統計の総合窓口 (e-Stat) による「統計でみる都道府県・市区町村のすがた (社会・人口統計体系)」(1999, 2005, 2008, 2015)[11] によると、公共のダンス場数が 1999 年から 2015 年で 4 倍以上に増加していることや、民間のダンス場数が 2008 年に

1,000 箇所を超えたことがわかる。このように、2000 年代以降、ダンスがより一般的で身近なものになっているとともに、ダンスが現代文化に根付いてきたことがわかる。

さらに、SNS や動画投稿サイトの拡張とサービスの充実により、誰もが気軽にコンテンツの発信者となることができるようになった。投稿された動画は、閲覧者が共感することによって SNS 上で拡散され、「いいね」などの機能によって評価される。動画の共有に多く用いられている SNS のうち、TikTok や Youtube は「いいね」や「高評価」「低評価」などが数値として表示される。この機能により、投稿者はより高く評価される動画を投稿しようとするようになっていくと考えられる。このように、現代においてダンスは簡単に他人の目に触れ、評価されることが常になっているため、より高い評価を得るダンスを作り出すことが必要とされている。

1.2 関連研究

Ofii et al.(2012) は、音楽とダンスの間の多対多の関係に基づいた振付の生成手法の提案を目的として、ダンス映像と楽曲の小節間の相関関係と、映像から映像への遷移確率の観点からダンス映像間の相関関係に基づくダンス映像への多対多のマッピング手法を考案した。生成された振付に対して、評価実験を行った結果、全体の振付の音楽との一致度が高い場合、観客は振付のバリエーションに魅力を感じていることを明らかにした [8]。

Qi et al.(2019) は、MFCC を用いてダンス動画から音響特徴量を、映像からポーズ画像を取得し、振付の自動生成に利用した。生成された振付について、10 名の評価者 (ブ

¹ 同志社大学
Doshisha University, 1-3 Tatara Miyakodani, Kyotanabe-shi,
Kyoto 610-0394, Japan

口のダンサー 5 名、ダンスについての専門知識を持たない 5 名) に、身体の動きが自然であるか、音楽に合っているか、バリエーションが豊富であるかといった観点で評価させた。その結果、上記の観点において、高い評価を得るための合成モデルを明らかにした [9]。

Fukayama and Goto(2015) は、音楽コンテンツを利用して振付を自動生成する手法を提案した。ダンスモーションデータベースより取得した楽曲の音楽的構造、コード、ビートなどの音響特徴量を利用して新たな振付を生成した。生成された振付の自然さについて、ダンスの専門知識の無い 5 名に 2 つの振付からより自然に感じるものを選択させることで評価した。その結果、振付を生成する際、データベース内の連続した動きを連続したまま利用することでより自然な動きであると評価されることを明らかにした [2]。

Amelynck et al.(2014) は、動作と身体の動きの関わりにおける個人差について、楽曲を身体を用いて表現する際に選択される動作を計測する実験を行った。36 名の実験対象者に、提示した楽曲を体を使って表現してもらい、その動作をモーションキャプチャーを用いて記録した。その結果、音楽を身体で表現する場合、テンポなどの基本的な音楽構造の表現のために胴体が、より上位の構造的要素の表現には自由度の高い手などが用いられていることを明らかにした [1]。

阪田ら (2006) は、舞踊における感性情報の基礎資料を得ることを目的として、舞踊の視線計測実験を実施した。実験協力者は、ダンス未経験者 3 名、ダンス経験者 6 名、ダンス指導者 5 名の計 14 名であった。実験の結果、鑑賞者の視線には法則が存在していることや、舞踊習熟度と視線行動に関連があることを明らかにした [10]。

以上の関連研究より、ダンスモーションの自動生成では、音楽と合っているかという点が重要視され、楽曲のリズムやスタイルに合った生成手法が提案されてきたことがわかる。また、人がダンスを主観的に評価する際には、音楽との関係の他にも、振付そのものや身体の動きなども用いられている。さらに、音楽を身体で表現する場合には、音楽のどのような点を表現するかによって、利用する身体の部位が異なることが明らかにされている。このことから、ダンスを鑑賞する場合でも、音楽のどのような点に注目しているかによって、重要視する身体の部位が異なることが予想される。また、舞踊鑑賞時の視線行動とダンス経験の関連が明らかにされていることから、ダンスを評価する際の視線行動もダンス経験と関連があることが予想される。しかし、ダンスに対する主観的評価において、身体のどの部分を注視するかという点で生じる個人差は明らかにされていない。

そこで、本研究はダンスを質的に評価する際にダンス経験者と未経験者がそれぞれ身体のどのような点を注視しているのかを明らかにすることを目的とする。

2. 分析方法

2.1 視線計測実験

本研究では、1 年以上のダンス経験者 11 名と未経験者 9 名を対象に、AIST Dance Database[13] より取得した 5 映像を用いて視線計測実験を実施した。5 映像は、いずれもダンサーは 1 名でジャンルは LA-style HIPHOP であるものを用いた。視線計測実験は、Gazepoint GP3 HD Desktop で行い、下記の手順で実施した：

- (1) 5 種類のダンス映像を 1 回ずつ提示し、鑑賞時の視点移動を記録した。
- (2) 2 回目の視聴では、視聴後にダンスに対して評価を行うことを事前に伝えた上で、5 種類のダンス映像を 2 回連続で提示し、視点移動を記録した。
- (3) 各ダンス映像の視聴後に全国高等学校ダンスドリル選手権 [12] が公開しているジャッジシートのうち、HIPHOP 部門の審査で用いられている評価項目を回答させた。

2.2 特徴量の算出

本研究では、各映像から記録した注視点（単位時間あたりの x 座標と y 座標の値）と、それらの前後のイベントから、次の 5 つの特徴量を算出した：(i) 移動速度、(ii) 注視時間、(iii) 移動距離、(iv) 水平移動距離、(v) 垂直移動距離。ただし、福田ら (1996)[3] に倣い、眼球移動速度が 10deg/sec 以上の移動と、まばたき等によって正しく計測できなかった注視点については、無効データとして分析対象から除外した。

2.3 ダンス経験と視点移動の比較分析

本研究では、算出した視点移動データの 5 項目について、以下の 2 つを要因とする 2 標本 Wilcoxon 検定^{*1}をそれぞれ実施した：

- 実験協力者のダンス経験の有無の差異
- (1) 鑑賞時の視線と (2) 評価時の視線の差異

2.4 ダンス経験と注視部位の比較分析

実験で使用した映像に対して、Hidalgo(2020) による Openpose[5] を用いて計 33 箇所の骨格検出を実施した (図 1)^{*2}。検出のサンプリングは 25Hz で、各ポイントの x 座標、 y 座標とその精度を取得した。取得した各検出ポイントの座標と、実験協力者の視点座標との Euclid 距離を算出し、最も距離が近い部位を注視点としてマークした。こ

^{*1} 各項目のデータの分布に正規性が仮定できないため、ノンパラメトリック検定である 2 標本 Wilcoxon 検定を用いて中央値による 2 群の比較を行った。

^{*2} 検出したポイントは、関節位置であるキーポイント 18 箇所、身体の部位の重心を示す部位重心が 14 箇所、身体の重心 1 箇所であった。

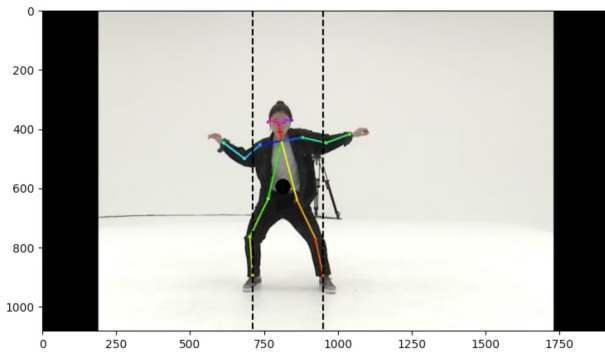


図 1 骨格検出したポイント

表 1 鑑賞時と評価時における 5 項目の中央値

	(1) 鑑賞時		(2) 評価時	
	(a) 経験者	(b) 未経験者	(c) 経験者	(d) 未経験者
(i) 注視時間	0.0166	0.0169	0.0166	0.0171
(ii) 移動距離	0.0138	0.0158	0.0141	0.0164
(iii) 水平移動距離	0.0040	0.0047	0.0042	0.0049
(iv) 垂直移動距離	0.0104	0.0121	0.0107	0.0123
(v) 移動速度	0.7557	0.8643	0.7812	0.8882

ここでは、阪田ら (2006)[10] に倣い、加藤・福田 (2002)[7] の示した基準を用いて、133ms 以上視線が停留した場合を注視した状態であると定義した。本分析では、骨格検出を 25Hz(0.04sec/frame) で記録したため、4frame 以上同一部位がマークされた場合を注視状態と定義して、部位ごとの注視回数と注視部位を記録した。

この記録に対して、ダンス経験者とダンス未経験者の差異を χ^2 検定の残差分析によって比較することで、注視傾向が特徴的な部位を明らかにした。キーポイント、部位重心それぞれの同じ部位の左右を合算したもの、キーポイントの目・鼻・耳を合算して顔としたものを作成し、検出ポイント全て・左右を合算した数値・顔を合算した数値（キーポイントのみ）のそれぞれに対して χ^2 検定を実施した。

3. 分析結果

表 1 は、ダンス映像の (1) 鑑賞時と (2) 評価時における 5 項目に対するダンス経験者と未経験者の中央値の一覧である。例えば、表 1(c) は、視線計測実験の (2) 評価時においてダンス経験者から収集した 5 項目に対する観測値の中央値を表している。以下に、表 1(a) から (d) に対して統計的検定を実施し、ダンス経験の有無による差異を検証した結果を述べる。

3.1 視点の移動に関する分析結果

3.1.1 経験の有無による比較

(1) 鑑賞時の視線計測データから算出した 5 項目について、ダンス経験の有無における差異を検証するために、表 1(a)(b) 間で 2 標本 Wilcoxon 検定を実施した。同様に、(2) 評価時についてもダンス経験の有無における差異を検証す

表 2 χ^2 検定の結果

変数	ポイント	χ^2 統計量	p 値	V
注視回数	A	143.45	< 0.05	0.084
注視時間	A	74.875	< 0.05	0.097
注視 frame 数	A	1871.9	< 0.05	0.097
注視回数	B	134.11	< 0.05	0.081
注視時間	B	71.13	< 0.05	0.095
注視 frame 数	B	1778.3	< 0.05	0.095

るために、表 1(c)(d) 間で 2 標本 Wilcoxon 検定を実施した。検定の結果、いずれの関係においても p 値が有意水準 $\alpha = 0.05$ を下回ったことから、帰無仮説は棄却され、ダンスの鑑賞時および評価時の視線の動きはダンス経験の有無によってそれぞれ異なっていることが明らかになった。

3.1.2 鑑賞時と評価時の比較

次に、ダンス経験の有無によって、(1) 鑑賞時と (2) 評価時における差異が生じているのかを検証するために、表 1(a)(c) 間および表 1(b)(d) 間で 2 標本 Wilcoxon 検定を実施した。検定の結果、未経験者の (i) 注視時間以外の 4 項目において、 p 値が有意水準 $\alpha = 0.05$ を下回ったことから、帰無仮説は棄却され、ダンス経験の有無に関わらず、(ii) 移動距離、(iii) 水平移動距離、(iv) 垂直移動距離、(v) 移動速度は、鑑賞時よりも評価時の方が値が大きいことが明らかになった。

3.2 注視部位とダンス経験に関する比較分析

骨格検出により取得した身体座標データと視点座標データを用いて χ^2 検定を実施した結果と、算出した Cramer's V のうち、数値が高かった部位重心を用いた結果を表 2 に示す。部位ごとの注視回数、注視時間、注視 frame 数の全てにおいて p 値が 0.05 を大きく下回ったため、帰無仮説は棄却され、注視部位とダンス経験には関連があるといえる。また、全ての検定において Cramer's V が $0.07 \leq V \leq 0.10$ であるため、注視部位とダンス経験の関連は弱いことが明らかになった。表 2 において、身体座標の検出ポイント A は、部位重心の検出ポイントをそのまま用いたものであり、B は、部位重心の左右を合算した数値を用いたものである。

表 3 から表 5 に、実施した χ^2 検定のうち、キーポイント (顔・左右合算) と部位重心 (左右合算) における各部位の注視回数、注視時間、注視 frame 数の観測値を示した。残差分析の結果、統計的に有意に多いとされたものを表中に太字で示している。

以上の結果から次の 3 点が読み取ることができた：

- 注視回数、注視時間、注視 frame 数のいずれにおいても、ダンス経験者は、未経験者よりも顔や頭部を有意に多く (長く) 注視する。
- 注視回数、注視時間、注視 frame 数のいずれにおいても、ダンス未経験者は、経験者よりも首や肩を有意に

表 3 注視回数

	頭部	胴体	大腿部	脚	足
経験者	3,565	721	816	568	782
未経験者	2,552	824	626	409	719
	腕	前腕	手	身体重心	
経験者	2,709	518	709	648	
未経験者	2,704	407	522	513	

表 4 注視時間 (sec)

	頭部	胴体	大腿部	脚	足
経験者	2217.4	185.9	282.4	205.1	407.1
未経験者	1448.0	197.6	228.4	137.8	381.6
	腕	前腕	手	身体重心	
経験者	711.4	106.3	154.2	198.1	
未経験者	717.3	88.5	109.1	137.4	

表 5 注視 frame 数

	頭部	胴体	大腿部	脚	足
経験者	55,434	4,647	7,061	5,127	10,177
未経験者	36,199	4,941	5,711	3,446	9,540
	腕	前腕	手	身体重心	
経験者	17,785	2,658	3,855	4,953	
未経験者	17,932	2,213	2,727	3,434	

多く (長く) 注視する。

- 注視回数, 注視時間, 注視 frame 数のいずれにおいても, ダンス未経験者は, 経験者よりも部位重心の胴体や腕を有意に多く (長く) 注視する。

4. 考察

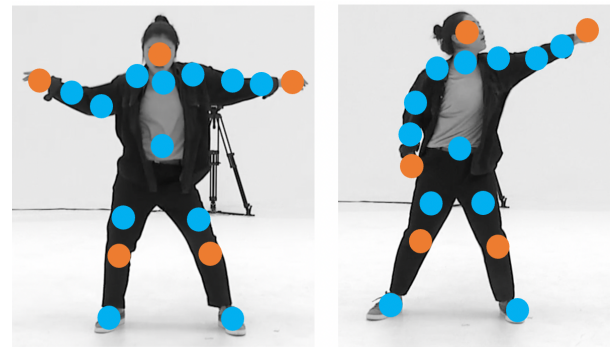
4.1 ダンス経験の有無による視点の移動の差異

ダンス経験の有無による 2 標本 Wilcoxon 検定の結果より, ダンス経験者と未経験者では, ダンス鑑賞時, ダンス評価時でともに視点の動きが異なるということが明らかになった。したがって, ダンス経験者は効率よく身体を全体的に注視して動きを評価している傾向が読み取れた。一方で未経験者は, 1 箇所を長く注視し, 大きく視点を移動させていることから, 1 つの動きを把握するのにある程度の時間を要することがわかった。

また, 鑑賞時と評価時を比較した 2 標本 Wilcoxon 検定の結果より, 鑑賞時には注視時間が長いことから 1 箇所を長く注視しており, 速度や距離の値が小さくなっていることから, 注視した部位を起点としてその周辺に徐々に視点を動かしていく傾向が読み取れた。一方で評価時には, 注視時間が短く, 移動速度や移動距離の値が大きくなっていることから, 様々な部位を評価しようとするために視点の移動が大きく, また注視回数も多くなっていることがわかった。

4.2 注視部位の傾向とダンス経験

χ^2 検定と残差分析の結果より, 経験者と未経験者ではダンス評価時に注視している部位に違いがあることが明らか



オレンジ：ダンス経験者 シアン：ダンス未経験者

図 2 ダンス経験の有無による注視部位の差異

になった。具体的な部位を図 2 に示す。図 2 において, オレンジの点は経験者が優位に多く注視していた部位を, シアンは未経験者が優位に多く注視していた部位を表している。これより, 経験者は顔や手先, 膝などの身体先端を特に注視していたことがわかる。それに対して未経験者は首, 腕, 胴体などの身体中心に近い部位をよく注視していたことが明らかになった。その要因として, 経験者は自身のダンス経験からダンスを評価する際には身体先端の部位を注視する必要があると考えて細かい動作まで把握しようとしたということが考えられる。一方, 未経験者はダンスを評価する際に注視すべき点がわからず, 目で見てわかりやすい動作を把握して評価しようとした結果, 身体中心部分の注視傾向が強くなったのではないかと考えられる。また, 経験者で頭部の注視傾向が強く見られた理由として, 経験者がダンサーの表情まで含めてダンスの表現であると考えた結果であると考えられる。本研究において, ダンス経験者はダンスを評価する際に, 抽象的な表現のひとつとしてダンサーの表情を重要視したのではないかと考えられる。

さらに, 経験者では頭部以外に注視傾向が強くなった部位が少なかったことについて, 評価者それぞれが主体的に, ダンスを評価するための部位を決定し, 注視することで, 経験者が評価する際に共通して注視する部位が少なかったのではないかと考えられる。一方で, 阪田ら (2006)[10] はダンス未経験者は舞踊を鑑賞する際に頭部に強く着目していたことを明らかにしたが, 本研究では, 未経験者がダンスの評価時に頭部をよく注視しているという結果は得られなかった。その理由として, ダンスを評価するよう伝えた場合, 今までに経験のない作業を行う必要があったため, 事前に伝えていた評価項目に対してきちんと評価しようとする意識が強まり, 比較的わかりやすく動き, 大きく動く首よりも下をより多く見ていたのではないかと考えられる。

ダンス経験者が, 身体先端の部位に注意していた理由として, 自身のダンス経験から, ダンサーがより抽象的な表現を行う際に手先など自由度の高い部位を用いていると

判断して、手先などに注目してダンサーの表現を評価しようとしていたことが考えられる。一方で未経験者は、ダンスの表現について理解しようとしていたのではなく、曲とダンスのリズムが一致しているかといった基本的な音楽構造との関わりをより評価しようとしたため、経験者と比較して胴体の注視傾向が高くなったと考えられる。つまり、経験者と未経験者の注視傾向から、経験者はダンサー個人の主体的な表現について、未経験者はテンポに合っているかといった基本的な概念について理解をしようとしていたと考えられる。

また、3.2 節より、キーポイントを用いた χ^2 検定における Cramer's V は全て 0.09 未満であり、全ての検定のうち、Cramer's V の上位であったものは部位重心を用いたものであった。このことから、ダンスを評価する際には、身体を関節同士の点と点で見ているのではなく、各部位ごとに大まかなエリアとして認識しているのではないかと考えられる。しかし、 $0.07 \leq V \leq 0.10$ であることから、ダンス評価時の注視部位とダンス経験の関連が強いと言うことはできなかった。以上より、ダンス評価時の注視部位はダンス経験によって完全に異なるのではなく、共通している部分があることが読み取れた。

5. 今後の課題

本研究では、ダンス経験とダンス評価時の注視傾向の関連を検証した。ダンス経験者と未経験者では、ダンスを質的に評価する際に、注視する部位の傾向に差があり、ダンス経験者は頭部や手先を、未経験者は胴体や首、腕といった身体を中心に近い部位を注視する傾向にあることが明らかになった。

今後の課題として、Watanabe et al.(2011)[15] は、人物に対する注視傾向に性差があることを明らかにしていることから、性別の違いを考慮した比較分析を検討することが考えられる。また、HIPHOP 以外のスタイルとの比較を行うことによって、ジャンル間での差異が得られることも予想される。また、土田ら (2020) は、振付の構造解析については未解明であると述べている [14]。そこで、本研究で用いた骨格検出を用いた振付の構造解析や、実験後のアンケートを解析することにより、注視した意図が得られることが予想される。

謝辞 本論文を作成するにあたり、多大なるご指導を賜りました同志社大学文化情報学部の阪田真己子教授に厚く御礼申し上げます。

参考文献

- [1] Amelynck, D., Maes, P.J., Martens, J.P., and Leman, M.: Expressive body movement responses to music are coherent, consistent, and low dimensional, *IEEE Transactions on Cybernetics*, **44**(12), pp.2288–2301 (2014).
- [2] Fukayama, S., and Goto, M.: Music content driven automated choreography with beat-wise motion connectivity constraints, *Proceedings of SMC*, pp.177–183 (2015).
- [3] 福田亮子・佐久間美能留・中村悦夫・福田忠彦: 注視点の定義に関する実験的検討, *人間工学*, **32**(4), pp.197–204 (1996).
- [4] 一般社団法人ストリートダンス協会: 入手先 (<https://www.streetdanceyoukai.com/>) (2021.02.15)
- [5] Hidalgo, G.: 入手先 (<https://github.com/CMU-Perceptual-Computing-Lab/>) (2021.02.15)
- [6] 株式会社ヤマハミュージックジャパン: 親世代と子供世代を比較するとダンスを習っている人が約 5 倍に増加! 入手先 (<https://prtmes.jp/main/html/rd/p/000000075.000010414.html>) (2021.02.15)
- [7] 加藤貴昭・福田忠彦: 野球の打撃準備時間相における打者の視覚探索ストラタジー, *人間工学*, **38**(6), pp.333–340 (2002).
- [8] Ofii, F., Erzin, E., Yemez, Y., and Tekalp, A.M.: Learn2Dance: Learning statistical music-to-dance mappings for choreography synthesis, *IEEE Transactions on Multimedia*, **14**(3), pp.747–75 (2012).
- [9] Qi, Y., Liu, Y., and Sun, Q.: Music-driven dance generation, *IEEE Access*, **7**, pp.166540–166550 (2019).
- [10] 阪田真己子・徳池雅子・原田純子: 舞踊鑑賞者の眼球運動に着目した感性情報処理の試み—アイカメラを用いた鑑賞者の視線分析, *じんもんこん 2006 論文集*, pp.167–172 (2006).
- [11] 政府統計の総合窓口: 社会・人口統計体系 入手先 (<https://www.e-stat.go.jp/regional-statistics/ssdsvview>) (2021.02.15)
- [12] 全国高等学校ダンスドリル選手権大会 <https://www.dancedrilljapan.com/competition/championship.html> (2021.02.15)
- [13] Tsuchida, S., Fukayama, S., Hamasaki, M., and Goto, M.: AIST Dance Video Database: Multi-genre, multi-dancer, and multi-camera database for dance information processing, *Proceedings of the 20th International Society for Music Information Retrieval Conference (ISMIR 2019)* (2019).
- [14] 土田修平・深山寛・濱崎雅弘・後藤真孝: AIST Dance Video Database: ダンス情報処理研究のためのストリートダンス動画データベース, *情報処理学会研究報告*, **2020-MUS-126**(9), pp.1–10 (2020).
- [15] Watanabe, K., Matsuda, T., Nishioka, T., and Namatame, M.: Eye gaze during observation of static faces in deaf people, *PLoS One*, **6**(2), e16919. (2011).