

能楽映像記録の舞動作解析手法

内記 綾子

津田塾大学大学院理学研究科

青柳 龍也

津田塾大学大学院理学研究科

能の理解には知識と経験が必要とよく言われる。知識なく能を観ても理解は難しい。第2次大戦前までの日本では、謡や仕舞を学ぶ者も多く、家庭で自然に能の知識と経験が得られていたと聞かすが、戦後は謡や仕舞を学ぶ者は減少し、自然に能の理解を得ることは難しくなっている。書籍やインターネットで知識を得ることはできても、能を経験する機会は少ない。本研究は、現存する能楽映像記録を解析し、得られた知識を公開することにより、能を深く知りたい観客を支援することを研究目的としている。能の舞は基本的な型の組合せから構成されている。本論文では、既存の能楽映像記録を解析し、立体的な舞の動きとこれに照合する能の舞の型を推定する手法を提案する。過去の名演を理解することは能の経験に通じる。

The Method to Analyze Noh Dance

Ayako Naiki
TSUDA College

Tatsuya Aoyagi
TSUDA College

Noh is said to be difficult to understand without the appropriate background knowledge and experience. Thus, it is usually taught to children by their parents who have learned it from their own parents. Before the Second World War, there were a good number of learners of Utai and Shimai. However, after the war, the number parents who teach Utai and Shimai to their children had decreased. The aim of this research is to assist Noh audiences with intellectual appetites by analyzing existing and obtainable Noh movie contents and by sharing acquired knowledge. We propose a method that estimates three-dimensional Noh dance motions from a two-dimensional Noh movie and to recognize the forms of the Noh dance. The result of the analysis affirms the variety of Noh performances. This research further provides a way to understand the characteristics of the performances of each Noh actor which allows deeper discovery and pleasure of watching Noh.

1. まえがき

能はユネスコの無形文化遺産に指定された日本の伝統芸能の一つである。しかし能で用いられる言語は中世の日本語であり、独特の節で謡われる。このため、能の音声も聴いても内容や表現を理解する事は難しい。一方、能の視覚面では、舞は基本的な型の組み合わせとして定義されるが、抽象的な意味を内包しており、知識なくては舞の深い理解は不可能である。能を深く味わうには、知識と経験が必要であると言われている。

聴覚面の謡や視覚面の仕舞はかつて日本中に広く普及した芸能であった。現在では、謡や仕舞の学習人口は激減している。かつて、家庭で自然に伝承されてきた能の知識や経験は殆どの現代日本人には無縁のものとなっている。

能の知識は書籍やインターネットから得ることができる。数多くの舞台を観る事で経験は可能だが、能をよく知り、疑問に答えてくれる者が周囲にいない状態で舞台を観ても、疑問の解消は難しい。

一方、能には貴重な映像記録が存在する。名演の記録には能の理解に繋がる情報が含まれていることが推測できる。

本研究の最終的な目標は、こうした映像記録を活用し、能の新たな普及に役立てることである。そのための解析ツールや解析結果の公開を考えている。本論文ではその第一歩として、入手可能な既存の能楽映像記録から能の舞動作を解析し、舞がどのような型から構成されているかを推定する手法を提案する。

解析結果は、同じ型から構成されているにもかかわらず、能楽師により舞は多彩な表現が取られていることを示唆している。それぞれの舞の特徴を解析する事で、型と舞表現の関係を推定出来れば、伝承芸術である能の深い理解に通じると考えられる。

2. 能の舞の特徴

能楽における舞は、予め定められた型の組合せから構成されている。その時々に取りべき型は型付書[1]に謡の詞章と併記されており(図1)、基本的に能の舞は謡の詞章との対応関係が定められている。型付書の表記は流派により異なる他、時代による変化もある。本論文では、喜多流の初学者用の動線図解(図2)付きの型付書を用いて、舞の動作を検証するものとする。図1中に示される番号と図2中の番号は同時刻を意味している。

個々の型は名称により識別される。図で型が示されるケースもあるが、基本的に型は名称と言葉による動作解説により定義されている。

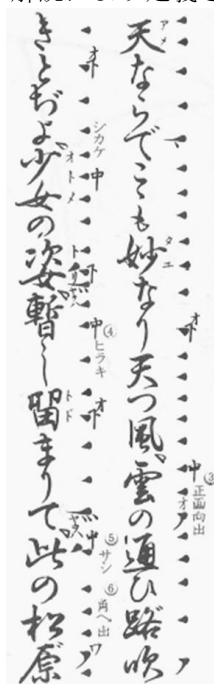


図1 型付[1]
Figure 1 Noh Dance Indication

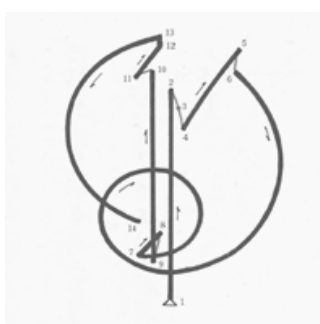


図2 動線図解[1]
Figure 2 Dance Indication by Orbit

型付書に記載されるのは型の名称のみであり、一般的には、型付書からは舞動作の予想は難しい。

謡本や型付け書に記載されないルールも存在する。例えば能の舞においては、基本的な姿勢は「カマエ」という型を口頭で伝承される。歩く際は基本的には、足を高く上げず、摺り足で歩く。型から型に移行する際も、原則は、無駄な動きを避ける事にある。

型や原則は流派により異なる。同じ曲も流派や演者により多様な表現となる。囃子との渡り合いもある。流派や型の伝承形態が能の伝承を分散し、理解を難しくしている。

幸い、過去の公演の一部は撮影され、動画として入手可能であるが、貴重な記録としてのみ

に留まっている。2次元の情報ではあるが、これらの解析が進めば、伝承を難しくしている部分の表現や理解に繋がる事が期待できる。

3. 研究対象

解析対象は、能、舞囃子、または仕舞を単一の視点から撮影した通常の2次元動画とする。

一般的に3次元の動作を計測する際には、1視点映像だけでは解は不定となる。この問題に対し、本研究では、能の舞の基本的な型や原則についての情報を活用することから始め、3次元動作解析と推定手法開発へ展開する。

4. 既存手法

岩月ら[2]は能楽の型付け書を利用して能の舞の動作を解析した事例を報告している。しかし、舞の計測にはモーションキャプチャーを用いており、特定の場における舞の解析になっている。

他にもモーションキャプチャーを用いた舞踊動作の解析事例が多く存在する。しかし本研究は、既存の能楽映像記録から、個々の舞の特徴を捉えることを目的としているため、モーションキャプチャーを前提とした解析手法は使用できない。

三上ら[3]は、遮へいなどにより追跡対象を見失った場合にも対応可能な追跡手法として、メモリアベースパーティクルフィルタ提案している。

しかしこの方法で全身動作の追跡を行うとなると、計算量が増大する。能の型付け書の情報を用いて計算量を削減できる可能性はあるが、その適用は容易ではない。

5. 提案手法

能楽映像記録の2次元画像より3次元の舞動作を推定するためには、映像以外の情報が必要である。また、能楽の衣装は、和服を纏って舞われるため、肘や膝の位置の観測は難しく、推定する他ない。

不足する情報を補うため、本手法は、能の舞の定義や規則を活用したいが、動きを文書化したものは少なく、師弟間などの伝承に頼ることになる。

以下に、本手法で使用する人間モデルと解析手順について述べる。

5.1 人間モデル

能の舞は型の規定はあるものの、具体的な表現方法は個々の能楽師が決定する。舞表現の多様性に対応できる認識手法が必要である。

本研究では、舞の型の識別に必要な最低限の情報を持つ型定義モデル(図3左)と、人の3次元的な姿勢の規定に必要な最小限の情報を持つ基本姿勢モデル(図3右)の2モデルを併用

する。図 3 における黒丸は端点または関節ノードを示し、白丸は位置と方向を規定する原点である。



図 3 型定義モデル (左) と基本姿勢モデル (右)
Figure 3 left : Form Definition Model
right : Basic Posture Model

(1) 型定義モデル

能衣装や和服でも比較的観測のしやすい 5 点のノードのみから構成されるモデル。型を識別できる最小限の情報を持つ。各ノードの局所座標系を表 2 に記す。各ノードの座標値は個別の閾値を持ち、極めて粗い離散値を取る。

図 3 は、能の基本型と言える「カマエ」に相当する。この時、型定義モデルの全ノードは全座標が 0 となる。

表 1 型定義モデルノード別座標値
Table 1 Node Value of Dance Form Definition Model

ノード	x座標(左右)	y座標(前後)	z座標(高さ)
首	0 固定	-1: 原点より後	-1: 屈む
		0: 原点の上	0: 立つ
		1: 原点より前	1: 座る
手	-1: 体の内側	-1: 首より後	0: 下がっている
	0: 内側でも外側でもない	0: 首と同じ	1: 上がっている
	1: 体の外側	1: 首より前	
足	0 固定	-1: 後に下がる	0: 床
		0: 同じ位置	1: 床より高い
		1: 前に出る	

(2) 基本姿勢モデル

基本姿勢モデルは頭頂、首、肩、肘、手、胸、腰、膝、踵、足先をノードとし、図 3 右図における黒線をエッジとするグラフ構造を持つ。但し、両肩と首をつないだ図形は、二等辺三角形になるものとする。

原点は床に接した両踵の midpoint とし、演者の前方向を x 軸方向、右手方向を y 軸方向、鉛直上向きを z 軸方向に置いた座標系において、各ノードは x, y, z 座標共に身長を 1.0 に正規化した実数値を持つ。身長データを乗じることにより、容易に実寸データを求めることができる。

人体のプロポーシヨン比はノード間距離の比に相当する。現在は能の型の認識を目標に置いたため、高精度の計算結果は不要であり、プロポーシヨン比は定数とした。

5.2 舞動作解析手法

2次元の画像より能の特徴を活用し、3次元動作の推定を行う。まずは能の型や規則を登録する。

(1) 情報登録

- 1) 型の登録 型定義モデルのノードの一連の運動の組として、舞の型を登録する。
- 2) 謡曲の登録 謡曲の名称、並びに詞章の読み方を登録する。
- 3) 型付けの登録 型付書に記載された型名や立ち位置を、謡曲の詞章との関係と共に登録する。

以上登録情報から、謡曲の詞章各部分における舞の型の候補が絞られ、動作解析の効率化を図ることができる。

表 2 に、代表的な型定義例を示す。1行で定義されている型は静的な型であり、複数行で定義されている型は動的な型である。多くの動的な型は、能の基本型である「カマエ」に帰着する。また、「カザシ」のように、他の型の一部に組み込まれるケースもある。

表 2 主要な型の定義例
Table 2 Examples of Form Definition

型名	首			右手			右足			左手			左足			注
	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z	
カマエ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
カザシ	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	
サシ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	カマエ
	0	0	0	0	1	1	0	1	0	0	0	0	0	0	1	
ヒラキ	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	カザシ
	0	0	0	0	1	1	0	0	0	0	1	1	0	-1	0	
	0	0	0	0	1	1	0	-1	0	0	1	1	0	0	0	
	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	
シトメ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	カマエ
	0	0	0	0	1	0	1	0	0	0	0	0	0	-1	0	
	0	0	0	0	1	1	0	0	0	0	0	0	0	-1	0	
	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
	0	0	0	0	1	1	0	0	0	0	0	0	0	-1	0	
	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	カマエ

(2) 動作解析手法

動画を 0.5 秒間隔でキャプチャーして得た画像に対して、基本姿勢モデルのノードに相当する観測点を追跡する。カメラ視線や演者の方向により座標の補正計算を行った上で、観測点の座標を基本姿勢モデルに受け渡し、3次元としての姿勢を推定する。人体としての成立し得ない解は排除する。残った解に対し、条件付き確

率場の計算モデルにより型の照合を行い、高確率の解を選定する。

型付情報が存在する場合は、音声表現指示や立ち位置の軌跡から舞の進行がわかり、型の候補を絞ることができ、効率的な解析が可能である。

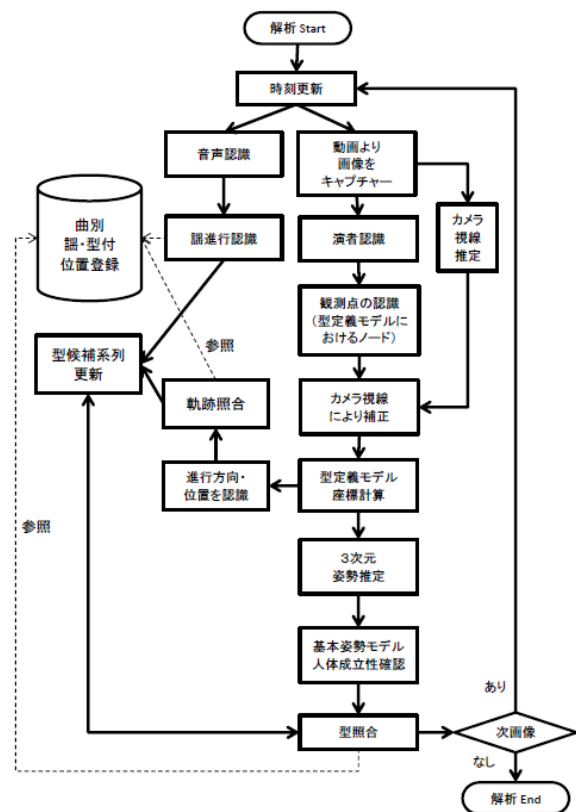


図4 動作解析手法
Figure 4 Method of Motion Analysis

図4に解析手順を示す。各プロセスは、起動条件が整えば自律的に起動するエージェントとして設計されている。各プロセスは、全プロセスが共有する型の候補系列情報を更新しながら、並行的に解析を進める。

以下に、主要なプロセスの処理概要を示す。

- 1) 演者の認識 一連の画像から、運動物体を検知し、演者を認識する。
- 2) 観測点の認識 型定義モデルの各ノードの運動を追跡する。能衣装をつけている場合でも、比較的追跡はやりやすい。
- 3) カメラ視線により補正 比較的認識のしやすい、能舞台の天井、床、柱の線を用いてカメラ視線を算出し、2)で追跡した各点の座標を能舞台の絶対座標系に換算する。本研究では、能舞台の中央に原点を置き、演者の立ち位置の軌道を算出している。

4) 型定義モデル座標計算 両踵の midpoint を人体モデルのローカル原点とし、右手方向を x 軸、正面向向きを y 軸とし各ノードの座標変換を行う。身長を 1.0 とした正規化した数値を基本姿勢モデルに受け渡す。型定義モデル独自の座標系定義 (表 2) に従い、各ノードの座標値を算出する。この数値は、型の照合に使われる。

5) 3次元姿勢推定 人体のプロポーシオン比を用いて、基本姿勢モデルの各中間ノードを推定する。中間ノードの存在し得る位置の集合は、3次元円の方程式として表現できる (図 5)。これを適度な粒度で離散化する。能は不要な動きを避けるため、ノードの経時変化における加速度ベクトルが大きいほど存在確率は下がるとみなし、一定の閾値で候補を絞った上で、候補系列を作成する。

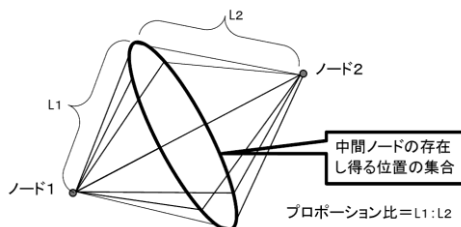


図5 中間ノードの存在し得る位置 (円)
Figure 5 Circle with existence probability of intermediate node

- 6) 基本姿勢モデル人体成立性確認 人体として成立し得ない角度を持つ解を除去する。
- 7) 型の照合 型定義モデルに立ち戻り、予め登録した型との照合を行う。型は、身体部位により若干時間的にずれて演じられることがある。このため、型の照合は先ず各ノードに分けて行い、検知時刻の差、並びに登録外運動の持続時間をコストとみなして信頼確率を減ずることにより、確率計算を行い、一定の閾値で候補を絞りこむ。

以下は必須ではないが、解析の効率化に役立つ。

- 8) 音声認識 動画が音声付きの場合は、音声認識を行い、詞章との照合を図る。能や舞囃子の場合は、音声には謡の他、囃子による音響が含まれる。しかし囃子の音響の内、能管は周波数が高いため、ローパスフィルタを用いて除

去することができる。残る囃子音響は打楽器と掛け声であるが、いずれも音響の減衰時間が短く、謡の音声とは識別が可能である。

謡の音声には、母音部を引き伸ばして発声するものや、一旦中断した後に母音のみを繰り返し発声するケースがよくある。謡音声は子音よりも母音の系列に特徴的なパターンが顕れると言える。そこで、フォルマント解析結果から各母音に対する存在確率を求め、各母音の確率場の系列を求め、謡との照合を図ることができる[4]。

音声認識ができなくとも謡曲部位との推定は可能であるが、認識できた場合は探索空間を絞ることができる。

- 9) 謡進行認識 8)で得た母音音列と予め登録しておいた謡の読み方とを比較し、一致した音列の長さにより信頼確率を計算する一方、差異についてはコストとみなして信頼確率を減ずることにより、確率的に謡の進行を認識し、その時点で舞われる型の可能性確率を算出し、型の候補系列を更新する。
- 10) 進行方向・立ち位置の認識 演者のボトムラインは最も容易に画像認識できる情報である。足も比較的検出しやすい。検出できている足の位置とボトムラインから、立ち位置を推定する。立ち位置は最終的には他のプロセスにより修正されることもある。立ち位置の軌跡から、進行方向を求めることができる。
- 11) 軌跡照合 立ち位置の軌跡と型付書の動線図解(図2)との照合から、舞の進行を認識することができる。動線上の各位置と舞の型は強い相関があり、その時点で取られている型を絞り込むことができる。この結果を型の候補系列に反映させる。

以上のプロセスを並列的に進めることにより、各時点の3次元姿勢と型が確率値と共に求まる。最終的には、最高確率の解を選定することで、舞の進行を時刻、3次元姿勢、型の名称により把握することができる。

同じ曲でも、進行速度や姿勢は演者による差異があり、個々の舞の特徴把握に役立つ情報が得られる。

6. 評価実験

本研究において最も難しいのは検証方法である。型の認識については、型付書との照合が可能であり、検証は可能である。しかし、3次元姿勢については正解が存在しないため、検証は

困難である。

そこで本論文では、型認識については、型付書と照合により定量的な評価が可能な実験を行い、3次元姿勢の認識については、定性的な評価を行った。

6.1 型認識実験

仕舞についての動画2種「田村」[5]と「羽衣キリ」[6]に対し、認識した型の推移を、型付書と照合する型認識評価実験を行った。

(1) 実験内容

予め「カザシ」「サシ」「ヒラキ」「シトメ」の型(表)を登録しておく。

型付けの登録を行わずに、動画から型認識を試みる。サンプリング間隔は0.5秒とする。

(2) 評価方法

得られた型認識結果を型付書[5]と比較する。

(3) 結果

型認識結果を図7~10に表示する。白い帯の部分がその型の認識された時間帯である。

2曲共に、認識した型の個数は完全に一致する。一方、その型の開始から終了までの出現のタイミングは、型付書と若干ずれる箇所が存在する。これは能楽師による舞表現の多様性の結果と考えられる。

型付書中の型が誤認なく認識できていることから、本手法による型認識は、少なくとも一部の基本的な型について、これまでに検証した映像記録に対しては機能している。しかし、今後も引き続き多様な映像記録に対し検証する必要がある。

6.2 3次元姿勢認識の評価

洋服で舞われた仕舞の動画[5]に対して得られた3次元認識結果の内、直接観測していない肘と膝の推定箇所を元映像に投影し、目視で位置に問題がないことを確認した。図6の白丸は、投影箇所を意味している。

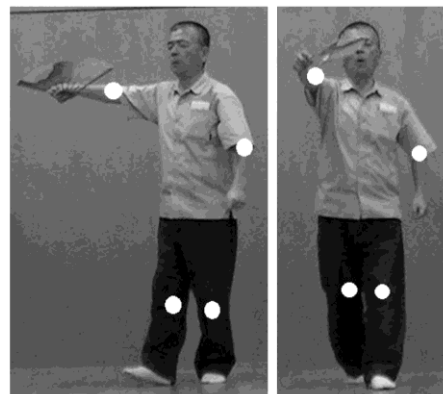


図6 肘・膝推定位置の確認

Figure 6 Checking Elbows and Knees' recognition points

語	読	型付書	カザシ	サシ	ヒラキ	シトメ	時刻(sec)
青橋の影絵にて	せいよおのかげみどりにて	ヒラキ			ヒラキ		32.0
							42.0
長閑なも書羽の道の白練の	のしらいとのな	ヒラキ			ヒラキ		57.0
		ヒラキ					63.0
地主権現の花の色も異なり	つしゆごんげんのはなのいろもことなり				ヒラキ		91.5
							37.5
		シトメ				シトメ	38.0
							44.5
在らん舞りはの舞舞舞にじものも木々の舞もさすや青橋の	た						
	よみずのみどりもさすや				ヒラキ		159.0
							166.5

図7 型認識結果1 (田村 1/2)
Figure 7 Experiment 1 (Tamura 1/2)

語	読	型付書	カザシ	サシ	ヒラキ	シトメ	時刻(sec)
.....	あおやぎの	ヒラキ		サシ			167.0
							174.5
げにも枯れたる木なりとも	げにも枯れたる木なりとも	ナン					177.0
		ヒラキ			ヒラキ		185.0
花櫻木のよそほひ	よそおい						191.0
いづくの春もおしなめて	いづくのはるもおしなめて	サシ		サシ			201.0
		カザシ					202.5
高麗けき影は青柳の	のどけ						208.5
天も花に舞へりや舞の舞の舞の舞の舞の	てんも	カザシ	カザシ				215.5
							217.0
の	の	シトメ				シトメ	247.0
	はる						252.0
型付書指示数		2	2	5	2		
認識数		2	2	5	2		

図8 型認識結果2 (田村 2/2)
Figure 8 Experiment 1 (Tamura 2/2)

題	読	型付書	カザシ	サシ	ヒラキ	シトメ	時刻(sec)
東遊の数々に	あずま						
	あそびのかずかずに						
其の名も月の	そのな				ヒラキ		36.0
	もつきの						
色人は	いろびとは	ヒラキ			ヒラキ		40.0
	はさんご	サシ					40.0
三五夜中の空に又	ちゆうのそらにまた			サシ			57.5
	また	ヒラキ				ヒラキ	58.0
	また						65.0
満眼真如の影となり	まんがんしん						
	よのかげとなり						
天の羽衣	あまのは						
	ごるも						

図9 型認識結果2 (羽衣キリ 1/2)
Figure 9 Experiment 2 (Hagoromo Kiri 1)

題	読	型付書	カザシ	サシ	ヒラキ	シトメ	時刻(sec)
清風に	うらかぜに						
	うきしまがくもの						
浮島が雲の	あしたかやまや	サシ		サシ			178.0
	あしたかやまや						185.0
愛鷹山や	ふじのたかね	カザシ		カザシ			191.0
	かすかになりて	カザシ					
富士の高嶺	あまつみそらの	カザシ		カザシ			
	かすみにまぎれて	ヒラキ					
天つ御空の	かすみにまぎれて				ヒラキ		215.5
	うせにけり	シトメ					シトメ
失せにけり	うせにけり						225.0
							232.0
型付書指示数		3	2	3	1		
認識数		1 ※	2	3	1		

※ 範囲は同じ。連続と判定。

図10 型認識結果2 (羽衣キリ 2/2)
Figure 10 Experiment 2 (Hagoromo Kiri 2)

7. 考察

一視点から撮影された能楽映像記録から、能の舞の特徴を活用することにより、3次元姿勢と型の認識手法を開発した。

本手法では、動画をキャプチャするサンプリング間隔の時間粒度で、舞の進行を把握することができる。能の基本原則である「序破急」の概念を、個々の能楽師がどのように舞に反映しているかといった芸術的意図も数値的に把握することができる。

入手可能な能楽映像記録の数は決して多くはない。しかし、能楽の多様な芸術表現を理解する上で、実に貴重な情報が含まれている重要な記録である。能楽の伝承のために活用できる情報もあれば、鑑賞者の理解を支援するために有用な情報もある。

現在はゲーム機や容易に入手可能な深度センサ等を使い、容易に3次元姿勢を認識することができる。教師データとの類似度の評価も容易い。今後はこれらの機能との結合も検討した上で、伝承に役立つツール開発も手がけていきたい。

鑑賞者のためには、情報をわかりやすく提供する手法の確立が必要である。本論文では、舞がどのような型から構成されており、個々の型はどの時点でどのように舞われているかについての情報を得たが、まだ鑑賞者に役立つ表現になっていない。今後の課題としたい。

8. あとがき

能楽映像記録から舞動作を解析し、個々の舞の特徴を把握することは能の経験を積むことに通じると考え、本研究を進めてきた。知識と経験が必要と言われる能の理解を助けられるよう、今後も研鑽に努め、能楽の普及に貢献できる研究を進めていきたい。

謝辞

喜多流能楽師出雲康雅氏に、能楽について貴重な御教示を賜りましたことを感謝申し上げます。

参考文献

- [1] 喜多流仕舞型附十曲，喜多流刊行会，平成22年。
- [2] 岩月正見,山中玲子ら: 能の型付資料に基づく所作単元の分析と舞の3Dアニメーション合成, 情報処理学会研究報告, 2009年。

- [3] 三上弾ら:メモリベースパーティクルフィルタ: 状態履歴に基づく事前分布予測を用いた頑健な対象追跡, 電子情報通信学会論文誌. D, 情報・システム J93-D(8), 1313-1328, 2010-08-01.
- [4] 内記 綾子, 青柳 龍也: 能楽音声記録の謡解析手法, 情報処理学会 人文科学とコンピュータシンポジウム論文集, 2012
- [5] 仕舞 田村, ビデオ映像, 出雲康雅, 2012年.
- [6] 栗谷明生仕舞集第壹巻より羽衣キリ, DVD.