Kobe University Repository : Kernel

PDF issue: 2025-07-06

文楽人形の演技動作における情緒性の解析に関する研究

服部，元史

(Degree)
博士（工学）
(Date of Degree)
2000-04-28
(Date of Publication)
2012-07-27
(Resource Type)
doctoral thesis
(Report Number)
乙2405
(JaLCDOI)
https://doi.org/10.11501/3173161
(URL)
https://hdl.handle.net/20.500.14094/D2002405

※ 当コンテンツは神戸大学の学術成果です。無断複製・不正使用等を禁じます。著作権法で認められている範囲内で、適切にご利用ください。

神戸大学博士論文

文楽人形の演技動作における情緒性の解析に関する研究

2000年4月

服部元史

神戸大学博士論文

文楽人形の演技動作における情緒性の解析に関する研究

2000年4月

服部元史

要旨

人間の形をした人工物に生き生きとした情緒豊かな動作をさせたいという願いは、自動人形のような物を考案し続けてきた昔から、人間にとって根強いものであった。アミューズメントやエンターテイメントにも価値をおく時代が到来したことから、人間型ロボットの研究においても、アミューズメント・ロボットのような物を作る必要性が提唱され、ハードウェア進歩に従って、多くの研究者や民間企業によって実際に製作されている。このような研究におけるロボット達が、生き生きとした動きの情緒豊かな表情をみせるためには、その動作を設計するにあたって様々な工夫がなされたはずであるが、そのような動作設計法は体系的には論じられていない。

そこで筆者らは、下記の理由によりヒューマノイド (人間型ロボットや文楽などの人形そして CG 内の人物など、人間の形をした人工物を、ここではヒューマノイドと呼ぶことにする) の情緒豊かな動作の設計方法の指針を得るために、日本の伝統芸能である文楽の人形の動作を解析した。

文楽人形の動きでは、長年に渡って修行を積んだ人形遣いによって、日本の伝統的な情緒表現が典型的に表現されている。本来 無生物である人形が、人形遣いによって操られた動きだけで、あたかも生きているかのように見える。このような特別な芸の力で操られた文楽人形の動きを解析すれば、無生物であるヒューマノイドに情緒豊かな生き生きとした動作を生成させる方法が得られるものと期待できる。

文楽人形の動作を記録して解析する研究においては、吉永孝雄や大西重孝といった演劇の研究者達による先駆的な仕事があり、具体的な演目で実演されている人形の動作の様々な型が考察されている。筆者らは工学者としての立場から、文楽人形の動作をモーション・キャプチャを用いて、客観的に正確な信号として記録して、時系列解析や力学解析などの工学的な手法に基づいて科学的な解析を試みた。つまり、上記のような文楽の具体的な演目において上演される文楽人形の複雑な演技動作を解析するための準備として、より基本的で単純な演技動作を解析し、ヒューマノイ

ドの動作に特有なデータ処理方法・主要な情報の抽出方法などを考案する。また、それらに基づいて、ヒューマノイドの動作設計に利用可能な 動作時系列の加工方法を幾つか提案した。

国立文楽劇場で活躍中の人形遣い　吉田玉女氏に、文楽人形の様々な動きを生成してもらい、時々刻々の人形の各部分の位置と姿勢をモーション・キャプチャで測定し、動作時系列を採集した。動作には、「挨拶する」「お茶を差し出す」などの特定の目的を達成するための機能の側面と、「好きである」「嫌いである」などの気持ちを表す情緒の側面があると考え、それぞれの機能の動作を幾つかの情緒で使い分けて、人形の動きを生成してもらった。このとき同一の機能と情緒の動作について、全く同一の動作（であると人形遣いが考える所のもの）を複数回生成してもらって測定した。動作時系列の測定については、第２章で詳しく述べた。

これらの測定された文楽人形の演技動作を、以下のように解析し、ヒューマノイドの動作の設計方法を提案した。

第３章では、文楽人形の動作時系列の確率的な揺らぎを補正して標準化を行うために、同じ条件のもとで複数回観測された時系列たちから、同一の事象に対応する時刻たちを求めて対応付ける方法を提案した。鍵となる考え方は、時系列たちのグラフが局所的に似ている時刻を対応させるという点であり、グラフの局所的な形を数量的に表す物として、時系列の各時刻における種々のスケールのウェーブレット係数に着目した。つまりウェーブレット係数が近い時刻どうしを対応する時刻と考えて、時系列の標準化を行う手法を提案した。比較するウェーブレット係数のスケールを適切に選ぶことによって、高周波雑音の影響を受けない時系列の標準化が可能である。

これら時系列の標準化手法を、文楽人形の動作時系列に対して適用し、標準化された動作時系列のグラフを検討し、それらに基づいてコンピュータ・グラフィックス上に生成された動作を観察することにより、上記の手法の問題点について考察し、標準化手法を時間区間・時間発展・人形の部位どうしの関連を考慮するように改良した。

また上記の標準化手法を用いて、時刻と状態の双方が確率的に揺らいでいる複数

の時系列から平均の時系列を求める方法を提案した。そのようにして平均化された時系列を参照信号として標準化を行い、良好な結果を得た。

提案された時系列標準化や平均時系列の計算手法は、文楽人形の演技動作の解析の今後の研究に有効であるだけでなく、異なる種類の複数の信号から対応する時刻を求める必要のあるセンサフュージョンや音声信号・ロボットへの教示信号などの応用分野の信号処理においても有用である。

第4章では、機能・情緒・確率的揺らぎの線形構造に基づいて、文楽人形の演技動作を解析し、ヒューマノイドの動作生成に使用する動作時系列の情緒を、強調させたり減衰させたりして加工するシステマティックな手法を提案した。様々な機能と様々な情緒の文楽人形の動作動作時系列から、機能の要素と情緒の要素を推定・抽出し、情緒の要素を強調させたり減衰させたりして、新しい動作を計算することができる。

第5章では、ヒューマノイドの動作を設計するにあたって、身体の各部位の動きを　ばらばらに設計するのではなく、それらの総体としての整合性のある動作を設計する技法を求めて、文楽人形の動作時系列を解析した。

測定された文楽人形の動作時系列を、（1）人形の慣性主軸の動きという主要な部分と（2）慣性主軸に対する相対動作という副次的な部分　の2つに分解する事を考えた。そして、動作の主要な部分を表している慣性主軸の動きが、動作の情緒の違いによって変化する様子を、時間伸縮と振幅値の増幅・減衰によって説明できる事を示した。さらに時間伸縮と振幅値の増幅・減衰を適切に設計する事で、情緒のない動作の軸の動きを加工して、情緒のある動作の軸の動きを生成する事を提案した。動作の主要部分を設計する一つの技法を示した事になる。時間伸縮を示す簡単なグラフの作成と振幅を増幅・減衰する数個のゲインの設定だけで、動作軸の動きの設計を行える事が、本手法の特色である。

上記に研究された動作解析および設計の手法の有効性を検証するために、文楽人形や動作軸などの動きを CG で視覚的に表示しながら、同時に動作時系列や動作軸のグラフなども確認できる マルチモーダル動作表示システムを開発したので、その概要を第6章にて紹介した。

以上のようにヒューマノイドの動作の設計方法を求めて、文楽人形の演技動作の解析方法を工夫して来た。人形の動作時系列（各時刻の各部分の位置と姿勢の値）という数値データを出発点とし、最初は数値に依存した動作の解析・設計方法を模索していたが、やがて第５章で提案された動作軸のように、細かい具体的な数値から余分なものをそぎ落として、動作の本質的な表現を求めて行く方向で進んで来たし、今後もこの方針に従って研究を進めるべきであると考える。

ヒューマノイドの動作を通じて、機能を実現したり情緒的な精神的な意図を伝達するにあたって、身体の各部位の各時刻の位置と姿勢という膨大な量の数値が本質なのではなく、それらからもっと簡略化された動作の本質だけを保存できる表現方法を、さらに深く究明すべきであると考える。その表現方法は、ヒューマノイドの身体の動きを表現するにあたって汎用的なものであり、人形や人間型ロボットだけでなく人間やアニメーションの人物など人間型の身体を有する物に共通なものであり、身体の形や大きさなどの個性に依存しないものが望ましいと考える。

このようなヒューマノイドの動作の記述・解析方法の研究は、人間型ロボットの動作設計にとって必要であるだけでなく、ビデオに記録された舞踊や文楽の身体の動きを３次元データに変換して、それらの動きの設計技法の解析や文化的な考察による分類などを客観的に行う比較舞踊学などの分野においても有用であると考える。このような動きの科学を客観的な考察に基づいて進めるにあたって、本論文で報告された内容は、基礎的な方向付けを提案するものである。

関連発表論文

(1) 服部元史、北川智広、田所諭、高森年、山田和人：Wavelet 係数を用いた動作時系列の標準化手法– 文楽人形の演技動作時系列の標準化への応用 – 日本機械学会論文集 C 編、投稿中

(2) 服部元史、中坊保則、田所諭、高森年、山田和人：文楽人形演技の動作軸の位相要素の解析– 人間型アミューズメント・ロボットの情緒豊かな動作生成を目指して – 日本機械学会論文集 C 編、掲載予定

(3) 服部元史、辻誠彦、田所諭、高森年、山田和人：人間型ロボットの動作生成において情緒を強調する一手法– 初期時刻を変化させた KM2O-Langevin 方程式を用いた文楽人形の演技動作の解析 – 日本機械学会論文集 C 編、掲載予定

(4) Motofumi Hattori, Masahiko Tsuji, Satoshi Tadokoro, Toshi Takamori and Kazuhito Yamada : An analysis and generation of Bunraku puppet's motion based on linear structure of functional factors, emotional factors and stochastic fluctuations for generation of humanoid robots' actions with fertile emotions, Journal of robotics and mechatronics, vol.11, no.5, pp.393-398, 1999

(5) 服部元史、辻誠彦、中坊保則、田所諭、高森年、山田和人：文楽人形の演技動作における確率的要素の解析– ホームロボットの情緒豊かな行動生成を目指して – 電気学会論文誌 C 電子・情報・システム部門誌, vol.117-c, no.5, pp.540-547, 1997

(6) Motofumi Hattori, Yasunori Nakabo, Satoshi Tadokoro, Toshi Takamori, and Kazuhito Yamada : An analysis of the Bunraku puppet's motions based on the phase correspondence of the puppet's motion axis – For generation of humanoid robots' actions with fertile emotions –, Proc. of 1999 IEEE international conference on systems, man, and cybernetics (SMC'99), vol. II, pp.1041-1046, Oct. 12-15, 1999, Tokyo, Japan

(7) Motofumi Hattori, Tomohiro Kitagawa, Masahiko Tsuji, Satoshi Tadokoro,

Toshi Takamori, and Kazuhito Yamada : Time normalization of action time series using their wavelet coefficients – For generation of humanoid robots' actions with fertile emotions –, Proc. of 1998 IEEE international conference on systems, man, and cybernetics (SMC'98) pp.4268-4273, October 11-14, 1998, Hyatt La Jolla, San Diego. California, USA

(8) Motofumi Hattori, Masahiko Tsuji, Satoshi Tadokoro, Toshi Takamori and Kazuhito Yamada : A motion analysis of the Bunraku puppet based on linear structure of functional factors, emotional factors and stochastic fluctuations – For generation of home robots' actions with fertile emotions –, Proc. of the 3rd international conference on advanced mechatronics (ICAM'98), vol.2, pp.873-878, Okayama, Japan, Aug. 1998

(9) Motofumi Hattori, Masahiko Tsuji, Satoshi Tadokoro, Toshi Takamori : A motion analysis of a Bunraku puppet based on KM2O-Langevin equation with varying initial time – For generation of home robots' actions with fertile emotions . Intelligent Autonomous Systems (Editors: Y.Kakazu, M.Wada, T.Sato), IOS Press, Ohmusha, pp.384-390, 1998 The 5 th international conference on intelligent autonomous systems (IAS-5)

目 次

1 緒論 **1**

2 文楽人形の演技動作と その測定 **5**

2.1 文楽人形の演技動作 5

2.2 演技動作の測定 7

3 ウェーブレット 係数を用いた動作時系列の標準化 **10**

3.1 緒言 10

3.2 繰り返された時系列の間で同一の事象を対応付ける手法 13

3.3 ウェーブレット係数を用いた時系列の標準化 15

3.4 高周波雑音の影響を取り除く対応付け 18

3.5 時系列データの標準化アルゴリズムの改良 27

3.5.1 単一の時刻における対応付け 29

3.5.2 時間区間における対応付け 30

3.5.3 時間発展を考慮した対応付け 32

3.6 人形の複数の部位を考慮した対応付け 35

3.6.1 参照信号の選出 35

3.6.2 各部位ばらばらの標準化 37

3.6.3 複数の部位を同時に考慮した標準化 40

3.7 平均動作時系列の作成方法 44

3.8 標準化された信号の評価 46

3.9 標準化された動作時系列に対する考察 48

3.10 結言 50

4 線形構造モデルに基づく機能・情緒・揺らぎの解析 52

4.1 緒言 52

4.2 動作を構成する 機能・情緒・揺らぎの要素 54

4.3 揺らぎの最小化による機能・情緒の要素の抽出 57

4.3.1 動作時系列の各要素の抽出 57

4.3.2 確率的要素の解析 58

4.3.3 結果と考察 63

4.4 KM_2O-ランジュバン方程式による機能・情緒の要素の抽出 65

4.4.1 外生入力を有する KM_2O-ランジュバン方程式による動作時系列の表現 65

4.4.2 動作時系列の各要素の抽出 67

4.4.3 数値シミュレーションによる同定方法の検証 69

4.4.4 動作時系列の各要素の抽出 76

4.4.5 結果と考察 78

4.5 機能的要素と情緒的要素による情緒的動作の生成 80

4.6 結言 85

5 文楽人形の動作軸の位相・振幅の解析 86

5.1 緒言 86

5.2 人形の動きを代表する動作軸 88

5.2.1 動作軸の提案 88

5.2.2 動作の構成に関する考察 92

5.2.3 情緒表現の要素に関する仮説 93

5.3 情緒表現における位相要素の抽出 96

5.3.1 ウェーブレット係数 97

5.3.2 スケールと approximation 99

5.3.3 位相調整手法の提案 103

5.3.4 位相要素の抽出 105

5.3.5 結果と考察 107

5.4 動作軸の振幅要素の抽出 108
5.4.1 位相抽出の結果の検討 108
5.4.2 情緒動作に対する機能動作のゲイン 111
5.4.3 結果と考察 117
5.5 動作軸の位相・振幅成分の調整による動作設計 118
5.6 結言 121

6 文楽人形のマルチモーダル動作表示システム 122
6.1 緒言 122
6.2 全体の構成 124
6.3 オブジェクトの構成 126
6.3.1 オブジェクトの階層構造 126
6.3.2 人形オブジェクト 127
6.3.3 グラフオブジェクト 130
6.3.4 動作軸オブジェクト 131
6.3.5 人形部位の動作軌跡オブジェクト 133
6.4 シーンの構成 134
6.4.1 オブジェクトの設定 134
6.4.2 カメラとライトの設定 141
6.5 マルチモーダル表示システム 143
6.5.1 一つの動作時系列データに基づいて表示する場合 143
6.5.2 二つの動作時系列データに基づいて表示する場合 145
6.6 動作表示の例 146
6.7 結言 153

7 結論 154

謝辞 156

参考文献 157

第 1 章　緒論

人間の形をした人工物に生き生きとした情緒豊かな動作をさせたいという願いは、自動人形のような物を考案し続けてきた昔から、人間にとって根強いものであった。アミューズメントやエンターテイメントにも価値をおく時代が到来したことから、人間型ロボットの研究においても、アミューズメント・ロボットのような物を作る必要性が、井口信洋らによって提唱されていた [1]。

ロボットのハードウェアの進歩に従って、池浦良淳と猪岡光による解説論文 [2] にて紹介されているような、井口信洋らによるダンシングロボット [3] や藤原春夫らによる阿波踊りロボット [4] などが、実際に製作されている。また遊園地や博物館などで使用するために、Kokoro などの民間企業が製作したアミューズメント・ロボットが見せてくれる動作の質の高さには目をみはるものがある [5]。

上記の研究におけるロボット達が、生き生きとした動きの情緒豊かな表情をみせるためには、その動作を設計するにあたって様々な工夫がなされたはずであるが、そのような動作設計法は体系的には論じられていない。そこで筆者らは、コンピュータ・グラフィックス (CG) で確認しながら、人間型の情緒豊かな動作を体系的に設計する方法を研究する。

CG 内の人物の動作を設計するにあたって、実際の人間の動作をモーション・キャプチャ(motion capture) を用いて計測し、得られた動作時系列を利用することが、広く行われている。このとき計測された動作時系列を、設計者が望むように加工したうえで CG 内の人物を動作させるので、その加工する技法が重要である。特に わらび座のデジタル・アート・ファクトリは、3 次元デジタル舞踊譜を開発し、様々な舞踊をモーション・キャプチャで計測して得た動作時系列を 3 次元デジタル舞踊譜で記述しており、CG を利用した舞踊の創作と 伝統舞踊の双方で注目を集めている [6]。舞踊や人形芝居などの伝統芸能をモーション・キャプチャで計測し、その動作

の本質を記述する技術は、CG を利用した芸術作品の創作において重要であるとともに、八村広三郎らによっても主張されているように、伝統芸能の記録・整理・保存という面からも、今後ますます研究される必要がある [7]。

そこで本研究では、日本の伝統芸能の文楽における人形の動作をモーション・キャプチャで計測し、得られた動作時系列を 様々な側面から解析した。この解析を通じて、人間型の動作から その本質を記述する方法を得ることによって、情緒豊かな動作の設計方法や伝統芸能の動作の記録方法に、改良を加えらるものと考える。

文楽人形の演技動作には、日本の伝統的な情緒表現が典型的に表現されている。日本の伝統的な人形芝居である文楽における人形の動きは、実に生き生きとしており、無生物であるとは思えない。これは、人形遣いの芸の力による動きだけで、無生物である人形があたかも生きているように見えるのである。この事は、無生物である人間型ロボットにも、その動かし方を適切に設計することによって、あたかも生きているような情緒豊かな表情を与えることができる可能性を示している。

文楽人形の動作を記録して解析する研究においては、吉永孝雄や大西重孝といった演劇の研究者達による先駆的な仕事がある [8] [9]。これらの研究においては、実演されている文楽人形の動作が、豊富な言葉と絵によって記録されており、ビデオ撮影やモーション・キャプチャによる記録が可能になった今日においても、文楽人形の動きを考察するにあたって貴重な資料となっている。とくに文楽という伝統芸能の具体的な演目を上演するために伝承されている人形の動きの様々な型を、記録・記述・解析するためには、上記の研究から学ぶことが多大であると考えるが、人間の目で見た人形の動作を言葉と手で描いた絵だけで記録する方法は、客観性という点で難点がある。筆者らは工学者としての立場から、文楽人形の動作をモーション・キャプチャを用いて、客観的に正確な信号として記録して、時系列解析や力学解析などの工学的な手法に基づいて科学的な解析を試みた。

本論文では、上記のような文楽の具体的な演目において上演される文楽人形の複雑な演技動作を解析するための準備として、より基本的で単純な演技動作を解析し、人間型の動作に特有なデータ処理方法・主要な情報の抽出方法などを考案する。また、それらに基づいて、人間型ロボットの動作設計に利用可能な 動作時系列の加工方法を幾つか提案する。

国立文楽劇場で活躍中の人形遣い　吉田玉女氏に、文楽人形の様々な動きを生成してもらい、時々刻々の人形の各部分の位置と姿勢をモーション・キャプチャで測定し、動作時系列を採集した。動作には、「挨拶する」「お茶を差し出す」などの特定の目的を達成するための機能の側面と、「好きである」「嫌いである」などの気持ちを表す情緒の側面があると考え、それぞれの機能の動作を幾つかの情緒で使い分けて、人形の動きを生成してもらった。このとき同一の機能と情緒の動作について、全く同一の動作（であると人形遣いが考える所のもの）を複数回生成してもらって測定した。動作時系列の測定については、第２章で詳しく述べる。

これらの測定された文楽人形の演技動作を、以下のように解析し、人間型の動作の設計方法を提案した。

まず第３章では、同一の動作を複数回繰り返して得た動作時系列を標準化する問題を扱った。全く同一の動作を複数回繰り返した場合、プロの人形遣といえども、それらの動作時系列は完全に同一のものが再現されるわけではなく、確率的に多少揺らいでいる。時系列のウェーブレット係数に基づいて、動作時系列の確率的な揺らぎを補正する標準化手法を提案した。この手法によって標準化された動作時系列を、以下の章で解析した。

第４章では、動作時系列を、機能・情緒・揺らぎの成分に分解する手法を提案した。動作時系列を、機能・情緒・揺らぎを表現する外生入力を有する KM2O-ランジュバン方程式で表現し、その外生入力を同定した結果に基づいて、機能・情緒・揺らぎの要素を計算した。求まった情緒の要素を増幅させたり減衰させたりして、新しい動作を設計する手法を提案した。

第５章では、動作の主要な部分を表現するものとして、人形の動作軸を提案し、その動きの解析に基づいて、人形遣いが情緒を使い分けて動作を生成することを、動作時系列の位相伸縮と振幅の増幅・減衰で説明する。この知見に基づいて、情緒のない動作時系列を加工して、情緒のある動作を設計する手法を提案した。

上記に研究された動作解析および設計の手法の有効性を検証するために、文楽人形や動作軸などの動きを CG で視覚的に表示しながら、同時に動作時系列や動作軸

のグラフなども確認できる マルチモーダル動作表示システムを開発したので、その概要を第6章にて紹介した。

本論文においては、日本の伝統芸能である文楽の人形の演技動作の単純で基本的な物を解析した段階までしか研究されていない。しかし本研究で得られた動作解析に関する幾つかの概念・方法は、以下に述べる意味において、文楽の具体的な演目における より複雑な演技動作を解析するにあたっても基本的な方向を示しており、文楽や舞踊などの身体の動きの伝統芸能を記録・記述・解析するにあたって有益な役割を果たすと考えられる。

複雑な演技動作は、それの要素となる演技動作 (phrase, フレーズ) が幾つか経時的に組み合わさって構成されており、それぞれのフレーズは さらに細かい複数の事象が経時的に合わさって構成されていると考えられる。複雑な演技動作を個々のフレーズに分解するためには、より細かい事象について考察する必要がある。第3章の動作時系列のウェーブレット係数や第5章の動作軸のウェーブレット係数は、演技動作の事象を定義する有力な概念になる。

またモーション・キャプチャで記録された動作時系列そのものは、身体の各部分の位置と姿勢を記録した時系列であるので、非常に高い次数の自由度を持ち、複雑すぎて そのままでは解析できない。また その動作を実演した人間型の身体 (人間や人形) の形状・大きさの個性に依存したデータであるので、別の形状・大きさを有する人間型の身体で再現できるデータではない。つまりモーション・キャプチャで記録された動作時系列のままでは、解析することが困難であり、自由度が低く しかも身体の個性に依存しないより本質的な記録へと変換したうえで解析する必要がある。このような動作の本質的な部分だけを記録する方法の一例を示した物が、第5章の動作軸である。

第 2 章　文楽人形の演技動作と その測定

本研究では、情緒表現を伴った動作の生成を目指して文楽人形の演技動作の解析を行う。本章では、その文楽人形の演技動作の測定方法と、測定する演技動作の種類、動作時系列の処理について述べる。

2.1　文楽人形の演技動作

今回の測定では測定実験用に製作した女の人形を使用した。文楽の女人形は、「かしら」と呼ばれる頭部、「肩板」と「胴輪」を布で繋げた胴体からなり、肩板から下は着物で被われている。

主遣い・左遣い・足遣いの 3 人の人形遣いによって文楽人形は操られ、主遣いが頭・肩 (胴体)・右腕・右手を、左遣いが左腕・左手を操る。女の人形には両足はなく、あたかも足があるかのように着物の形を足遣いが作っている。人形を動かすに当たって、主導権を握っているのは主遣いであり、左遣い・足遣いは主遣いに従って操作を行う。従って、文楽人形の各部のうち主遣いが操作する部分である頭、肩 (胴体)、右腕、右手の位置と姿勢を測定すれば、人形の動きは一意に決定されることになる。

文楽人形の演技動作には、日本の伝統的な情緒表現が典型的に表現されており、具体的な演目を上演するために人形の動きの様々な動きの型が伝承されている (吉田玉男氏 談)。そのような人形の動きの数多くの種類にわたる型を、記録・整理・解析することを筆者は目指しているが、そのような研究で必要となる有益な概念・手法を整備するために、本論文では、より基本的で単純な動作を対象とする。

文楽は人形によって演じられる演劇であり、人間によって行われる日常の動作が演じられる。しかし「ウソをまことしやかに演じるのが芸である」(竹本住大夫氏 談 [10]) といわれるように、日常動作を そのまま演じるのではなく、人形の役柄の持つ

感情や精神が観客に印象深く伝わるように、人形遣いによって芸術的に練り上げられた動きが演じられる。

ここでは最も基本的で単純な人形の動作から扱うので、文楽人形の、日常動作的な側面を機能と呼び、役柄の持つ精神的側面を情緒と呼んで、機能と情緒の２つの要素を動作が有していると考える。そこで、文楽協会の人形遣い 吉田玉女氏に依頼し、文楽人形に様々な「機能」(挨拶をする・お茶を渡す)、様々な「情緒」(感情を込めずに・好意を持って・敵意を持って) の動作を生成してもらい、測定を行った。また、同じ機能、同じ情緒の動作それぞれにつき、$K = 5$ 回ずつ繰り返して測定を行った。

より具体的に述べると次のようになる。「挨拶する」$(i = 1)$ という機能を「感情を込めない（情緒なし)」$(j = 1)$ という情緒で行った動作を測定して得られた時系列を $x^{11}(t)$ とし、「相手を好きという気持ち」$(j = 2)$ という情緒で行った動作を測定して得られた時系列を $x^{12}(t)$ とし、「相手を嫌いという気持ち」$(j = 3)$ という情緒で行った動作を測定して得られた時系列を $x^{13}(t)$ とする。

「お茶を差し出す」$(i = 2)$ という機能を「感情を込めない（情緒なし・普通)」$(j = 1)$ という情緒で行った動作を測定して得られた時系列を $x^{21}(t)$ とし、「相手を好きという気持ち」$(j = 2)$ という情緒で行った動作を測定して得られた時系列を $x^{22}(t)$ とし、「相手を嫌いという気持ち」$(j = 3)$ という情緒で行った動作を測定して得られた時系列を $x^{23}(t)$ とする。機能の添え字 i と情緒の添え字 j の意味を、以下の表にまとめた。

Table 2.1: Subscript i.

$i = 1$	$i = 2$
greeting	serving tea

Table 2.2: Subscript j.

$j = 1$	$j = 2$	$j = 3$
no emotion	friendly	hostile

また同一の機能 i　同一の情緒 j　に対して、全く同じ動作を $K = 5$　回、人形遣いに生成してもらった。それらのうち、k 回目 $(k = 1, 2, \cdots, K)$ に測定した動作

時系列を

$$x_k^{ij}(t) \quad (t = 0, 1, 2, \cdots, T) \tag{2.1}$$

で表す。

2.2 演技動作の測定

式 (2.1) のような文楽人形の動作時系列において、各時刻 t の $x_k^{ij}(t)$ は、人形の各部分の位置と姿勢から成り立っている。文楽人形の各部分の位置と姿勢の測定には、米国 POLHEMUS 社の FASTRAK というモーション・キャプチャを使用した。この FASTRAK は磁界を利用して 3 次元位置を測定する装置であり、コントロールユニット、トランスミッタ (磁場発生装置)、レシーバ (磁気センサ) からなる。各レシーバ毎に位置の 3 自由度 (x, y, z)、姿勢の 3 自由度 $(yaw, pitch, roll)$ が測定される。また、レシーバを 4 つ使用した場合、サンプリングタイムは各レシーバとも 0.2[s] である。

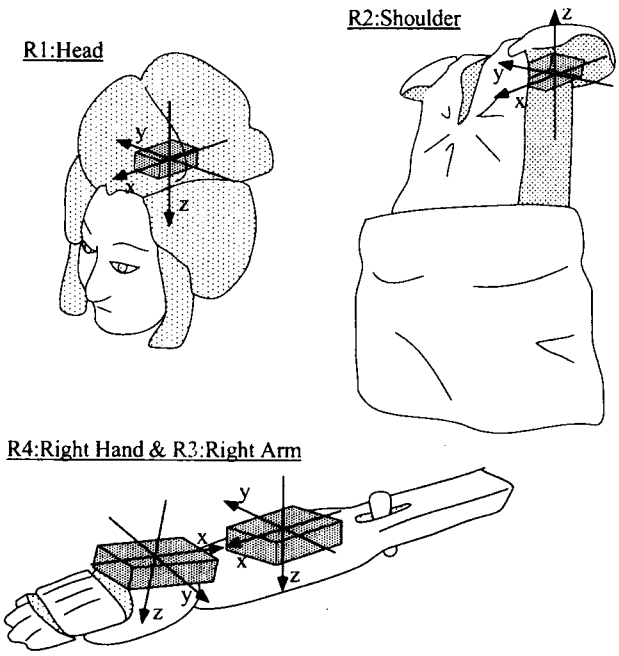


Figure 2.1: Setting of 4 Receivers.

Fig. 2.1に示すように、人形の R1 頭 $(p = 1)$・R2 肩 $(p = 2)$・R3 右腕 $(p = 3)$・

R4 右手 $(p=4)$ にセンサを取り付けて、それぞれの部分の x 座標 $(q=1)$・ y 座標 $(q=2)$ ・z 座標 $(q=3)$ という位置の値と、yaw$(q=4)$・pitch$(q=5)$・roll$(q=6)$ という姿勢の値を、時々刻々と測定した。

人形の動作時系列 $x(t)$ （但し　$t=0,1,2,\cdots,T$　は離散時刻）は、
頭 $(p=1)$ の時系列 $x_1(t)$ ・肩 $(p=2)$ の時系列 $x_2(t)$ ・右腕 $(p=3)$ の時系列 $x_3(t)$ ・右手 $(p=4)$ の時系列 $x_4(t)$ を成分とする縦ベクトル

$$x(t) = (x_p(t); p \downarrow 1, \cdots, 4) \tag{2.2}$$

$$= (x_1(t), \cdots, x_4(t))^T \tag{2.3}$$

であり、その成分の $x_p(t)$ は、x 座標 $(q=1)$ $x_{p1}(t)$・y 座標 $(q=2)$ $x_{p2}(t)$・z 座標 $(q=3)$ $x_{p3}(t)$・yaw$(q=4)$ $x_{p4}(t)$・pitch$(q=5)$ $x_{p5}(t)$・roll$(q=6)$ $x_{p6}(t)$ を成分とする縦ベクトル

$$x_p(t) = (x_{pq}(t); q \downarrow 1, \cdots, 6) \tag{2.4}$$

である。添字 p, q　の意味を整理すると以下の表のようになる。

Table 2.3: Subscript p.

$p=1$	$p=2$	$p=3$	$p=4$
R1:head	R2:shoulder	R3:arm	R4:hand

Table 2.4: Subscript q.

$q=1$	$q=2$	$q=3$	$q=4$	$q=5$	$q=6$
$x-coordinate$	$y-coordinate$	$z-coordinate$	yaw	$pitch$	$roll$

文楽人形の動作を表す時系列は、

$$x(t) \quad (t=0,1\cdots,T) \tag{2.5}$$

と、表される。ここで $t=0,1,...,T$ は離散時間であり、T は動作が終わる最終時刻を表している。

人形の動作時系列 $x_k^{ij}(t)$ は、頭 $(p=1)$ の時系列 $x_{k(1)}^{ij}(t)$ ・肩 $(p=2)$ の時系列 $x_{k(2)}^{ij}(t)$ ・右腕 $(p=3)$ の時系列 $x_{k(3)}^{ij}(t)$ ・右手 $(p=4)$ の時系列 $x_{k(4)}^{ij}(t)$ を成分とする縦ベクトル

$$
\begin{aligned}
x_k^{ij}(t) &= \left(x_{k(p)}^{ij}(t); p \downarrow 1, \cdots, 4\right) && (2.6)\\
&= \left(x_{k(1)}^{ij}(t), \cdots, x_{k(4)}^{ij}(t)\right)^T && (2.7)
\end{aligned}
$$

であり、その成分の $x_{k(p)}^{ij}(t)$ は、x 座標 $(q=1)$ $x_{k(p1)}^{ij}$ ・y 座標 $(q=2)$ $x_{k(p2)}^{ij}$ ・z 座標 $(q=3)$ $x_{k(p3)}^{ij}$ ・yaw$(q=4)$ $x_{k(p4)}^{ij}$ ・pitch$(q=5)$ $x_{k(p5)}^{ij}$ ・roll$(q=6)$ $x_{k(p6)}^{ij}$ を成分とする縦ベクトル

$$
\begin{aligned}
x_{k(p)}^{ij}(t) &= \left(x_{k(pq)}^{ij}(t); q \downarrow 1, 2, \cdots, 6\right) && (2.8)\\
&= \left(x_{k(p1)}^{ij}(t), \cdots, x_{k(p6)}^{ij}(t)\right)^T && (2.9)
\end{aligned}
$$

である。

このようにして得られた動作時系列を元に解析を行う際に、幾つか問題点が生じた。以下に問題点とその解決のための処理を示す。

1. 測定された動作時系列の長さ (時系列長) が一定ではなかった。そこで、測定動作時系列の中で最も長いものを基準とし、他は最終時刻の値を用いて長さを伸ばすことで、すべての動作時系列の長さを揃えた。

2. 姿勢をあらわす 3 自由度 $(yaw, pitch, roll)$ の測定範囲が $-180° \sim 180°$ であるため、$\pm 180°$ 付近で値が大きく変わるときに波形が不連続になった。そこで、測定範囲を $0° \sim 360°$ に変えることで、波形を連続にした。

3. 同一の機能・情緒の動作を $K=5$ 回ずつ測定したが、これらの動作開始時刻が一定ではなかった。そこで、$K=5$ 回のうち一つを基準とし、残りを基準に合わせた。

4. それぞれの動作開始時における開始位置が一定ではなかった。そこで、各機能の動作の中から一つを基準とし、トランスミッタ (座標系の原点に相当) を中心としてその他の動作の開始位置を基準の位置に揃えた。

第 3 章　ウェーブレット 係数を用いた動作時系列の標準化

3.1　緒言

第 2 章で述べたように、プロの人形遣い（文楽協会の吉田玉女氏）によって操られた文楽人形の種々の動作を測定した。これらは、人形の各部分の 3 次元の位置と 3 自由度の姿勢とを時々刻々と計測して得られた時系列である。様々の機能と情緒について、同一の機能と同一の情緒のもとでは同じ動作になるように人形遣いにお願いして、同一の動作を複数回繰り返してもらったが、もちろん完全に同一の時系列にはならず、時刻と振幅値の双方についてある程度確率的に揺らいでいる。つまり、時系列を構成する　それぞれの事象が実現する時刻が、繰り返される度に多少異なっている。そこで、本来同一であるべき複数の時系列の間で、共通の事象を探し出して、繰り返すたびに多少異なってしまっている実現時刻を補正する必要がある。このような時刻に関する確率的な揺らぎを補正する問題は、この文楽人形の動作時系列に限らず、音声認識や動作認識においても重要な問題であり、時系列の標準化と呼ばれている。[11]

時系列の標準化を行うにあたって、通常は DP マッチングのように、時系列の振幅値のみに注目して、繰り返された時系列の間で振幅値の最も近い事象を共通する状態とみなして対応付け、時刻を補正する手法が多く用いられている。[11][12][13]

しかし、そのような時系列の振幅値が近い時刻どうしを対応付ける手法は、雑音を含んだ時系列に用いると時刻の誤対応が生じ標準化が適切に行われないという欠点がある。

雑音を含んだ時系列の標準化を行うためには、振幅値が近い時刻を単に対応付けるのではなく、時系列をグラフで表示したときそのグラフの形が似ている時刻どうしを対応付けるべきである。つまり、横軸に時刻を縦軸に振幅値を表すグラフで複

数の時系列を表現したときに、そのグラフの局所的な形が似ている時刻どうしを対応付けるのが自然な考え方である。

時系列のグラフの局所的な形を数量的に表現する物として、その局所的な時刻における種々のスケールのウェーブレット係数が考えられる。本章では繰り返された時系列の各時刻のウェーブレット係数に基づいて標準化を行う手法を提案する。繰り返された時系列において、単に振幅値が近い時刻どうしを同一の事象とみなして対応付けるのではなく、ウェーブレット係数が近い時刻どうしを同一の事象とみなして対応付ける事にする。

まず基準となる信号 (参照信号) を 1 本決め、他の信号 (サンプル信号) の時刻の確率的揺らぎを補正して、参照信号に合わせるものとする。それぞれの信号の各時刻のウェーブレット係数を求め、参照信号のウェーブレット係数に最も近い時刻をサンプル信号から選び、時刻の揺らぎを補正して、新たなサンプル信号を生成するアルゴリズムの検討を行う。単一の時刻どうしのウェーブレット係数の比較だけで対応付けを行うことが不十分であることがわかったので、幅を持った時間区間上のウェーブレット係数を比較するようにアルゴリズムを修正する。また対応付けにあたって、時間の流れに逆行した対応付けがおこり得ることが認められたので、時間の発展を考慮したアルゴリズムに修正する。

このように 1 自由度の時系列に関して標準化のアルゴリズムが完成したので、それを用いて、実際の文楽人形の時系列データの各自由度ごとに標準化を行う。文楽人形の時系列データは頭、肩、右腕、右手の 4 部位それぞれが位置と姿勢の合計 6 自由度を持っている。まずそれぞれの部位、自由度共に個々で標準化を行ない、作成した時系列データの評価を行なう。その結果、24 自由度のうち、「腕と手」「頭と肩」などの複数の部位を同時に考慮して、これらの時系列データの標準化を行う必要性が認められたので、複数の部位を同時に考慮して対応付けを行うように、アルゴリズムを修正する。

また同一の現象から得られる複数の時系列から、時間的な揺らぎと振幅値の揺らぎの双方を考慮した平均の時系列を作成する方法を提案し、そのような平均時系列を参照信号として標準化を行った結果についても考察する。

第２章で測定された文楽人形の動作時系列データを、この章で確立された手法によって標準化した物を第４章 で用いた。

3.2 繰り返された時系列の間で同一の事象を対応付ける手法

時系列 $x(t)$ $(t = 0, 1, 2, \cdots, T)$ とは、各離散時刻 t における振幅値 x を並べた系列である。同じ条件のもとで同一の現象の時系列を複数回（K 回）観測して得られる K 本の sample path を $x_k(\cdot)$ $(k = 1, 2, \cdots, K)$ としたとき、これらは全く同一の信号にはならない。サンプル 1 の時系列 $x_1(\cdot)$ の時刻 t_1 の振幅値 $x_1(t_1)$ とサンプル 2 の時系列 $x_2(\cdot)$ の時刻 t_2 の振幅値 $x_2(t_2)$ が同じ事象を表現しているとしても、$t_1 \neq t_2$ となって時刻が異なり、$x_1(t_1) \neq x_2(t_2)$ となって振幅値が異なるのが普通である。

Fig.3.2 のように、同一の現象を 2 回観測して得られた時系列 $x_1(\cdot)$ (信号 1) と時系列 $x_2(\cdot)$ (信号 2) のグラフにおいて、a_1 と　a_2 は同一の事象を表現しているが、発生時刻も実現値も異なっているし、b_1 と　b_2 は同一の事象を表現しているが、発生時刻も実現値も異なっている。

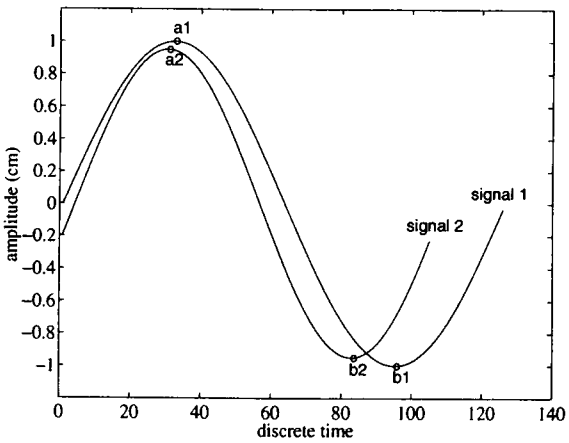


Figure 3.1: The basic idea of time normalization

例えば、「挨拶する」という動作を行う人間の頭の高さの時刻 t における値を $x(t)$ とする。同じ動作を K 回観測し、k 回目に観測された時系列を $x_k(\cdot)$ とする。「挨拶する」という動作は、「頭が一番上にある」事象に始まり、頭が段々と低くなる幾つもの事象を経て、「頭が一番低い」事象に至り、それから頭が段々と高くなる事象を経て、最後に「頭が一番上にある」事象で終わる。$x_1(t_1)$ と $x_2(t_2)$ とが共に「頭

が一番低い」事象を表しているとしても、一般には、$t_1 \neq t_2$ となり $x_1(t_1) \neq x_2(t_2)$ となるであろう。

このような本来同じである複数の時系列から同一の事象を表す時刻を対応させ、それぞれの時系列の時刻を修正して、複数の時系列が同じ時刻に同じ事象を表すように「時刻合わせ」を行うことが、「時系列の標準化」である。

3.3 ウェーブレット係数を用いた時系列の標準化

時系列の標準化を行う手法として、DP マッチングのように、状態の値そのそのものに注目する手法が従来多くの場面で用いられていた [11][12][13]。これは、時系列の振幅値が近い時刻どうしを対応付ける手法であるため、雑音を含んだ時系列に用いると時刻の誤対応が生じ標準化が適切に行われないという欠点がある。

雑音を含んだ時系列の標準化を行うためには、振幅値が近い時刻を単に対応付けるのではなく、時系列をグラフで表示したときそのグラフの形が似ている時刻どうしを対応付けるべきである。つまり、横軸に時刻を縦軸に状態を表すグラフで複数の時系列を表現したときに、そのグラフの局所的な形が似ている時刻どうしを対応付けるのが自然な考え方である。Fig.3.2 のように、同一の現象を 2 回観測して得られた時系列 $x_1(\cdot)$ (信号 1) と時系列 $x_2(\cdot)$ (信号 2) のグラフにおいて、a_1 と　a_2 を対応付けたり、b_1 と　b_2 を対応付けたりするのは、理に適っている。

時系列のグラフの局所的な形を数量的に表現する物として、その局所的な時刻における種々のスケールのウェーブレット係数が考えられる。時系列 $x_k(\cdot)$ に関する時刻 τ におけるスケール c の ウェーブレット 係数を $d_k(\tau, c)$ で表す。このとき大きいスケール（低い周波数に相当する）c_1 から小さいスケール（高い周波数に相当する）c_L に至る種々のスケールのウェーブレット係数を並べて得られるベクトル

$$\begin{aligned} d_k(\tau) &= (d_k(\tau, c_k); k \downarrow 1, 2, \cdots, L) \\ &= (d_k(\tau, c_1), \cdots, d_k(\tau, c_L))^{\mathrm{T}} \end{aligned} \tag{3.1}$$

が時系列 $x_k(\cdot)$ のグラフの時刻 τ の周辺の局所的な形を表していると考えられる。

従って、時系列 $x_1(\cdot)$ の時刻 t_1 における ウェーブレット係数たち $d_k(t_1)$ と 時系列 $x_2(\cdot)$ の時刻 t_2 における ウェーブレット係数たち $d_k(t_2)$ が近いベクトルであるとき、$x_1(t_1)$ と $x_2(t_2)$ が同じ事象を表していると考えられる。

時系列 $x_{ref}(t)$ を参照信号（基準）として、サンプル信号 $x_{smp}(t)$ を標準化する問題を考える。ここで　$t = 0, 1, 2, \cdots, T$ は離散時刻である。

スケール

$$c_1 > c_2 > \cdots > c_L \tag{3.2}$$

に対して、参照信号 x_{ref} の時刻 t におけるウェーブレット係数たちを成分とするベクトルを

$$d_{ref}(t) = (d_{ref}(t, c_l); l \downarrow 1, 2, \cdots, L)$$

とし、サンプル信号 x_{smp} の時刻 t におけるウェーブレット係数たちを成分とするベクトルを

$$d_{smp}(t) = (d_{smp}(t, c_l); l \downarrow 1, 2, \cdots, L)$$

とする。

時系列 $x_{ref}(t)$ を参照信号（基準）として、サンプル信号 $x_{smp}(t)$ を標準化して新しい信号 $x_{nor}(t)$ を作るものとする。上記の時系列の ウェーブレット係数の考察に従うと、サンプル信号 $x_{smp}(t)$ の各発生時刻を下記のように訂正して、標準化された信号 $x_{nor}(t)$ を作る事になる。

任意の離散時刻 t_{ref} に対して、参照信号の時刻 t_{ref} における ウェーブレット係数のベクトル $d_{ref}(t_{ref})$ と、サンプル信号の時刻 t_{smp} におけるウェーブレット係数のベクトル $d_{smp}(t_{smp})$ との差が、最も小さくなるような離散時刻 t_{smp} を選んで、その離散時刻 t_{smp} におけるサンプル信号の値 $x_{smp}(t_{smp})$ を、離散時刻 t_{ref} における新しい信号 $x_{nor}(t_{ref})$ の値として採用する。この操作を、全ての離散時刻 $t_{ref} = 0, 1, 2, \cdots, T$ に対して行う事によって、サンプル信号 $x_{smp}(t)$ を標準化した新しい時系列 $x_{nor}(t)$ $(t = 1, 2, \cdots, T)$ が得られる。

つまり、任意の離散時刻 t_{ref} に対して、

$$x_{nor}(t_{ref}) \stackrel{\text{def}}{=} x_{smp}(t_{smp}) \tag{3.3}$$

とする。但し、離散時刻 t_{smp} は、ベクトルの差の大きさ

$$|d_{ref}(t_{ref}) - d_{smp}(t_{smp})| \tag{3.4}$$

を最小化するように選ぶということになる。

ここでベクトル

$$d = (d_l; l \downarrow 1, 2, \cdots L) \tag{3.5}$$

に対して、その大きさは

$$|d| = \left(\sum_{l=1}^{L} {d_l}^2 \right)^{1/2} \tag{3.6}$$

で定義されている。

3.4 高周波雑音の影響を取り除く対応付け

ここで 式 (3.1) のスケール

$$c_1 > c_2 > \cdots > c_L \tag{3.7}$$

における最も小さいスケール c_L に対応する最も高い周波数を時系列中の高周波雑音の周波数より低く設定することによって、雑音の周波数よりも低い周波数における参照信号のグラフとサンプル信号のグラフとの局所的な形が似た時刻を対応付けることになり、雑音の影響を取り除いた標準化を行うことができる。

具体例を以下に示す。参照信号 $x_{ref}(t)$ を

$$x_{ref}(t) = sin\,(0.125 \times t \times T_{smp}) \tag{3.8}$$

とする。ただしサンプリング時間を

$$T_{smp} = 9.76 \times 10^{-4}[\mathrm{sec}] \tag{3.9}$$

とし、離散時刻

$$t = 0, 1, \cdots, T \tag{3.10}$$

の最終時刻を

$$T = 5118 \tag{3.11}$$

とする。雑音を含まないサンプル信号 (信号 1) を

$$x_1(t) = sin\,(0.1375 \times t \times T_{smp} - 0.4) \tag{3.12}$$

とし、雑音を含んだサンプル信号 (信号 2) を

$$x_2(t) = x_1(t) + Noise(t) \tag{3.13}$$

とする。ただし雑音 $Noise(t)$ を

$$Noise(t) = 0.4 \times sin\,(0.5028 \times t \times T_{smp})$$

で定義する。

参照信号の周波数は 0.0199[Hz] 、サンプル信号の周波数は共に 0.0219[Hz] 、雑音の周波数は 0.0801[Hz] となる。

ウェーブレット係数を求める時のアナライジング・ウェーブレットを $\Psi(t)$ で表す。$\Psi(t)$ は特定のものではなく、様々な局在する波の総称であり、信号を切り出す時の単位として使われるものである。そして、$\Psi(t)$ で切り出した信号の部分が、それぞれ時間軸上における時刻と、その時刻の局所的な周波数を表すウェーブレット係数である。

また、任意の周波数までの信号の成分を求める時に、周波数成分と対応するものとしてスケール (scale) というものが存在する。ここでスケールとは、$\Psi(t)$ の圧縮の度合を表すものである。次に例を示す。

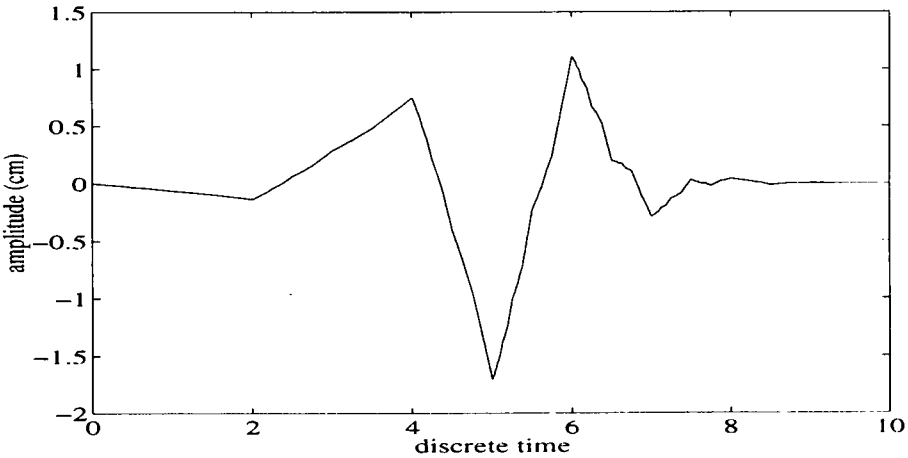


Figure 3.2: The analysing wavelet with scale 1.

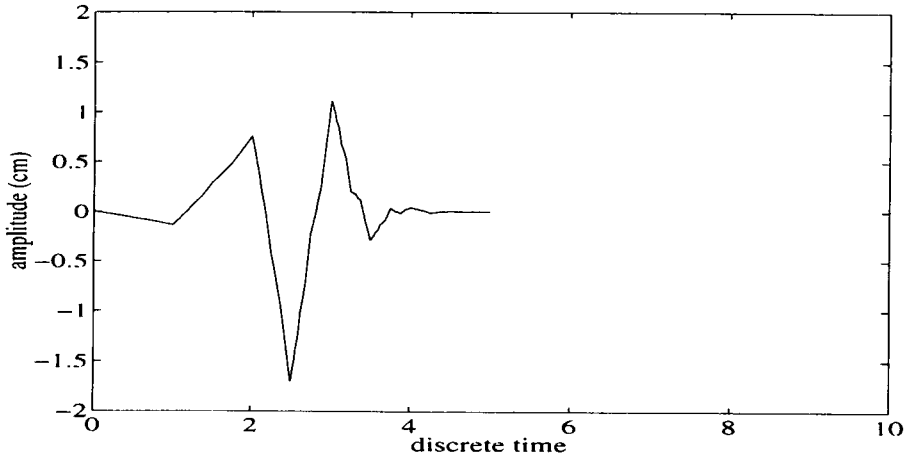


Figure 3.3: The analysing wavelet with scale $\frac{1}{2}$.

Fig. 3.2の $\Psi(t)$ のスケールを 1 とすると、Fig. 3.3のスケールは 1/2 であり、$\Psi(t)$ は 1/2 倍となっている。ここで、周波数の観点から見ると、Fig. 3.3は Fig. 3.2と比べて、周波数は 2 倍となっていることが分かる。スケールとそれに対応する周波数は逆数の関係にあり、スケールが小さいことは周波数が高いことに相当し、スケールが大きいことは周波数が低いことに相当する。

このように、ウェーブレット係数ではスケールが用いられ、これは周波数に大きく関係するため、次に $\Psi(t)$ と信号、そして雑音について、成分が最大となる周波数を調べる。

アナライジング ウェーブレット $\Psi(t)$、参照信号、サンプル信号 1、サンプル信号 2 の雑音部をそれぞれフーリエ変換し、その成分が最大となる周波数を調べた。アナライジング ウェーブレット $\Psi(t)$ のフーリエ変換のグラフを Fig. 3.4に示す。

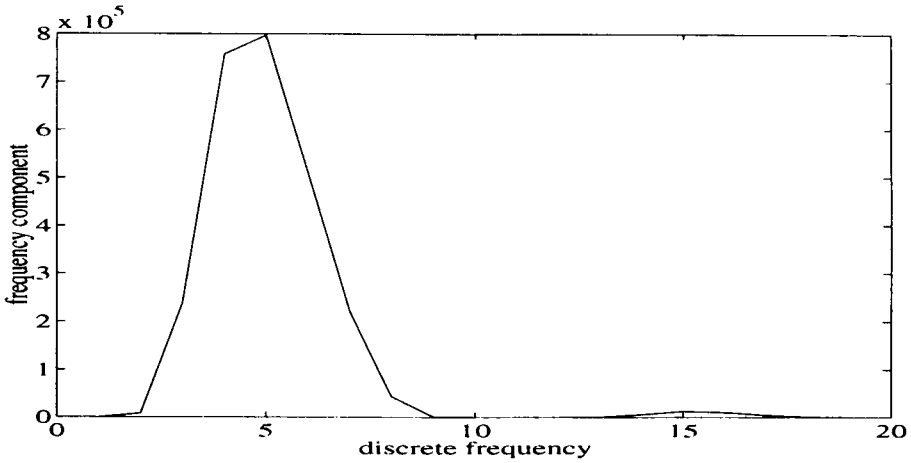


Figure 3.4: Frequency-domain representation of the analysing wavelet.

サンプリング時間が $T_{smp} = 9.76 \times 10^{-4}[s]$ であることを考慮すると、$\Psi(t)$ の成分が最大となる周波数は $F_{\Psi} = 0.8005[Hz]$ となる。

アナライジング・ウェーブレットの周波数、参照信号の周波数、サンプル信号の周波数、雑音の周波数を、Table 3.1に示す。

Table 3.1: The main frequencies of the signals.

signal	frequceny [Hz]
analysing wavelet	0.8005
reference signal	0.0199
sample signal 1	0.0219
noise of sample 2	0.0801

サンプル信号 1 は、周波数が参照信号より少し高く、時間が進んでいるものである。またサンプル信号 2 は、これに参照信号の約 4 倍の周波数の雑音を加えたものである。また、ウェーブレット係数を求める時に用いるアナライジング・ウェーブレット $\Psi(t)$ の成分が最大となる周波数は $F_{\Psi} = 0.8005[Hz]$ であった。$\Psi(t)$ の任意のスケールに含まれる成分が最大となる周波数を F_{scale} とする。例えばつまり、$\Psi(t)$

のスケール 20 に含まれる成分が最大となる周波数を F_{20} とすると、

$$F_{scale} = \frac{F_{\Psi}}{scale} \tag{3.14}$$

が成立する。

まず、サンプル信号 1 の標準化を行なう。結果を Fig. 3.5 と Fig. 3.6に示す。ただし、横軸は離散時間、縦軸は振幅値である。

ウェーブレット係数による標準化の結果を、Fig. 3.5 に示す。DP マッチングのように、ウェーブレット係数ではなく振幅値の最も近い時刻どうしを対応付けて標準化を行った結果を、Fig. 3.6 に示す。どちらの図も共に、上から順に参照信号・サンプル信号・標準化された信号が示されている。これらを見て分かるように、周波数成分が近く、時間的にずれている信号の標準化は、多少の違いはあるがどちらの場合もうまく標準化されていることが分かる。

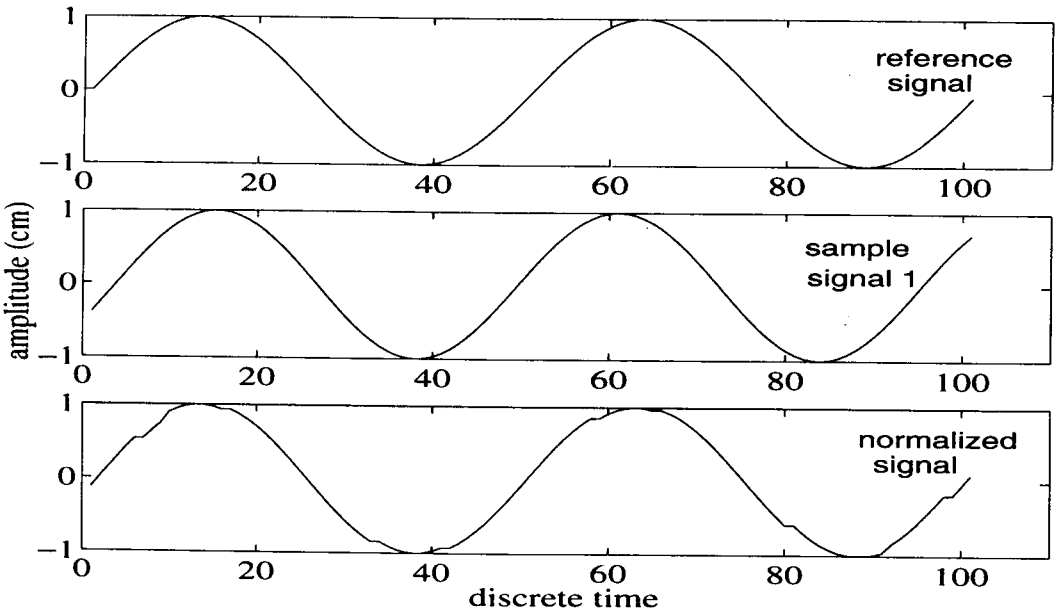


Figure 3.5: Normalization of a signal with NO noise by wavelet coefficients.

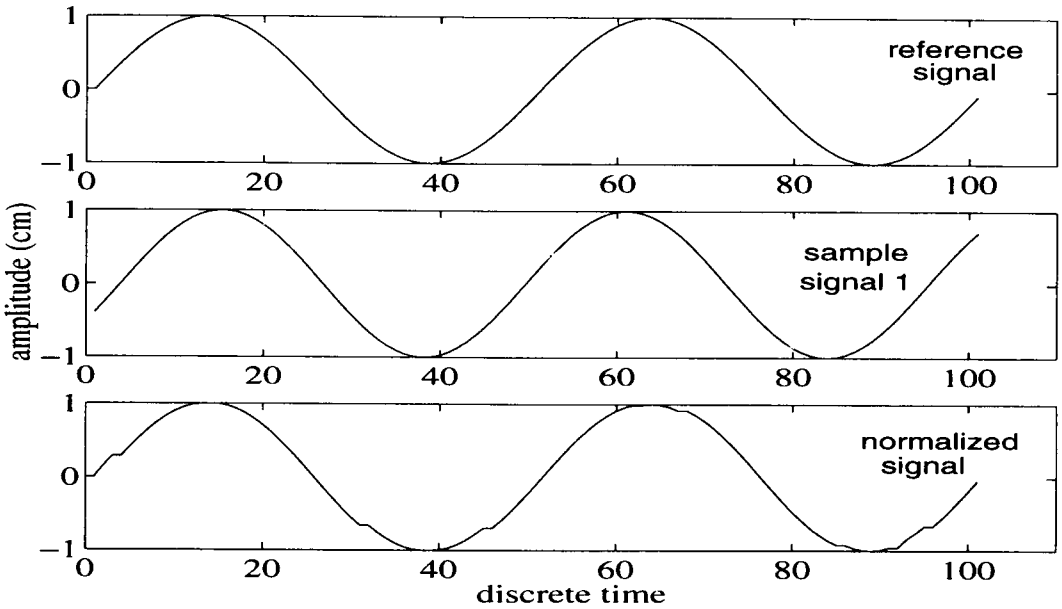


Figure 3.6: Normalization of sample signal with NO noise by amplitudes.

次に、高周波数成分の雑音を加えたサンプル信号 2 の標準化を行なう。その時の結果を Fig. 3.7と Fig. 3.8に示す。Fig. 3.7はウェーブレット係数による信号の標準化（考慮された周波数は 0.0178[Hz] から 0.0229[Hz] の範囲）であり、Fig. 3.8は振幅値による標準化である。

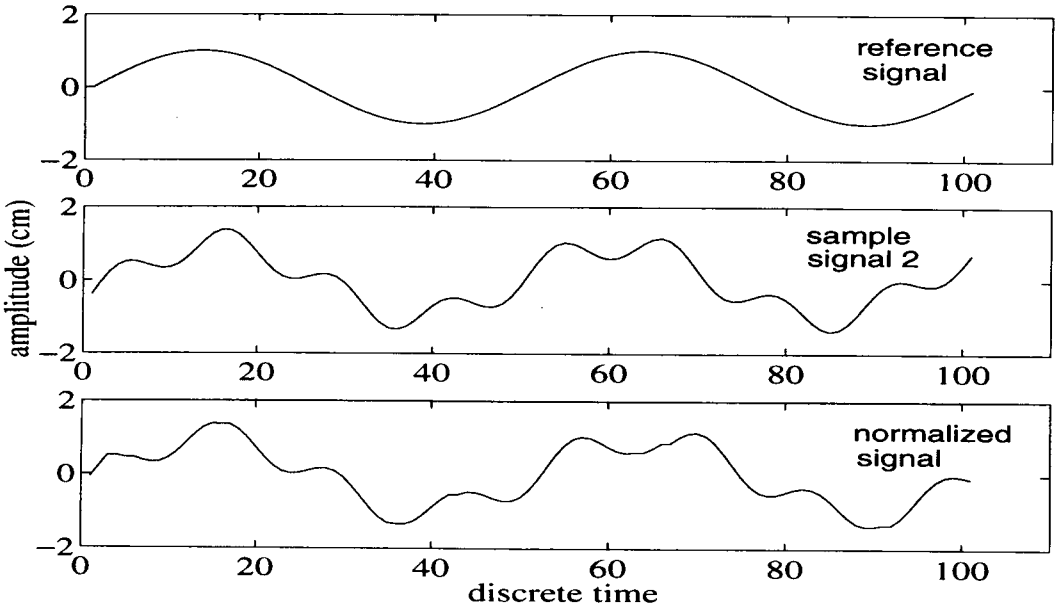


Figure 3.7: Normalizaton of a signal with NOISE by wavelet coefficients.

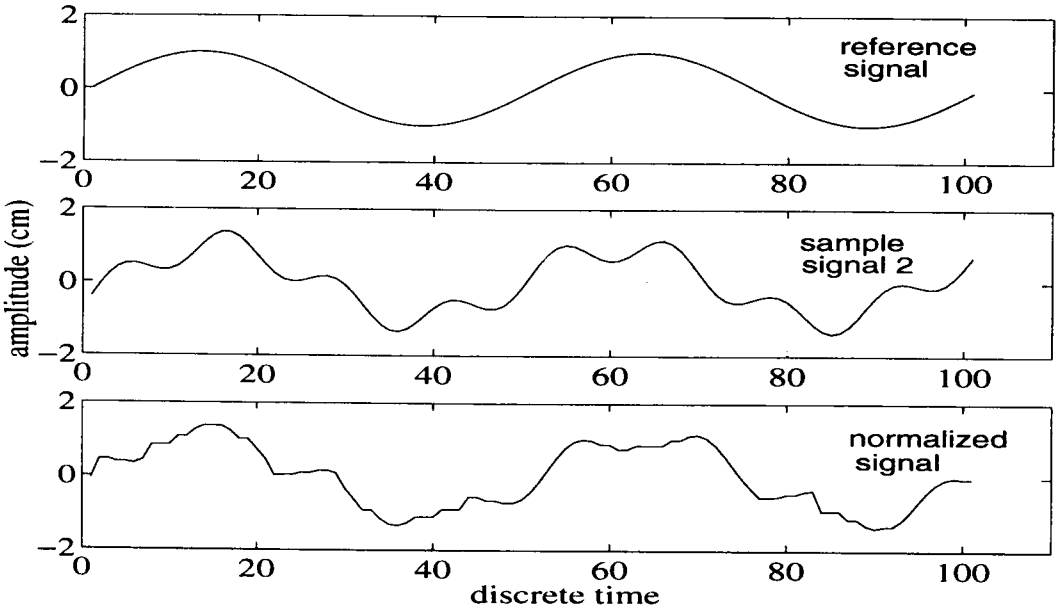


Figure 3.8: Normalizaton of a signal with NOISE by amplitudes.

標準化された結果を見ると、振幅値による標準化の方は、波形がずれてしまって、以前のように適切な標準化が行なわれていない。これは、標準化された信号が振幅値で評価されているため、様々な時刻から雑音部を対応づけられた結果であると考えられる。振幅値で標準化を行なうと、信号本来の成分が対応する時刻を見つけら

れず、動作時間を合わせるための標準化ができない危険性がある。

これに対し、ウェーブレット係数による標準化は、サンプル信号が持つ雑音まで含めた成分を残したまま、動作時間だけが適切に合っているとみなせる。ウェーブレット係数では、ある時刻での任意の周波数成分を比較できるため、サンプル信号2のように雑音が含まれている場合でも、雑音が持つ高周波数成分を無視することにより、信号の低周波数成分のみを適切に対応づけることが可能となる。つまり、サンプル信号2には、雑音である0.0801[Hz]の周波数成分が含まれているが、ウェーブレット係数を用いて標準化を行なう時に、その雑音の周波数以下の成分だけに注目して、適切な標準化を行った。

また、ウェーブレット係数を用いた場合でも、考慮する最高周波数を0.160[Hz]として、雑音の周波数0.0801[Hz]を含む範囲のウェーブレット係数を用いて標準化を行うと、結果はFig. 3.9のようになり(上から順に 参照信号・サンプル信号・標準化された信号が示されている)、標準化は良好に実現されない。

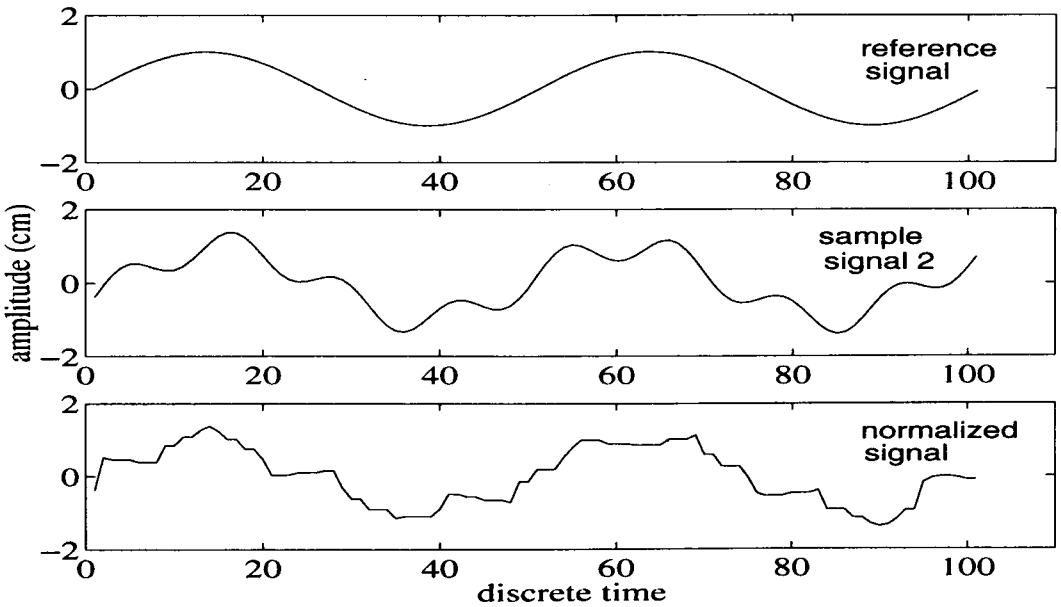


Figure 3.9: A bad example of normalizaton of a signal with NOISE by wavelet coefficients.

このようにウェーブレット係数を用いることで、様々な周波数の信号成分を評価

でき、信号や雑音の周波数特性を考慮した標準化を行なうことができる。以上の結果から様々な時系列データに対応できるものとして、DP マッチングのような振幅値による標準化ではなく、ウェーブレット係数に基づいた時系列データの標準化を、本研究では採用する。

3.5 時系列データの標準化アルゴリズムの改良

第2章で述べたように、人形の動作時系列 $x(t)$
(但し　$t = 0, 1, 2, \cdots, T$　は離散時刻) は、頭 $(p = 1)$ の時系列 $x_1(t)$ ・肩 $(p = 2)$ の時系列 $x_2(t)$ ・右腕 $(p = 3)$ の時系列 $x_3(t)$ ・右手 $(p = 4)$ の時系列 $x_4(t)$ を成分とする縦ベクトル

$$x(t) = (x_p(t); p \downarrow 1, \cdots, 4) \tag{3.15}$$

$$= (x_1(t), \cdots, x_4(t))^T \tag{3.16}$$

であり、その成分の $x_p(t)$ は、x 座標 $(q = 1)$ $x_{p1}(t)$・y 座標 $(q = 2)$ $x_{p2}(t)$・z 座標 $(q = 3)$ $x_{p3}(t)$・yaw$(q = 4)$ $x_{p4}(t)$・pitch$(q = 5)$ $x_{p5}(t)$・roll$(q = 6)$ $x_{p6}(t)$ を成分とする縦ベクトル

$$x_p(t) = (x_{pq}(t); q \downarrow 1, \cdots, 6) \tag{3.17}$$

である。

人形の一つの部位 p と一つの座標 q に対して実数値の時系列

$$x_{pq}(t) \quad (t = 0, 1, 2, \cdots, T) \tag{3.18}$$

$$(T = 105) \tag{3.19}$$

に対して、同一の動作を複数回測定している。その一例を Fig.3.10 に示す。これは同一の動作を5回測定したデータの、頭 $(p = 1)$ の pitch $(q = 5)$ の時系列 $x_{pq}(t)$ を5回に渡ってグラフに表示したものである。1回ごとに異なる時系列が得られているが、5回の時系列のグラフの外形が似ている様子が示されている。

このように同一の動作を複数回に渡って観測して得られた、よく似た複数の時系列を、ウェーブレット係数を用いて標準化を行って、生じてきた問題点と　その改良方法を以下で述べる。例えば、参照信号を　Fig. 3.11とし、サンプル信号を　Fig. 3.12　とした時の標準化の結果に基づいて説明する。

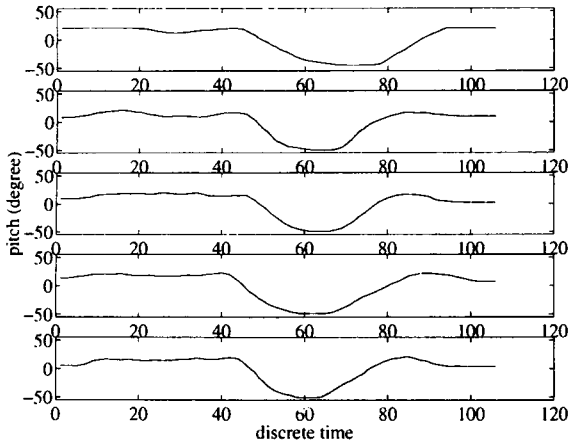


Figure 3.10: 5 signals of the same actions.

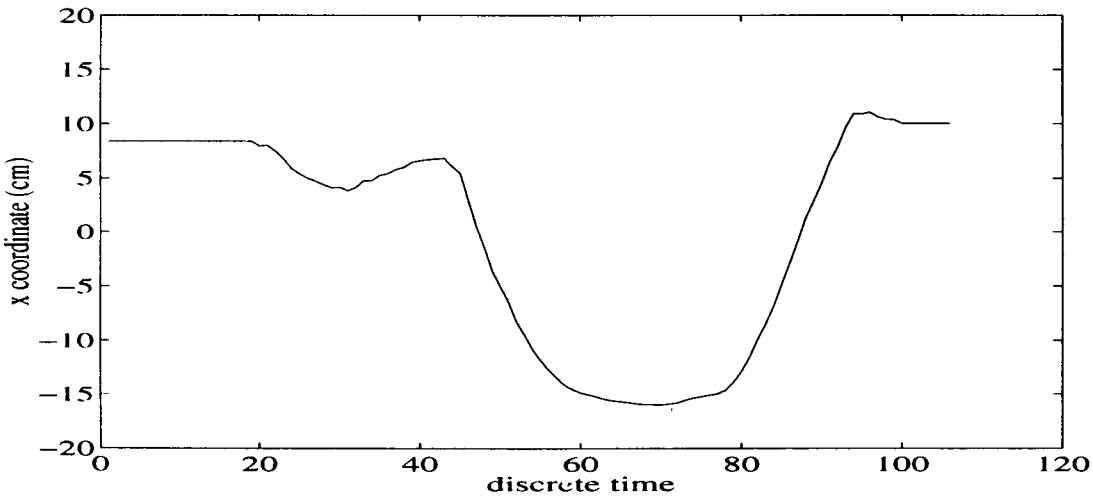


Figure 3.11: Reference signal.

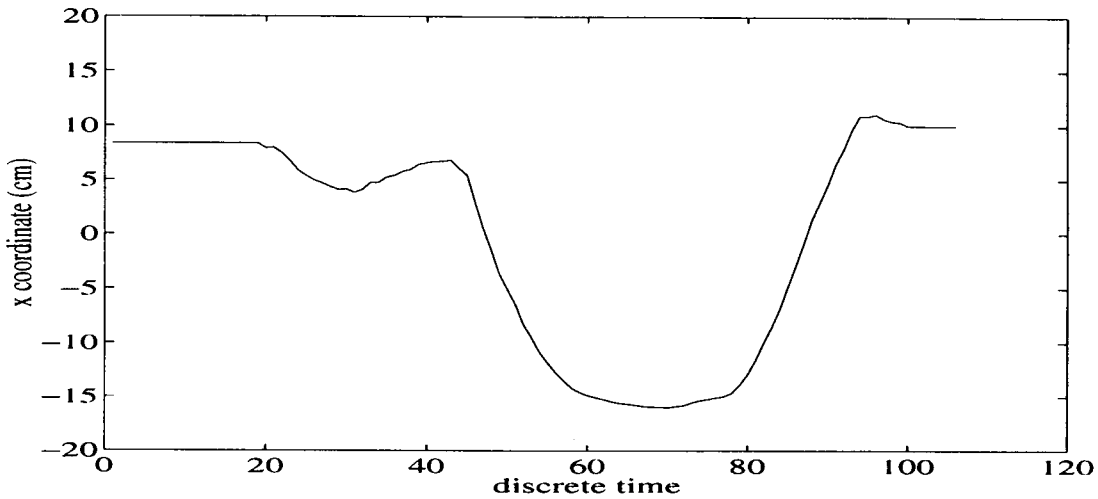


Figure 3.12: Sample signal.

3.5.1 単一の時刻における対応付け

参照信号の時刻 $t_{ref} = 0, 1, \cdots, T$ が対応しているサンプル信号の時刻を順に求め、それを新たな時刻としてサンプル信号を再構成する。ここで、参照信号とサンプル信号の時間的揺らぎはある範囲内に限られると考えられるため、参照信号の時刻 t_{ref} に対応する時刻を見つけるうえサンプル信号の時刻を

$$t_{ref} - T_{cmp}, \cdots, t_{ref} - 1, t_{ref}, t_{ref} + 1, \cdots, t_{ref} + T_{cmp} \tag{3.20}$$

の範囲から探索することにする。$T_{cmp} = 20$ としたときの結果を Fig. 3.13に示す。上から順に参照信号・サンプル信号・標準化された信号が示されている。

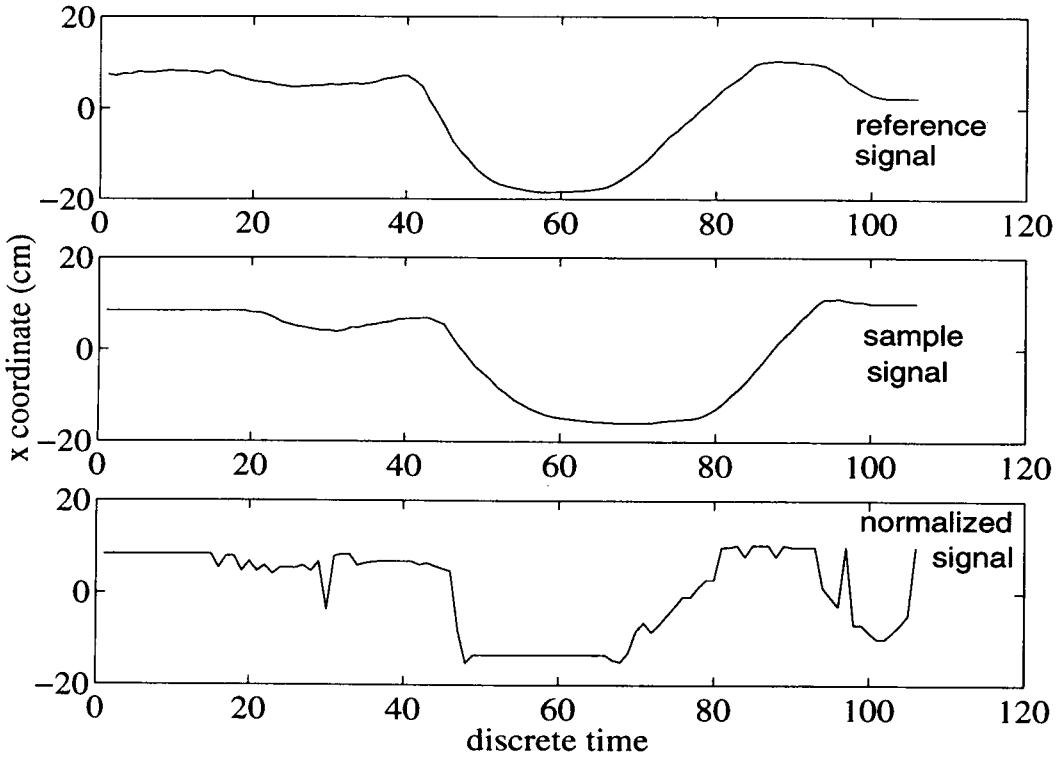


Figure 3.13: Time normalization at Each Time.

Fig. 3.13では、標準化された信号の谷 (振幅値が低い部分) の幅がサンプル信号の時よりも狭くなり、この点では標準化が適切に行なわれていることが分かる。しかし、全体的に見ると信号が様々なばらつきをとっており、元のサンプル信号と比較しても明らかに信号として不適切である。

3.5.2 時間区間における対応付け

前節では参照信号とサンプル信号を対応させるうえで、単一の時刻どうしのウェーブレット係数を比較した。しかし信号として特徴を持つある時刻 t におけるグラフの形をを考えるときに、その時刻の前後の時刻 $t-\tau, t-\tau+1, \cdots, t, \cdots, t+\tau-1, t+\tau$ も関係があるものと考えられる。

そこで標準化された信号を作る時に式 (3.3) で用いられる離散時刻 t_{smp} の選び方を、単一時刻の ウェーブレット係数の差である式 (3.4) を最小化するように選ぶの

ではなく、

$$\sum_{s=-\tau}^{\tau} |d_{ref}(t_{ref}+s) - d_{smp}(t_{smp}+s)|^2 \tag{3.21}$$

を最小化するようにした。つまり

$$-\tau, -\tau+1, \cdots, 0, \cdots, \tau-1, \tau \tag{3.22}$$

という時間区間の上で、参照信号 $x_{ref}(t_{ref})$ の ウェーブレット係数

$$\begin{aligned} &d_{ref}(t_{ref}+s) \\ &s = -\tau, -\tau+1, \cdots, 0, \cdots, \tau-1, \tau \end{aligned} \tag{3.23}$$

とサンプル信号 $x_{smp}(t_{smp})$ の ウェーブレット係数

$$\begin{aligned} &d_{smp}(t_{smp}+s) \\ &s = -\tau, -\tau+1, \cdots, 0, \cdots, \tau-1, \tau \end{aligned} \tag{3.24}$$

との差が最小になるように 離散時刻 t_{smp} を選んで、式 (3.3) に従って、標準化された信号 $x_{nor}(t)$ を作成するようにした。

時間区間で比較して標準化を行なった結果を Fig. 3.14に示す。ただし、ここでは $\tau = 4$ で式 (3.22) のような比較する時間区間を $t_{ref}-4, \cdots, t_{ref}, \cdots, t_{ref}+4$ とし、式 (3.20) におけるサンプル信号の比較範囲は $T_{cmp} = 20$ とした。

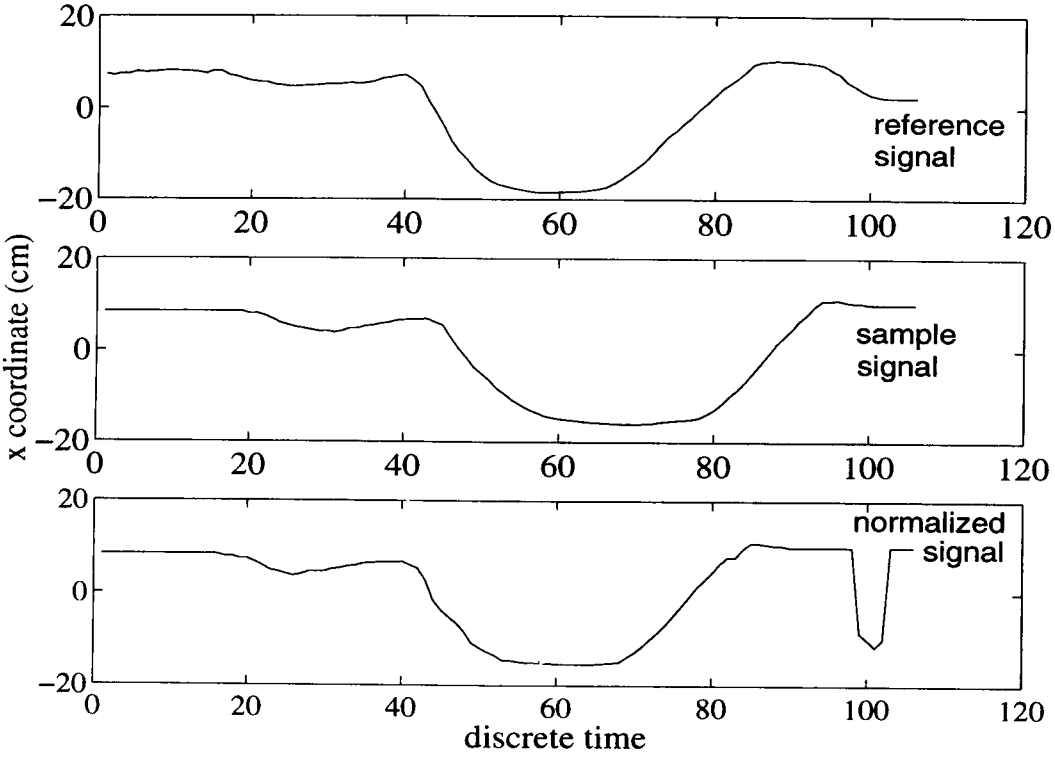


Figure 3.14: Time normalization on Each Time Interval.

Fig. 3.14を見ると、信号の様々なばらつきがなくなり、Fig. 3.13と比べて適切な標準化が行なえていることが分かる。しかし、標準化された信号では図からも明らかなように、時刻 100 付近でサンプル信号の時刻からはあり得ない波形が表れている。この付近の対応時刻を調べてみると、参照信号の時刻 100 付近と対応している時刻がサンプル信号の時刻 80 付近であることが判明した。このため標準化された信号が、突然不適切な振幅をとり得るという結果になった。

3.5.3 時間発展を考慮した対応付け

時間区間上の ウェーブレット係数を比較して標準化を行うことによって、Fig.3.14 の結果は Fig.3.13 に比べてかなり改善されたが、動作の終了時あたりの標準化がまだ誤っている。そこで誤っていると思われる時刻において、参照信号の時刻 t_{ref} に対して、式 (3.21) を最小化する t_{smp} が どのように選ばれているのかを調べた。それを Table 3.2 に示す。

Table 3.2: Time correspondence between the referecce signal and the sample signal. (bad case)

reference signal	$\cdots$	96	97	98	99	100	101	102	103	104	$\cdots$
sample signal	$\cdots$	102	102	102	83	82	81	82	106	106	$\cdots$

Table 3.2 より、$t_{ref} = 98$ に対して $t_{smp} = 102$ が選ばれているのに対して、$t_{ref} = 99$ に対して $t_{smp} = 83$ が選ばれている。つまり t_{ref} が 98 から 99 へと増加しているのに対して、それに対応する t_{smp} は 102 から 83 へと減少している事になり、時間発展が逆行した時刻の対応付けによって標準化が行われていた事が分かった。

このような時間発展の逆行した時刻の対応付けが起こらないようにするために、参照信号の時刻 t'_{ref} にサンプル信号の時刻 t'_{smp} が対応付けられたならば、t'_{ref} より大きな参照信号の時刻 t_{ref} に対しては、t'_{smp} 以上のサンプル信号の時刻のうちで、式 (3.21) を最小化する t_{smp} を対応付けるように、時間発展の制約を設けて標準化を行うようにした。

参照信号の時刻に対して

$$t'_{ref} < t_{ref} \tag{3.25}$$

という時間発展があるならば、それらに対応するサンプル信号の時刻も

$$t'_{smp} \leq t_{smp} \tag{3.26}$$

という時間発展が反映されるようにして、標準化を行うように改良した。その結果が Fig.3.15 であるが、良好な標準化が実現している。ただし、取り出した探索範囲 T_{cmp} と比較範囲 τ は前回と同じで、$T_{cmp} = 20$, $\tau = 4$ である。

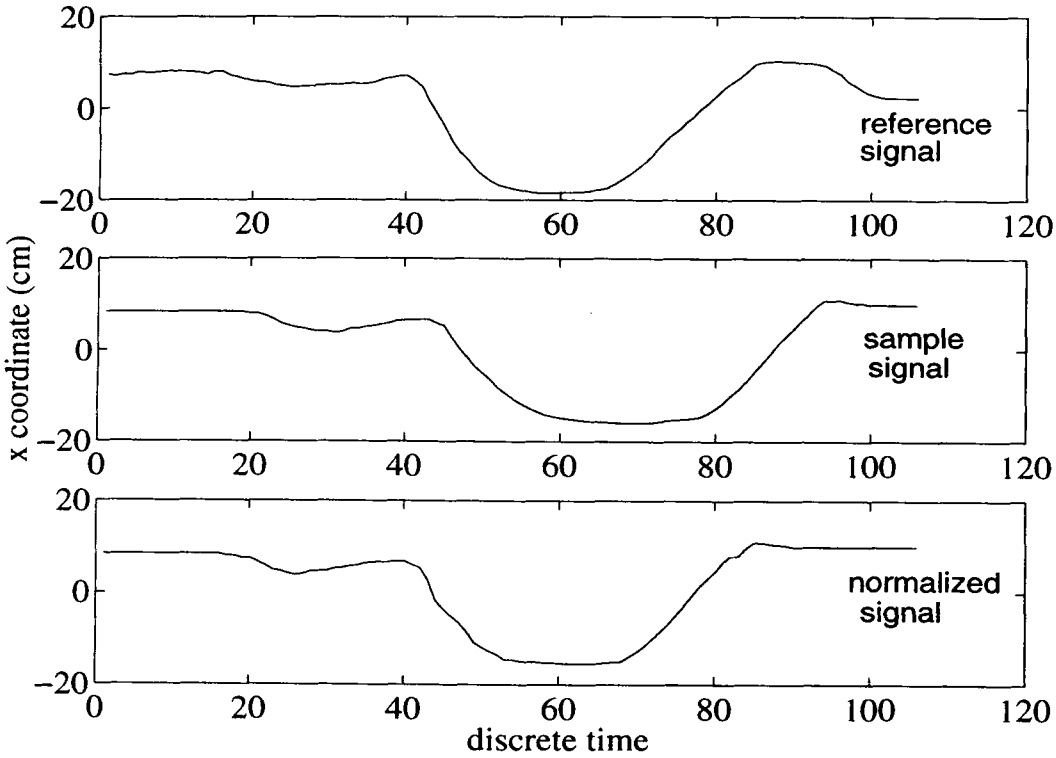


Figure 3.15: Time normalization considering time evolution.

この標準化における t_{ref} と t_{smp} の対応関係を Table 3.3 に示す。時間発展が逆行することなく参照信号の時刻にサンプル信号の時刻が対応付けられている。

Table 3.3: Time correspondence between the referecce signal and the sample signal.(good case)

reference signal	···	96	97	98	99	100	101	102	103	104	···
sample signal	···	102	102	102	102	102	102	102	106	106	···

以上のように、時間区間上での ウェーブレット係数の比較に基づたり、時間発展を考慮したりして、時刻の対応付けを行うことにより、時系列の標準化を良好に実現できるようになった。

3.6 人形の複数の部位を考慮した対応付け

ここでは前節で考案した時間区間や時間発展を考慮した対応付けの手法を用いて、実際の文楽人形の時系列データの標準化を行なう。文楽人形の動作には「挨拶、お茶出し」の2種類の機能それぞれについて「普通、好意、敵意」の3種類の情緒があるが、まず最初に「挨拶・普通」の時を考える。

3.6.1 参照信号の選出

本研究における時系列データの標準化は、1本の参照信号にその他のサンプル信号の動作時間を合わせるものであるため、ここではまず参照信号を選出する。参照信号として望まれるものは、動作時間が長く、他の信号と比較して動作 (信号の波形) に際だった特異点がないことである。

文楽人形の時系列データは、測定部位である「頭、肩、腕、手」の4部位それぞれが、位置・姿勢の6自由度を持っているものである。次頁 Fig.3.16 から Fig.3.21 に「挨拶・普通」の場合の文楽人形の時系列データとして、測定部位が手（式 (2.9) において $p=3$ である）のデータを示す。

ここで、測定回数の異なる時系列データを式 (2.9) のように $x_{k(pq)}(t)$ $(k=1,\cdots,5)$ とし、それぞれの図において上の信号から順に $x_{1(pq)}(t)$, $x_{2(pq)}(t)$, $\cdots$, $x_{5(pq)}(t)$ とする。Fig. 3.16から Fig. 3.21を見ると、信号 $x_{1(pq)}(t)$ が他の4本の信号と比べて明らかに動作開始時刻が遅れていることが分かる。これより、参照信号として $x_{1(pq)}(t)$ は不適切である。また $x_{1(pq)}(t)$ 以外の信号は比較的どれも近い波形をしているとみなせるが、ここでは動作開始時刻も早く、動作時間も長い信号 $x_{4(pq)}(t)$ を参照信号として選ぶことにする。以後、$x_{4(pq)}(t)$ を参照信号とし、残りの信号の標準化を行なう。

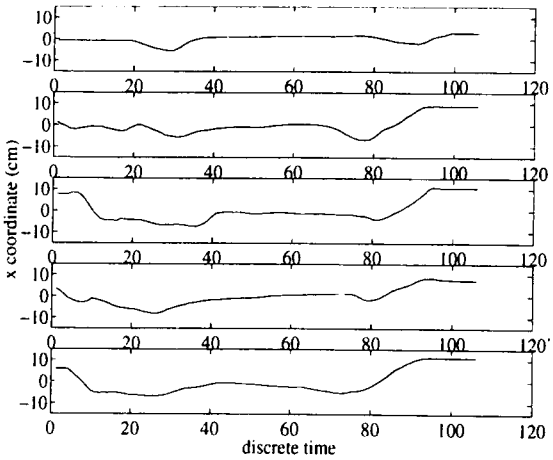


Figure 3.16: $x_{(41)}$ $(i = 1, j = 1)$

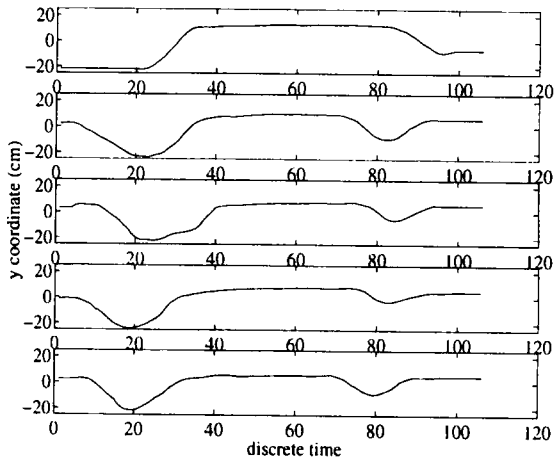


Figure 3.17: $x_{(42)}$ $(i = 1, j = 1)$

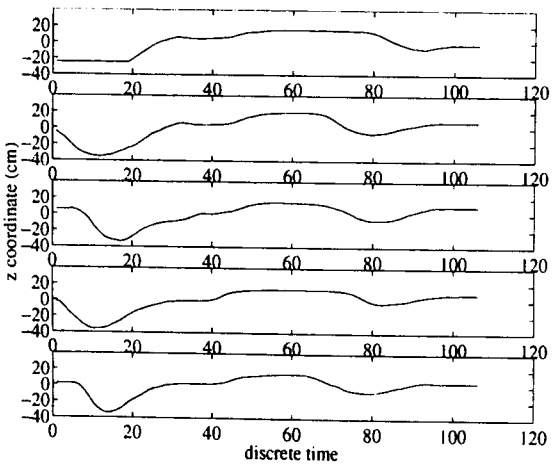


Figure 3.18: $x_{(43)}$ $(i = 1, j = 1)$

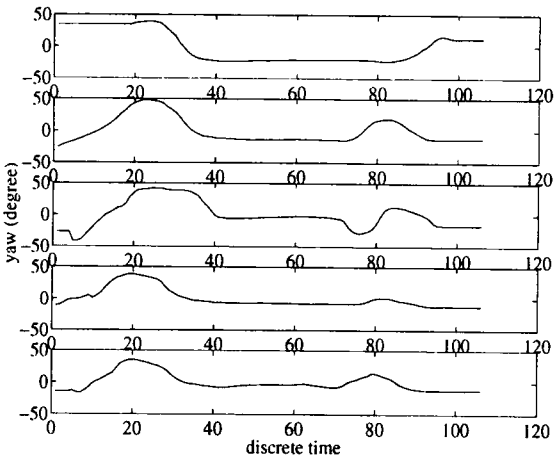


Figure 3.19: $x_{(44)}$ $(i = 1, j = 1)$

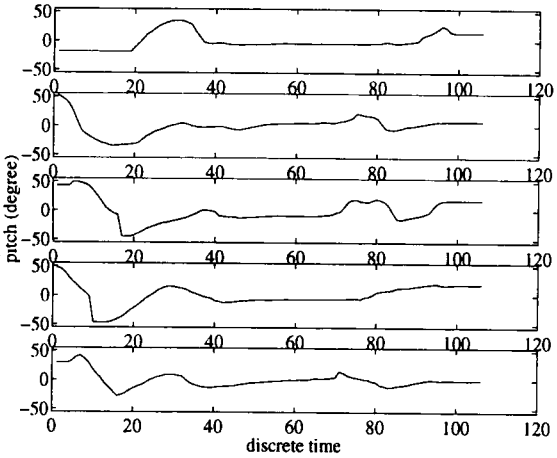


Figure 3.20: $x_{(45)}$ $(i = 1, j = 1)$

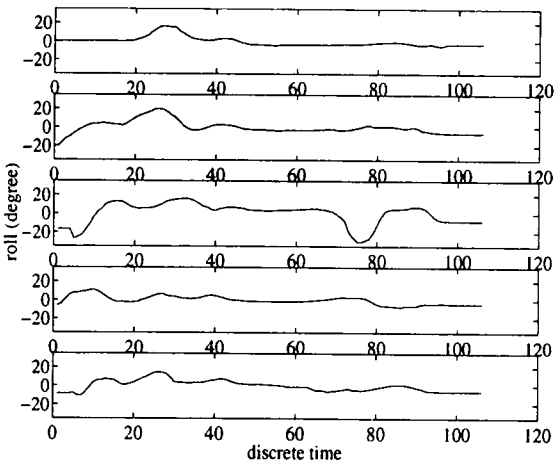


Figure 3.21: $x_{(46)}$ $(i = 1, j = 1)$

3.6.2 各部位ばらばらの標準化

信号 $x_{4(pq)}(t)$ を参照信号とし、その他のサンプル信号の標準化を行なう手順を述べる。文楽人形の時系列データは、測定部位 4ヶ所 $(p = 1, 2, 3, 4)$ がそれぞれ 6 自由度 $(q = 1, 2, \cdots, 6)$ を持っているが、標準化を行なう時は、それぞれの部位 p の各自由度 q ごとに個々ばらばらに行なうものとする。

部位 p　自由度 q の参照信号 $x_{ref(pq)}(t)$ の時刻 t におけるウェーブレット係数たちのベクトルを、式 (3.3) のように

$$d_{ref(pq)}(t) = \Big(d_{ref(pq)}(t, c_l); l \downarrow 1, 2, \cdots, L\Big)$$

とし、サンプル信号 $x_{smp(pq)}(t)$ のそれを、式 (3.27) のように

$$d_{smp(pq)}(t) = \Big(d_{smp(pq)}(t, c_l); l \downarrow 1, 2, \cdots, L\Big)$$

とする。

それぞれの部位 p の各自由度 q ごとに個々ばらばらに、参照信号 $x_{ref(pq)}(t)$ の時刻 t_{ref} に対して、サンプル信号 $x_{smp(pq)}(t)$ の時刻 t_{smp} を対応付けるにあたって、式 (3.21) のように

$$\sum_{s=-\tau}^{\tau} \Big| d_{ref(pq)}(t_{ref} + s) - d_{smp(pq)}(t_{smp} + s) \Big|^2 \tag{3.27}$$

を最小化するようにした。ここで、ウェーブレット係数を比較する時間区間に関しては $\tau = 4$ とし、対応付けを探索する範囲の式 (3.20) に関しては $T_{cmp} = 20$ とし、対応付けにあたっては時間発展をもちろん考慮した。

この対応付けに従って、式(3.3)のようにサンプル信号を標準化した信号を、$x_{nor(pq)}(t)$ で表す。測定部位が手の場合の標準化結果を Fig.3.22 から Fig.3.27 に示す。ただし、図の実線は標準化された信号で、破線は元のサンプル信号を表す。元のサンプル信号と比較すると、標準化された信号は多少のずれは見られるが、動作時刻はほぼ $x_{4(pq)}(t)$ に揃っているとみなせる。ただし、$x_{1(pq)}(t)$ は含めない。

そして、各部位 p 各自由度 q 毎にばらばらに標準化して得た信号 $x_{nor(pq)}(t)$ の全てから、式 (3.15) や式 (3.17) のように、人形全体としての動作時系列 $x_{nor}(t)$ を

$$x_{nor}(t) \;=\; \Big(x_{nor(p)}(t); p \downarrow 1, \cdots, 4\Big)$$

$$x_{nor(p)}(t) \quad = \quad \left(x_{nor(pq)}(t); q \downarrow 1, \cdots, 6\right)$$

として構成し、computer graphics 上に人形の動作を生成してみると、Fig. 3.28のように、腕 $(p = 3)$ と手 $(p = 4)$ が離れた不適切な動作が見られた。

それぞれの自由度の信号について、適切な標準化を行うことができても、文楽人形の時系列データ全体として見た時に動作として成り立たたなければ、適切な標準化が行なわれているとはいえない。

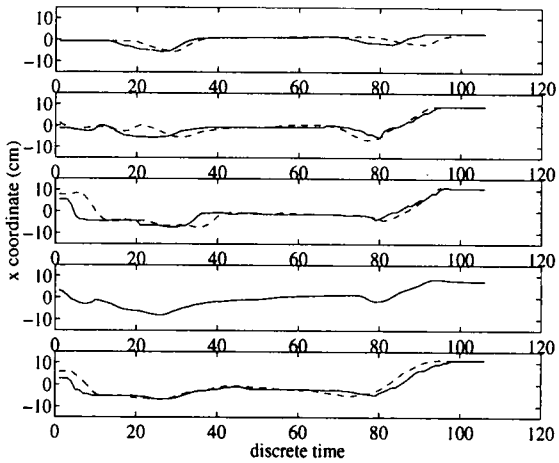


Figure 3.22: Normalized Figure 3.16.

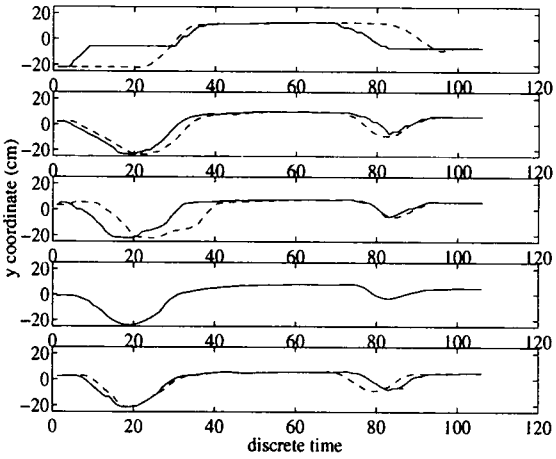


Figure 3.23: Normalized Figure 3.17.

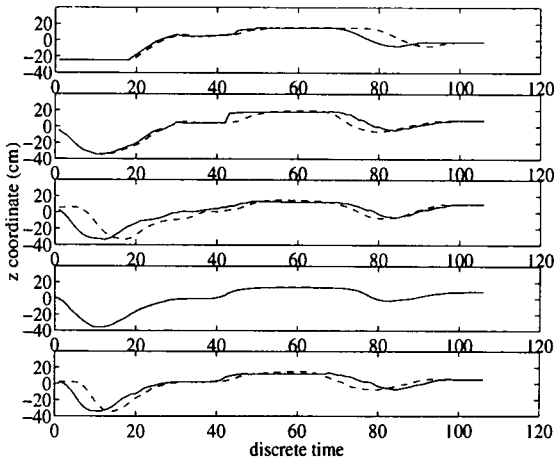


Figure 3.24: Normalized Figure 3.18.

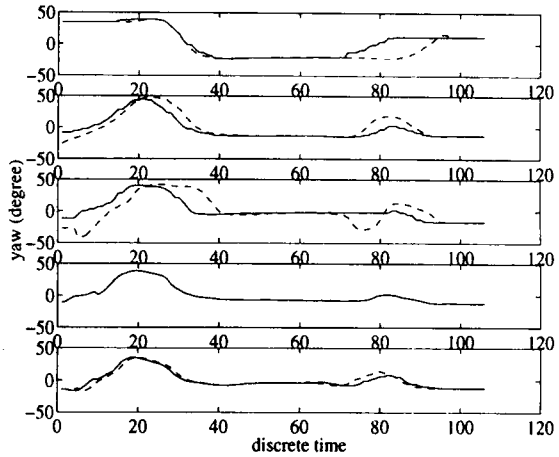


Figure 3.25: Normalized Figure 3.19.

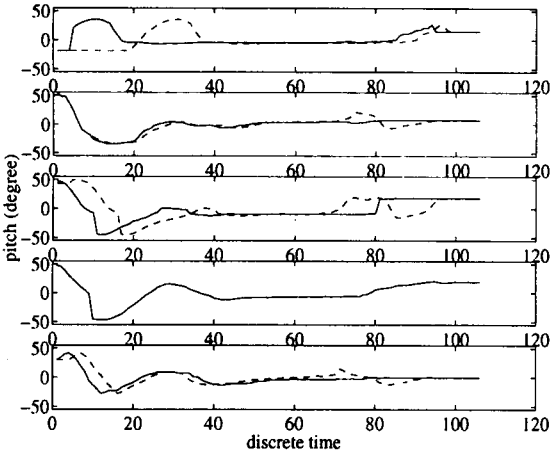


Figure 3.26: Normalized Figure 3.20.

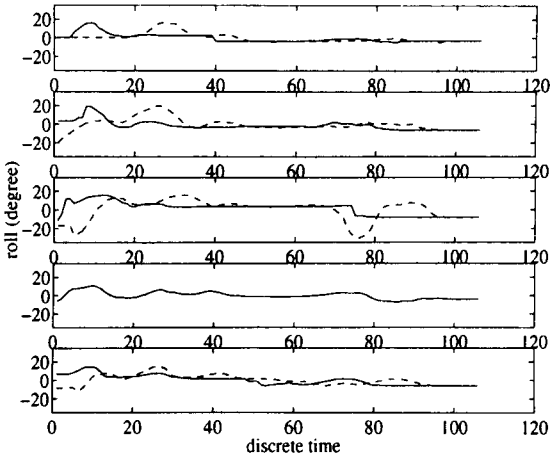


Figure 3.27: Normalized Figure 3.21.

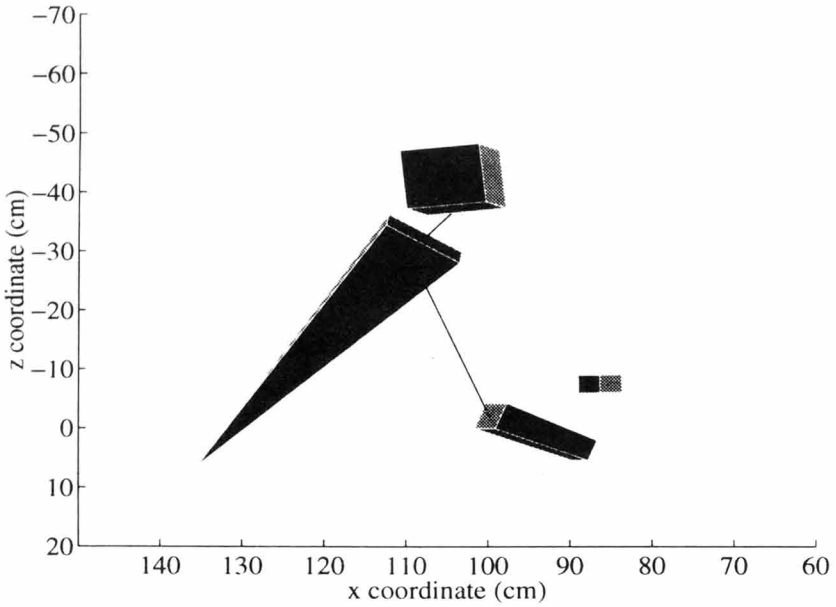


Figure 3.28: A bad example of the normalized action.

3.6.3 複数の部位を同時に考慮した標準化

前節では、標準化された時系列データの動作確認を行なったところ、腕と手が離れるという不適切な結果が得られた。文楽人形の動作において、腕と手が離れるということはあり得ないので、標準化を行なうにあたり、腕の時系列のウェーブレット係数と手の時系列のウェーブレット係数とを同時に考慮する事にした。

つまり、参照信号の時刻 t_{ref} に対してサンプル信号の時刻 t_{smp} を対応付けるにあたって、単に式 (3.27) を最小化するように t_{smp} を選ぶのではなく、腕 $(p=3)$ と手 $(p=4)$ を同時に考慮して

$$\sum_{p=3}^{4}\sum_{s=-\tau}^{\tau}\left|d_{ref(pq)}(t_{ref}+s)-d_{smp(pq)}(t_{smp}+s)\right|^2 \tag{3.28}$$

を最小化するように t_{smp} を選ぶようにした。ここで、それぞれの部位 p と座標 q に対して、$d_{ref(pq)}(t_{ref})$ は参照信号 $x_{ref(pq)}(t)$ の時刻 t_{ref} におけるウェーブレット係数であり、$d_{smp(pq)}(t_{smp})$ はサンプル信号 $x_{smp(pq)}(t)$ の時刻 t_{smp} におけるウェーブレット係数である。

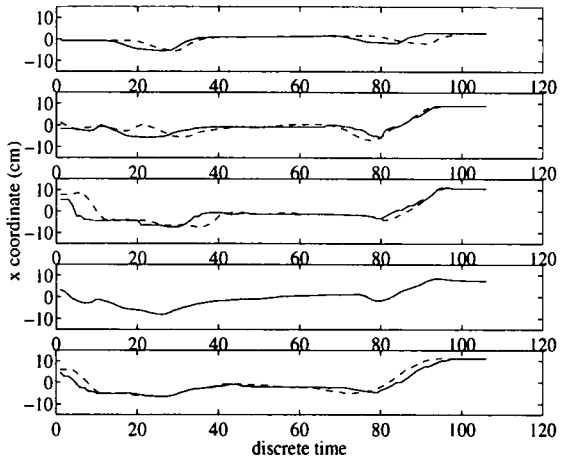


Figure 3.29: Normalized Figure 3.16.

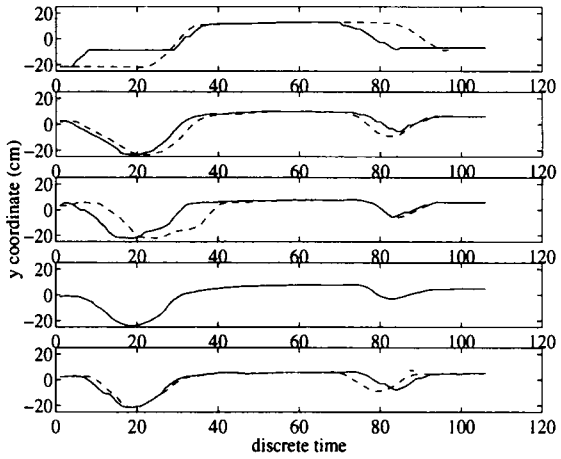


Figure 3.30: Normalized Figure 3.17.

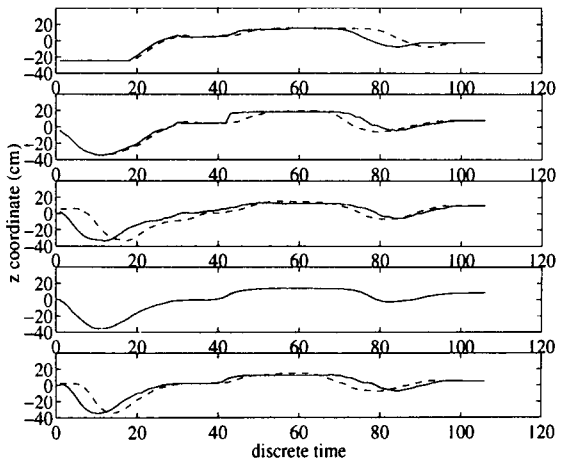


Figure 3.31: Normalized Figure 3.18.

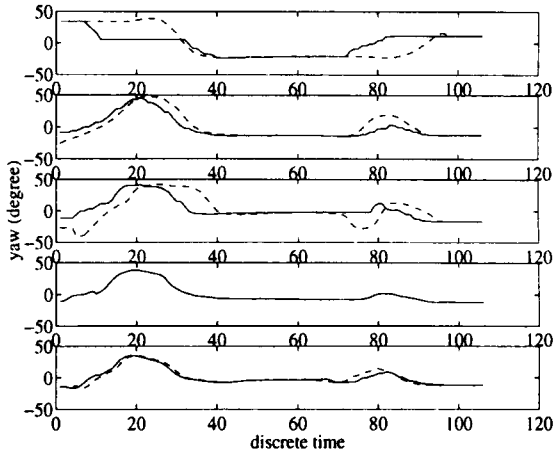


Figure 3.32: Normalized Figure 3.19.

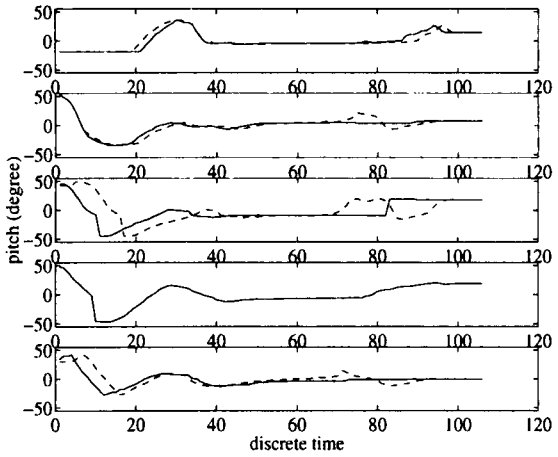


Figure 3.33: Normalized Figure 3.20.

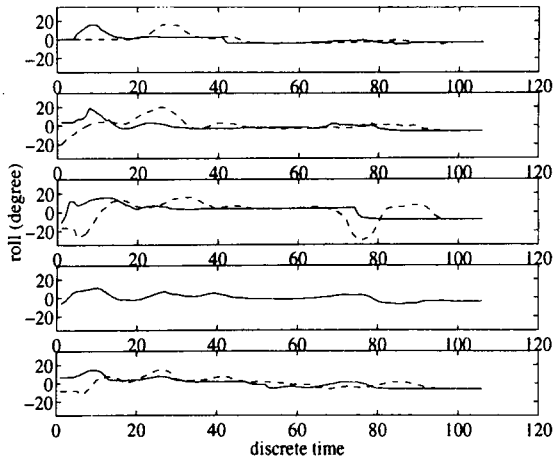


Figure 3.34: Normalized Figure 3.21.

このようにして再び全てのサンプル信号の標準化を行なった。つまり、腕と手の部分の標準化結果が以前と異なる。測定部位が手の場合の結果を、前ページの Figure 3.29 から Figure 3.34 に示す。実線が標準化された信号で、破線が元のサンプル信号である。

これら標準化された時系列データを用い、文楽人形の動作確認を行なった。その結果、以前のように腕と手が離れることがなくなり、挨拶動作として成り立つものになり、動作全体として見た場合でも、動作がぎこちない部分は見られなかった。また、標準化されたデータは、動作時間が信号 $x_{4(pq)}(t)$ に揃っており、この点で適切に標準化が行なわれたということが分かった。

次に、この腕と手を同時に考慮する標準化を用いて、「挨拶・好意」動作の場合の標準化を行なった。そして、標準化された時系列データの動作確認を行なったところ、腕と手の部分は離れることはなかった。しかし、文楽人形の頭と肩 (胴体) の部分が離れることがあり、適切な挨拶動作を行なっているとは考えにくい。頭が肩 (胴体) から離れている一例を Fig. 3.35に示す。

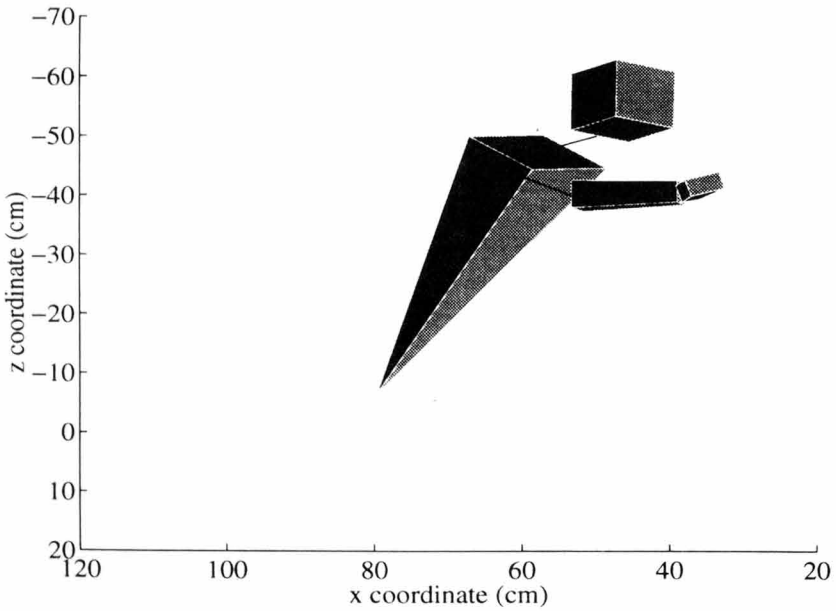


Figure 3.35: A bad example of the normalized action.

この問題を改善するために、腕と手が離れた場合と同様に、頭と肩 (胴体) の標準

化を同時に行うこととする。つまり、参照信号の時刻 t_{ref} に対してサンプル信号の時刻 t_{smp} を対応付けるにあたって、頭 $(p=1)$ と 肩 (胴体)$(p=2)$ を同時に考慮して

$$\sum_{p=1}^{2}\sum_{s=-\tau}^{\tau}\left|d_{ref(pq)}(t_{ref}+s)-d_{smp(pq)}(t_{smp}+s)\right|^2 \tag{3.29}$$

を最小化するように t_{smp} を選ぶようにした。

つまり全体として、「頭と肩、腕と手」の標準化を同時に行なうこととする。この様にしてサンプル信号の標準化を行ない、その時系列データの動作確認を行なった。その結果、「挨拶・情緒なし」と「挨拶・好意」のどちらの場合でも、全体的に見て適切な挨拶動作を行なっているとみなすことができた。そして動作時間も参照信号に揃っており、適切な標準化が行なわれたということが確認できた。

ここで「腕と手」だけを同時に考慮して標準化を行うものを手法 1 と呼び、「頭と肩、腕と手」の両方を同時に考慮して標準化を行うものをを手法 2 と呼ぶことにする。挨拶動作において、情緒が「情緒なし」、「好意」のどちらの場合でも適切な標準化としては手法 2 が考えられるが、「挨拶・情緒なし」の場合には手法 1、手法 2 ともに適切な動作を行なっているので、これら 2 種類の手法を用いて時系列データを標準化し、今後それらについて比較、考察を行なうこととする。

3.7 平均動作時系列の作成方法

Figure 3.10 のように、同じ条件のもとで複数回（K 回）観測して得られた K 本の時系列 $x_k(\cdot)$ $(k = 1, 2, \cdots, K)$ について、同一の事象に対応する時刻も振幅値もサンプル k が変われば異なる値になる。その意味で、時系列は時刻についても振幅値についても確率的に揺らいでいるといえる。それらのうちの 1 本を参照信号とし、その他の時系列（サンプル信号）を標準化する方法は、前節までで確立された。そこで今度は、それらの複数の時系列から平均となる時系列を計算する方法を考える。

まず K 本の時系列から基準となるものを 1 本定める。ここでは一例として、サンプル $k = 1$ を基準とするが、それ以外のサンプルを基準とする場合でも、平均の時系列の計算方法は以下と同じである。

サンプル $k = 1$ の時系列 $x_1(t)$ $(t = 0, 1, 2, \cdots, T)$ が基準として選ばれている。サンプル k の時系列の任意の時刻 t_1 とその振幅値 $x_1(t_1)$ の事象に対応する サンプル k $(k \neq 1)$ の事象の時刻と振幅値を、上記で述べられた ウェーブレット係数を用いた標準化法で求める。この方法によって、サンプル $k = 1$ の時刻 t_1 に対応する サンプル k の時刻が t_k として求まったものとする。

このように、時刻 t_1 に対応する時刻が t_k $(k = 2, 3, \cdots, K)$ として求まると、振幅対 $x_1(t_1)$ に対応する振幅値は $x_k(t_k)$ $(k = 2, 3, \cdots, K)$ という事になる。前説までで述べられた考え方と方法に従って、$x_k(t_k)$ $(k = 1, 2, \cdots, K)$ が同一の事象を表していると判断されたことになる。それらの平均を計算して

$$\bar{t} = \sum_{k=1}^{K} t_k / K \tag{3.30}$$

$$\bar{x} = \sum_{k=1}^{K} x_k(t_k) / K \tag{3.31}$$

とおく。すると、時刻 $\bar{t}$ の時の振幅値が $\bar{x}$ である時系列が、K 本の時系列 $x_k(\cdot)$ $(k = 1, 2, \cdots, K)$ の平均時系列と考えるのが自然である。

任意の $t_1 = 0, 1, 2, \cdots, T$ について、上記に述べられた時刻の平均 $\bar{t}$ と状態の平均 $\bar{x}$ を求めて、横軸を $\bar{t}$ とし縦軸を $\bar{x}$ とするグラフで表される時系列 $\bar{x}(\bar{t})$ を、$t = 0, 1, 2, \cdots, T$ に線形補間した $\bar{x}(t)$ $(t = 0, 1, 2, \cdots, T)$ を、K 本の時系列 $x_k(\cdot)$

$(k = 1, 2, \cdots, K)$ の平均時系列と定義する。

文楽人形の動作時系列の標準化にあたって、これまでは参照信号として $x_4(t)$ を用いていたが、上記のようにして作成した平均時系列も参照信号として標準化を行う事を考える。このとき前述したように「挨拶・情緒なし」の場合、信号 $x_1(t)$ は他の信号と比べて動作開始時刻がかなり遅く、明らかに特異であるため、平均をとるうえで信号 $x_1(t)$ を含める場合と、含めない場合で結果が異なると考えられる。このため、5 本の平均による参照信号 (5ave) と、信号 $x_1(t)$ を除く 4 本の平均による参照信号 (4ave) をそれぞれ作成する。また、「挨拶・情緒なし」の場合は標準化する方法を 2 種類用いるのであるが、手法 1 による参照信号を ref1、手法 2 による参照信号を ref2 とする。つまり、ref1-5ave、ref1-4ave、ref2-5ave、ref2-4ave の参照信号を作成した。そしてこれらの参照信号の動作確認をしたところ、ほとんど違いは見られず、動作としては適切な挨拶動作をしているとみなすことができた。今後はこれらの参照信号を用いて、時系列データの標準化を行なう。

3.8 標準化された信号の評価

新たに作成した4種類の参照信号を用いて5本のサンプル信号の標準化を行ない、得られた時系列データの動作確認を行なったところ、信号 $x_1(t)$ が若干ずれる程度で、全体としてはどれも適切な挨拶動作を行なっており、動作時間も参照信号に揃っていることが確認できた。しかし動作確認という評価においては、参照信号として $x_4(t)$ を用いて標準化を行なった場合と比較しても、あまり違いが見られることはなかった。

標準化後の信号として、どの場合が最も適切な標準化信号であるかを調べるため、ここでは信号の分散を用いて評価する。分散は

$$\sum_{t=0}^{T}(x_k(t) - x_{ref}(t))^2 \tag{3.32}$$

で定義される。ここで、k は測定回数 $(k = 1, 2, \cdots, K)$ であり、x_{ref} は参照信号である。$x_4(t)$ が参照信号の時は $x_{ref}(t) = x_4(t)$ である。式 (3.32) は、ある参照信号に対する標準化信号の離れ具合である。つまり分散が小さい程、標準化された信号が参照信号に近いということである。

式 (3.32) を用いて、「挨拶・情緒なし」の時の標準化された時系列データそれぞれの分散値を求めた。計算結果を次に示す。ただし、Table 3.4 は腕と手」だけを同時に考慮して標準化を行った手法1による標準化であり、Table 3.5は「頭と肩、腕と手」の両方を同時に考慮して標準化を行った手法2による標準化である。

Table 3.4: Each variance of normalized signals by method 1.

reference	$x_4(t)$		ref1-5ave		ref1-4ave	
signal	sample	normalized	sample	normalized	sample	normalized
$x_1(t)$	3.0077	2.4348	2.4151	1.6071	3.1817	2.7875
$x_2(t)$	1.0492	0.5795	0.7919	0.2605	0.7819	0.2720
$x_3(t)$	2.2888	0.5526	2.1837	0.4749	1.9348	0.2750
$x_4(t)$	——	——	0.2341	0.2126	0.1897	0.1705
$x_5(t)$	1.1202	0.5632	0.8681	0.3299	0.7631	0.2573

$(\times 10^5)$

Table 3.5: Each variance of normalized signals by method 2.

reference	$x_4(t)$		ref2-5ave		ref2-4ave	
signal	sample	normalized	sample	normalized	sample	normalized
$x_1(t)$	3.0077	2.4147	2.4177	1.6123	3.1847	2.7925
$x_2(t)$	1.0492	0.6088	0.7849	0.2653	0.7739	0.2688
$x_3(t)$	2.2888	0.5551	2.1759	0.4597	1.9301	0.2969
$x_4(t)$	——	——	0.2427	0.2247	0.1990	0.1872
$x_5(t)$	1.1202	0.5905	0.8721	0.3218	0.7660	0.2568

$(\times 10^5)$

Table 3.4、Table 3.5のどちらの場合も、参照信号が $x_4(t)$ の時の標準化より、平均により作成した参照信号の標準化の時の方が、分散は小さくなっている。ただし、信号 $x_1(t)$ は除く。このことより、参照信号として $x_4(t)$ より平均による参照信号の方が、信号全体の参照信号として適切であることが分かる。

また、参照信号が 5ave の場合と、4ave の場合を比較すると、$x_1(t)$、$x_2(t)$ を除くその他の信号では、4ave の時の方が分散がかなり減少している。ここで信号 $x_1(t)$ における分散の増加原因は、参照信号 4ave には信号 $x_1(t)$ が全く反映されていないためであると考えられる。そして信号 $x_2(t)$ での分散の増加の割合について考えると、これは信号 $x_3(t)$〜$x_5(t)$ での分散の減少の割合に比べて、比較的少ないものとみなせるので、「挨拶・情緒なし」の場合では、参照信号として 4ave が最も適切な参照信号であると考えられる。

次に「腕と手」の標準化を同時に行なう手法 1 と、「腕と手、頭と肩」の標準化を同時に行なう手法 2 を比較する。ここで、分散の減少の割合を比較すると、手法 1 の方が手法 2 よりも多く、手法 1 の方が適切であると考えられる。しかし、「挨拶・情緒なし」動作以外の場合でも適切な標準化を行なうためには、手法 2 が適していることは前述した通りであり、また、分散の減少の割合も手法 1 と比較しても大きな相違はないため、文楽人形の時系列データ全体を標準化する方法として、手法 2 が最も有効なものと考えられる。

3.9 標準化された動作時系列に対する考察

時系列の各時刻のウェーブレット係数に基づいて、全ての動作の時系列データの標準化を行なった。ここで文楽人形の動作には、「挨拶、お茶出し」の2種類の機能それぞれについて、「情緒なし、好意、敵意」の情緒があり、用いた標準化アルゴリズムは、「頭と肩、腕と手」の標準化を同時に行なうアルゴリズムである。

各機能・情緒ごとに、参照時間区間 $\tau = 4$、比較範囲 $T_{cmp} = 20$ として全てのサンプル信号の標準化を行なった。

そしてこれらの標準化信号の動作確認を行なったところ、ほとんど全ての信号については、動作時間が揃い、かつ適切な動作を行なっているものとみなせた。しかし、中には動作が少しぎこちない信号が見られ、この信号については適切な標準化が行なわれていないものと考えられる。

そこでこの信号について、参照範囲と比較範囲を様々な値に変えて再び標準化を行なった。その結果、ある参照範囲と比較範囲において、以前よりも適切な動作を行なうことが確認できた。参照範囲とは、信号間でどれだけの成分を比較するかという基準であり、これが小さいと信号の持つ特徴があまり反映されず、大きいと細かな部分の特徴を見失う恐れがある。また比較範囲とは、任意の時刻から前後数時刻までを比較・対応づける基準であり、これが小さいと必要な時刻を対応づけられず、大きいと余分な時刻を対応づける恐れがある。

つまりサンプル信号に適した参照範囲と比較範囲は、それぞれ微妙に異なるため、現在の標準化信号に用いている「参照範囲 $\tau = 4$、比較範囲 $T_{cmp} = 20$」というのが最適であるとは断定できない。今後、最適な標準化を行なうために、様々な参照範囲と比較範囲を調べる価値があると考えられる。

また標準化を行なううえで、時系列データそれぞれが個々で意味を持つ場合は、標準化が極めて単純である。しかし文楽人形の演技動作のように、合計24自由度の時系列データそれぞれがお互いに関連し、全体として一つの意味を持つような場合は、標準化は単純なものにはならず、全体としてまとまりのある標準化を行なわなければならない。

本章では、文楽人形の動作が適切に行なえるように、様々な制約を付けて標準化を行なった。現在、「頭と肩、腕と手」の標準化を同時に行なうという方法で、標準化信号による動作が適切な動きをしていることが確かめられている。この方法は、4ヶ所の測定部位のうち空間的に離れ得ない部位を、それぞれお互いに関連づけ、その時の位置と姿勢は別々に標準化しているものである。ここで位置とは各 x、y、z 座標であり、姿勢とは各軸周りの回転を表す *yaw*、*pitch*、*roll* である。しかし、現在の方法以外にも

- 「全ての部位」の標準化を同時に行なう

また、位置と姿勢に関しては

- 「各部位での位置と姿勢」の標準化を同時に行なう
- 「各自由度」の標準化を同時に行なう
- 「全ての位置と姿勢」の標準化を同時に行なう

など様々な制約が考えられる。

つまり複数の自由度を持ち、全体として一つの意味を持つような時系列データの標準化は複雑で困難なため、現在の文楽人形の時系列データの標準化信号も、上で挙げたような新たな制約を付けることにより、さらに最適な標準化が行なえる可能性があり、今後様々な標準化方法を行なうべきである。

3.10 結言

同じ条件のもとで複数回観測された時系列たちから、同一の事象に対応する時刻たちを求めて、標準化（時刻合わせ）を行う方法を提案した。鍵となる考え方は、時系列たちのグラフが局所的に似ている時刻を対応させるという点であり、グラフの局所的な形を数量的に表す物として、時系列の各時刻における種々のスケールのウェーブレット係数に着目した。つまりウェーブレット係数が近い時刻どうしを対応する時刻と考えて、時系列の標準化を行う手法を提案した。比較するウェーブレット係数のスケールを適切に選ぶことによって、高周波雑音の影響を受けない時系列の標準化が可能である。

これら時系列の標準化手法を、文楽人形の動作時系列に対して適用し、標準化された動作時系列のグラフを検討し、それらに基づいてコンピュータ・グラフィックス上に生成された動作を観察することにより、上記の手法の問題点について考察し、標準化手法を以下のように改良した。

- ウェーブレット係数が近い時刻どうしを対応する時刻と考えるにあたって、単一の時刻どうしのウェーブレット係数を比較するだけでなく、時間区間上のウェーブレット係数を比較して、参照信号とサンプル信号の時刻の対応を求める。
- 参照信号の時刻が時間発展しているのに、それに対応するサンプル信号の時刻が逆行することのないような制約条件の下で、ウェーブレット係数が最も近い時刻を対応付ける。
- 文楽人形の動作時系列を標準化するにあたっては、人形の単独の部位の時系列だけでウェーブレット係数が最も近い時刻を対応付けるのではなく、複数の部位の時系列の複数のウェーブレット係数が最も近い時刻を対応付ける。

また上記の標準化手法を用いて、時刻と状態の双方が確率的に揺らいでいる複数の時系列から平均の時系列を求める方法を提案した。

提案された時系列標準化や平均時系列の計算手法は、文楽人形の演技動作の解析の今後の研究に有効であるだけでなく、異なる種類の複数の信号から対応する時刻

を求める必要のあるセンサフュージョンや音声信号・ロボットへの教示信号などの応用分野の信号処理においても有用である。

第 4 章　線形構造モデルに基づく機能・情緒・揺らぎの解析

4.1　緒言

人間型のアミューズメント・ロボットが生き生きとした情緒豊かな表情をみせるためには、それらのロボットの動作を設計する手法が大切になる。人間の動作をモーション・キャプチャなどで測定して得られた動作時系列を使用する事が広く行われているので、そのような測定された動作時系列を、アミューズメント・ロボットで使用するために、望むように加工する技術が大切になる。動作をコンピュータ・グラフィックスなどで確認しながら、動作時系列を試行錯誤で修正していくよりも、もっとシステマティックな手法の確立が望まれる。

そこで本章では、様々な機能や情緒を表現すべくプロの人形遣いが操作して生成した文楽人形の演技動作の情緒を強調したり減衰させたりして、新しい動作を計算する一つの手法を提案する。文楽人形の演技動作の時系列から、機能の要素と情緒の要素を抽出し、情緒の要素を強調させたり減衰させたりして、新しい動作時系列を生成させる方法を考案する。

機能や情緒の要素に関して、具体的に述べる。「好意を持って挨拶する」・「敵意を持ってお茶を差し出す」などの動作は、「挨拶する」・「お茶を差し出す」などの機能の要素と「好意」・「敵意」などの情緒の要素から構成されると考えられる。つまり

(好意で挨拶する)=(挨拶する機能)+(好意の情緒)

と考える。情緒なしで挨拶する動作は、挨拶する機能そのものであると考えて、

(情緒なしで挨拶する)=(挨拶する機能)

と考える。従って、挨拶するという動作における好意の情緒の要素は、好意で挨拶

する動作と挨拶する機能の要素との差であるということになる。

機能や情緒の要素を具体的に求める事ができれば、機能の要素をそのままにして、情緒の要素だけ強調して新しい動作を計算する事により、情緒を強調した動作を生成する事ができる。情緒を減衰させて動作を生成する事も、同じく可能である。

このような考え方に基づいて、観測された動作時系列から、機能や情緒の要素を計算する方法の概略を、以下に述べる。文楽人形の演技動作を解析するにあたって、本章では、動作時系列が、機能・情緒・確率的な揺らぎの要素から構成されていると考えている。例えば「好意を持って挨拶をする」動作を考えた場合に、対象相手に向かって動作主の上半身を前方にかがめるという機能的要素と、動作主の好意という情緒を対象相手に伝えるという情緒的要素が存在する。また、同じ機能、情緒を伴った動作を数回行うと、それらの動作全てが全く一致するわけではない。機能的、情緒的役割への影響はなく、繰り返すごとによって、動作が微妙に変化する残差が存在する。それを、確率的要素と呼ぶことにする。

そこで、文楽人形の動きを測定して得られた動作時系列を、機能的要素、情緒的要素、確率的要素に分解する必要がある。まず、最も標準的手法である最小2乗法を用いて動作時系列から各要素を抽出する。その中の確率的要素には機能に依存しない情緒的な要素が含まれていると考え、その統計的な性質を解析する。

次に、確率的要素のモデルとして、広い範囲の確率課程を表現できる KM_2O-ランジュバン方程式 [14] [15] を採用する。これを用いて、文楽人形の動作時系列が、機能・情緒の外生入力を持つ KM_2O-ランジュバン方程式で表せることを提案する。そして、マルコフ推定を用いた外生入力の同定手法を確定し、数値シミュレーションにて手法の有効性を示す。

最後に、文楽人形の動作時系列から同定された外生入力から、機能的要素と情緒的要素を計算する。そして、同定された情緒的要素に重みを加えて、特定の情緒を強調したり減衰させたりする情緒的動作の生成を行なう。

4.2 動作を構成する 機能・情緒・揺らぎの要素

緒言で述べたように、文楽人形の動作時系列に関して、「人間の動作は、機能的要素、情緒的要素、確率的要素から成り立っている」という仮説のもとで議論を進める。文楽人形の、日常動作的な側面を機能と呼び、役柄の持つ精神的側面を情緒と呼んで、機能と情緒の２つの要素を動作が有していると考えている。同一の「機能的要素」「情緒的要素」を持った動作を何回も繰り返した時、その動作は全く同じものにならない。動作の基本的な機能、情緒は同じでも、動作にかかる時間、動きの大きさといった点では少しずつ異なってくる。つまり、動作の中には、「機能的要素」「情緒的要素」とはまた異なった要素が含まれていると考えられる。これは、動作の本質的なものとは違う、動作する度に変わる残差のようなものであり、ある確率分布にしたがった、「確率的揺らぎ」であると言え、これを「確率的要素」と呼ぶことにする。そうすると、ある機能、情緒を持つ動作は、無限のパターンの中から確率的に選ばれた一つの実現であると考えられる。これが、前節で文楽人形の動作を表す時系列を確率過程とした理由である。

これにより、機能的要素、情緒的要素、確率的要素の意味について述べたが、文楽人形の動作時系列を解析するにあたり、本研究では、さらに３つの仮説を立て、それぞれの意味を説明する。

(仮説 1)
人形の動作を表す時系列が、(1) 機能的要素、(2) 情緒的要素、(3) 確率的要素の３つの要素に分解される。

(仮説 2)
機能的要素は情緒的要素に依存しないが、情緒的要素は機能的要素つまり動作そのものに依存する。

(仮説 3)
機能的要素と機能に依存する情緒的要素は deterministic である。

(仮説 1) は先程述べた通りである。(仮説 2) では情緒的要素は機能的要素に依存するとあるが、これは「挨拶をする」という動作に含まれる「好意」「敵意」と、「お茶を渡す」という動作に含まれる「好意」「敵意」とは違うことを示している。つまり、ある情緒的動作が決定すると、その動作特有の感情が決定することである。そこで (仮説 1) を (仮説 2) に基づいて修正し、新たな (仮説 1) とする。

(仮説 1)
人形の動作を表す時系列が、(1) 機能的要素、(2) 機能に依存する情緒的要素、(3) 確率的要素の 3 つの要素に分解される。

(仮説 3) は、機能的要素、情緒的要素は繰り返し動作を行なっても、そのサンプル回数に依存せずに決定しているということである。

この 3 つの仮説を人形の動作時系列の確率過程の分解として表現すると、次式のようになる。

$$x_k^{ij}(t) = y^i(t) + z^{ij}(t) + \eta_k^{ij}(t) \tag{4.1}$$

$$y^i(t) : \text{機能的要素}$$
$$z^{ij}(t) : \text{機能に依存した情緒的要素}$$
$$\eta_k^{ij}(t) : \text{確率的要素}$$
$$i = 1, \cdots, I;\ j = 1, \cdots, J;\ k = 1, 2, \cdots, K$$
$$I = 2,\ J = 3,\ K = 5$$
$$t = 0, 1, \cdots, T \quad T : \text{終端時刻}$$

ここで機能的要素 $y^i(t)$ は、非常にはっきりした dynamics を有する非定常な要素であると考えられる。「挨拶する」「お茶を渡す」などのはっきりとした目的を持った動きであり、dynamics そのものであると考えられる。機能に依存した情緒的要素も、dynamics のはっきりした非定常な要素に依存しているという点で、ある程度の非定常性を有していると考えられる。

そして j=1 は「感情を込めない」情緒であるが、これに対応する機能に依存した情緒の要素はゼロであると考えられる。そこで次の仮説を提唱する。

(仮説 4)
感情を込めない情緒に対応する「機能に依存した情緒的要素」はゼロである。つまり

$$z^{i1}(t) = 0 \quad (i = 1, \cdots, I;\ t = 0, 1, \cdots, T) \tag{4.2}$$

最後に確率的要素 $\eta_k^{ij}(t)$ は、各回数の動きのばらつきだけでなく、機能に依存しない情緒的要素 (情緒そのもの) を含んでいるのではないかと考えられる。確率的要素は、はっきりした dynamics を有しない定常な要素であると推測される。

次節では、文楽人形の動作時系列 $x^{ij}(t)$ を機能的要素 $y^i(t)$、情緒的要素 $z^{ij}(t)$、確率的要素 $\eta_k^{ij}(t)$ への分解を試みて、

$$x^{ij}(t) = y^i(t) + z^{ij}(t) + \eta^{ij}(t) \tag{4.3}$$

と表すことを行なう。

本節では、文楽人形の動作は、機能的要素、情緒的要素、確率的要素から成り立っていると仮説をおき、各要素の意味を説明した。以下の節では、文楽人形の動作時系列から機能的要素、情緒的要素、確率的要素の抽出を 2 つの方法により行なう。そして、最も各要素の抽出に適した方法を選び、その方法により抽出した機能的要素と情緒的要素から、情緒的要素を強調させることで、情緒豊かな動作を生成する。

4.3 揺らぎの最小化による機能・情緒の要素の抽出

4.3.1 動作時系列の各要素の抽出

文楽人形の動作を測定して得られた動作時系列 $x_k^{ij}(t)$ $(i=1,\cdots,I;\ j=1,\cdots,J;\ k=1,\cdots,K)$ を、(仮説 1)(仮説 2)(仮説 3) から表される式 (4.1) のように、機能的要素 $y^i(t)$、情緒的要素 $z^{ij}(t)$、確率的要素 $\eta_k^{ij}(t)$ に分解する計算方法を述べる。(仮説 4) で表される式 (4.2) より

$$x_k^{i1}(t) = y^i(t) + \eta_k^{i1}(t) \tag{4.4}$$

$$i=1,\cdots,I;\ k=1,2,\cdots,K$$

と表せるので、機能的要素 $y^i(t)$ を

$$y^i(t) = \frac{1}{K}\sum_{k=1}^{K} x_k^{i1}(t) \tag{4.5}$$

によって求める。

機能的要素 $y^i(t)$ が決まったので、式 (4.1) の未知数を左辺に移行すると、

$$z^{ij}(t) + \eta_k^{ij}(t) = x_k^{ij}(t) - y^i(t) \tag{4.6}$$

$$i=1,\cdots,I;\ j=2,\cdots,J;\ k=1,2,\cdots,K$$

と表せるので、未知数である情緒的要素 $z^{ij}(t)$ と確率的要素 $\eta_k^{ij}(t)$ を求める事を考える。それには、確率的要素

$$-\eta_k^{ij}(t) = y^i(t) + z^{ij}(t) - x_k^{ij}(t) \tag{4.7}$$

の $i=1,\cdots,I;\ k=1,2,\cdots,K$ についての 2 乗和

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left|\eta_k^{ij}(t)\right|^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left|y^i(t) + z^{ij}(t) - x_k^{ij}(t)\right|^2 \tag{4.8}$$

が最小になるように、情緒的要素 $z^{ij}(t)$ と確率的要素 $\eta_k^{ij}(t)$ を求める。

Figure 4.1と Figure 4.2は、k=1 回目の各動作の確率的要素 $\eta_1^{ij}(t)$ の R_1(頭) の 6 自由度を示している。確率的要素は残差であり、時々刻々とデータの値は変化する

が、はっきりとした dynamics を有せず、時間がたっても統計的な性質は変わらない、ランダムな信号であると予想されていたが、これらのグラフから、「好意」「敵意」の場合は比較的ランダムな信号といえても、「普通」の場合は波形に大きな変化が見られ、偶然的な誤差とも考えられるが、ただの残差であるとは言えそうではない。つまり、これら分離された確率的要素の中に、機能に依存しない情緒、機能の本質、純粋な情緒といったような、まだ何か他の要素が残っているのではないかと考えられる。そこで、この確率的要素に注目し、その統計的な性質をつかむことにする。

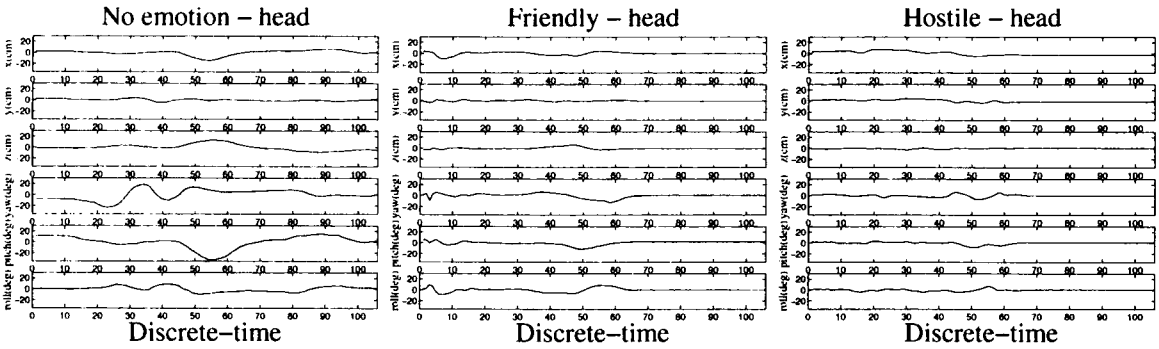


Figure 4.1: Stochastic factor of greeting.

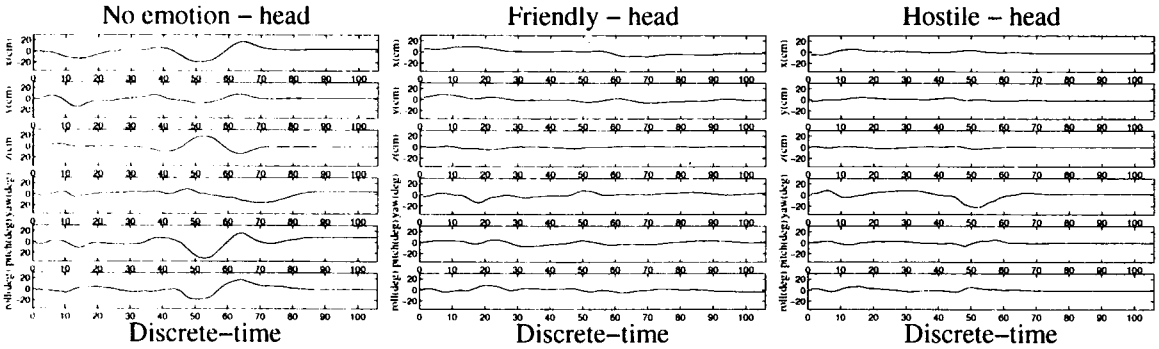


Figure 4.2: Stochastic of serving tea.

4.3.2 確率的要素の解析

本節では、前節で得られた確率的要素 $\eta_k^{ij}(t)$ について考える。この確率的要素は、本来動作の中に含まれる必要のない、同一の機能、情緒のもとで同じ動作を行なう度に違ってくる残差であるとしていた。しかし、前節で各要素へ分離された後、確率的要素を見ると明らかに何か傾向があり、ただの残差であるとは言えそうになかった。この理由として、前節の各要素への分離では情緒的要素は機能的要素に依存し

たものであると考えて分離を行っており、そのため確率的要素には、各回数のばらつきだけでなく、機能に依存しない情緒的な要素が含まれてしまったのではないかと考えた。また、この確率的要素は、はっきりとした dynamics を持たない定常な要素であると推測される、このことを実際のデータ解析に基づいて検証する事を試みる。

・スペクトル密度

まず、確率的要素 $\eta_k^{ij}(t)$ の各自由度 $\eta_{k(pq)}^{ij}(t)$ $p=1,\cdots,4;\ q=1,\cdots,6$ の統計的性質を調べるために、スペクトル密度 (パワースペクトル) を計算した。それらのうち、R_1(頭) の x 座標の確率的要素 $\eta_{k(11)}^{ij}(t)$ のパワースペクトルを Figure 4.3〜Figure 4.8 に示す。

Figure 4.3、Figure 4.4と Figure 4.5は、同一の機能、情緒のもとで同じ動作を N 回行なった時の確率的要素 $\eta_{k(pq)}^{ij}(t)$ $(p=1,\cdots,4;\ q=1,\cdots,6)$ のパワースペクトルを表し、比較している。機能 i と情緒 j が同一であれば、回数 k が変わっても、確率的要素 $\eta_{k(pq)}^{ij}(t)$ のスペクトル密度はよく似ていることが、グラフより見てとれる。従って、機能 i と情緒 j が同一であれば、回数 k が変わっても確率的要素は同一の確率過程 $\eta^{ij}(t)$ $(t=0,1,\cdots,T)$ の実現とみなせる。

Figure 4.6、Figure 4.7と Figure 4.8 は、同一の機能のもとで情緒を変えた時のパワースペクトルを、「感情なし」と「好意」、「感情なし」と「敵意」、「好意」と「敵意」の組み合わせで比較している。情緒 j=1、すなわち「感情なし」の時は、他の「好意」「敵意」に比べて値が大きいということがわかるが、機能が同一で、情緒が「好意」と「敵意」それぞれの確率的要素のパワースペクトルには顕著な変化を見出すことはできなかった。確率的要素は、情緒に依存せず、純粋にランダムな信号であるとも推察されるが、これについては、より詳しく統計的性質を調べるために ARMA 同定 [18] を用いる。

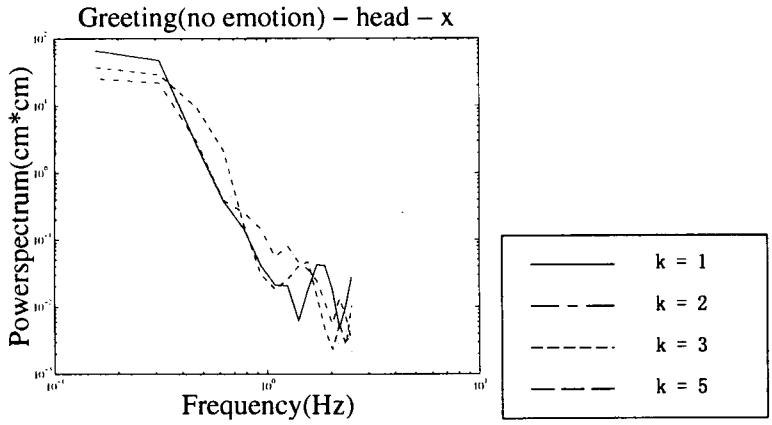


Figure 4.3: Power spectrum of stochastic factors(greeting(no emotion)).

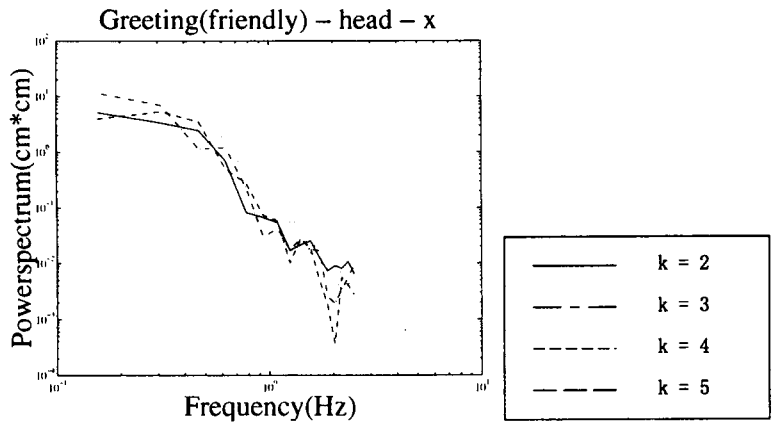


Figure 4.4: Power spectrum of stochastic factors(greeting(friendly)).

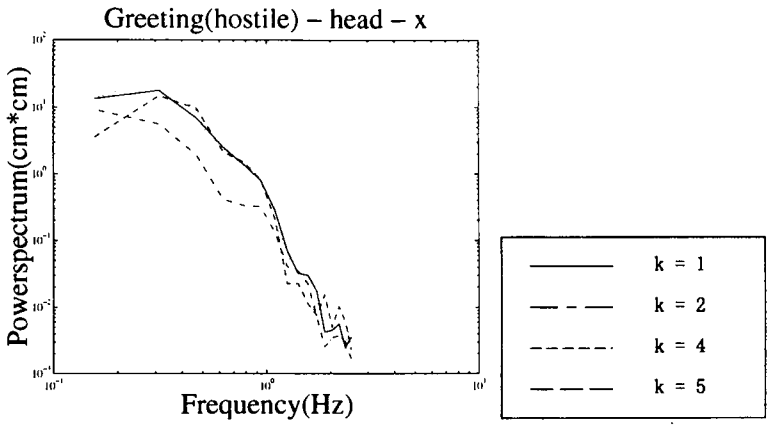


Figure 4.5: Power spectrum of stochastic factors(greeting(hostile)).

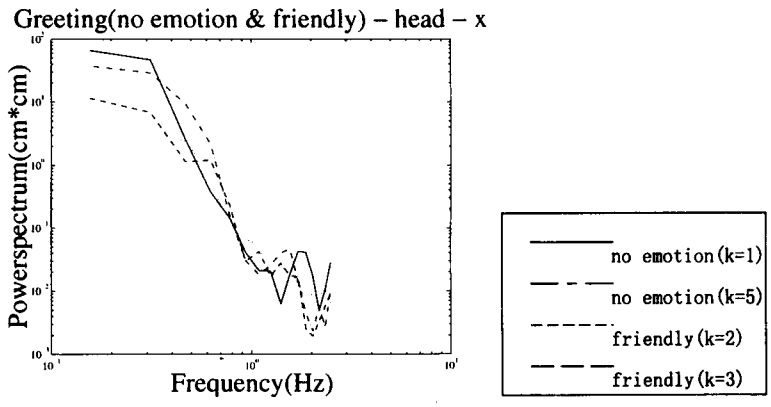


Figure 4.6: Power spectrum of stochastic factors(greeting(no emotion and friendly))).

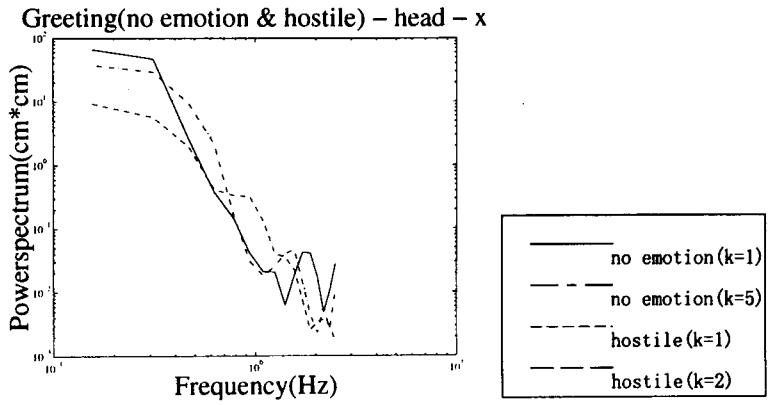


Figure 4.7: Power spectrum of stochastic factors(greeting(no emotion and hostile)).

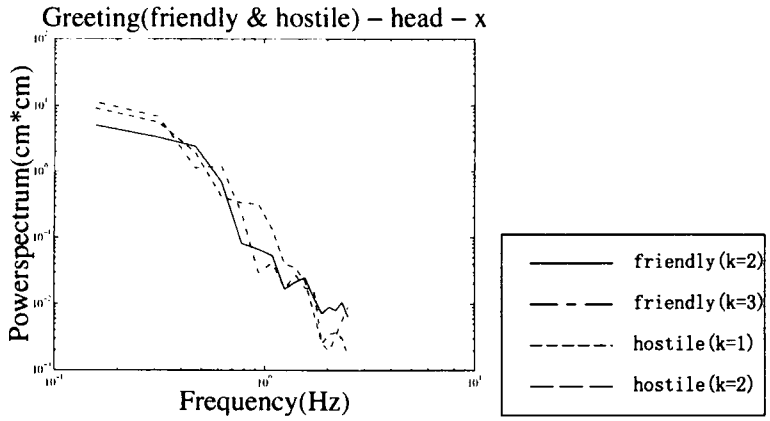


Figure 4.8: Power spectrum of stochastic factors(greeting(friendly and hostile)).

・ARMA(Auto Regressive Moving Average) 同定

先程の Figure 4.3～Figure 4.8を見ると、機能 i が同一で、情緒 j が変わった時の確率的要素のパワースペクトルにはっきりとした差異が見つけられなかった。そこで、より詳しく統計的性質を調べるために、ARMA モデルによるパラメータ同定を行い、これらのパラメータを比較する。

まず、ARMA モデルを利用するため、この確率的要素 $\eta^{ij}(t)$ が定常な確率過程であると仮定する。そして、得られた同定式から機能 i、情緒 j が変わると確率的要素の統計的性質がどのように変化するかを調べる。
ARMA モデル構造は、

$$\eta^{ij}(t) = -\sum_{m=1}^{p} a_m^{ij}\eta^{ij}(t-m) + \sum_{n=0}^{q} b_n^{ij} w(t-n) \tag{4.9}$$

$$a_m^{ij}, b_m^{ij} : \text{パラメータ}$$

$$w(t) : \text{ホワイトノイズ}$$

$$p, q : \text{次数}$$

で表す。ここで種々の次数をもったモデルの中から選択を行なう必要がある。選択の手法として最終予測誤差規範 (FPE) を用いて、最適の次数を決定する。FPE は次式で計算される。

$$FPE = \frac{1+n/T}{1-n/T} \times V \tag{4.10}$$

$$n = p+q+1 : \text{計算されるパラメータの総数}$$

$$T : \text{データ長 (時系列の長さ)}$$

V : 問題としている構造に対する損失関数 (2 次近似)

この結果、次数が 2 の時が最適であることがわかった。

$$\eta^{ij}(t) = -\sum_{m=1}^{2} a_m^{ij}\eta^{ij}(t-m) + \sum_{n=0}^{2} b_n^{ij} w(t-n) \tag{4.11}$$

同一の機能 i、情緒 j を持った動作 K 回分のパラメータを求め、その平均をとって、機能 i、情緒 j が変化すると a がどのように変化するかを調べる。以下に ARMA 同定により得られたパラメータの一部を示す。Table 4.1は R_1(頭) の x 座標を表している。

Table 4.1: Parameters a_m^{ij}, b_m^{ij}(head - x).

	a_1	a_2	b_0	b_1	b_2
$i=1, j=1$	-1.8290	0.8586	0.2125	0.5891	0.0035
$i=1, j=2$	-1.4239	0.5201	0.2414	0.6818	0.0486
$i=1, j=3$	-1.2734	0.3064	0.2164	0.6834	0.0309
$i=2, j=1$	-1.8791	0.9090	0.3267	0.6809	-0.0197
$i=2, j=2$	-1.7569	0.7812	0.1484	0.6383	0.0056
$i=2, j=3$	-1.4880	0.5361	0.1371	0.6121	-0.0085

4.3.3 結果と考察

この節では、文楽人形の動作時系列の仮説に基づき、最小 2 乗法により、機能的要素、情緒的要素、確率的要素への分解を行なった。分解された要素のうち本来、確率的要素は残差であり、時々刻々とデータの値は変化するが、はっきりした dynamics を有せず、時間がたっても統計的性質は変わらない、ランダムな信号であると考えていた。しかし、確率的要素のグラフ (Figure 4.1と Figure 4.2) を見ると、明らかに機能 i に関しては大きな値を有してあり、仮説では述べられなかった、何か他の要素が含まれているかと考え、確率的要素の統計的性質について詳しく検討した。まず、スペクトル密度 (パワースペクトル) を解析すると、同一の機能と情緒のもとでは、サンプル回数 k によらず、確率的要素の統計的性質が同一とみなせることが明らかになった。しかし、スペクトル密度からは、機能が同一で情緒が変化しても、違いを見出すことができなかった。よって、より詳しく統計的性質を調べるために ARMA 同定を用いたが、それでも違いを見出すことができなかった。しかしながら、スペクトル密度と ARMA 同定に共通する部分があった。スペクトル密度では、「感情なし」の値が「好意」「敵意」より大きく、ARMA 同定では、「感情なし」のパラメー

タは、「好意」「敵意」より値が大きいことが分かった。

つまり、「感情なし」の動作は、仮説より機能的要素、確率的要素から成り立っているが、正しく分解が行なわれていないと考えられる。「感情なし」の動作から機能的要素を求める時に、回数 K 個の動作時系列の平均をその動作の機能的要素として、その他の成分を仮説より確率的要素としていた。そして、「好意」「敵意」の動作は、確率的要素をできるだけ小さくなるように、求めている。だから、「感情なし」の確率的要素は「好意」「敵意」よりも値が大きくなり、それが、スペクトル密度と ARMA 同定に反映される。よって、機能的要素の求め方に問題があったと考えられる。そこで、次節でこの問題点を解決した、新たな動作時系列の分離方法について提案する。

4.4 KM_2O-ランジュバン方程式による機能・情緒の要素の抽出

4.4.1 外生入力を有する KM_2O-ランジュバン方程式による動作時系列の表現

文楽人形の動作時系列は、機能的要素、情緒的要素、確率的要素から成り立っていると考え、

$$x_k^{ij}(t) = y^i(t) + z^{ij}(t) + \eta_k^{ij}(t) \tag{4.12}$$

$$y^i(t) : \text{機能的要素}$$
$$z^{ij}(t) : \text{機能に依存した情緒的要素}$$
$$\eta_k^{ij}(t) : \text{確率的要素}$$
$$i = 1, \cdots, I;\ j = 1, \cdots, J;\ k = 1, 2, \cdots, K$$
$$I = 2,\ J = 3,\ K = 5$$
$$t = 0, 1, \cdots, T \quad T : \text{終端時刻}$$

とする。

式 (2.7) や式 (2.9) のように、人形の動作時系列 $x_k^{ij}(t)$ が、

$$x_{k(pq)}^{ij}(t)$$
$$(p = 1, \cdots, 4;\ \ q = 1, \cdots, 6)$$

を成分とする縦ベクトルであるので、$y^i(t)$, $z^{ij}(t)$, $\eta_k^{ij}(t)$ も同様の縦ベクトルで、$\eta_{k(pq)}^{ij}(t)$ などを成分とするものであるが、記号が煩雑であるので、(p, q) を省略して、$\eta_k^{ij}(t)$ などと書くことにする。従って以下、$x_k^{ij}(t)$ は 1 自由度の動作時系列を示すことになる。

確率的要素 $\eta_k^{ij}(t)$ のモデルとして、できるだけ広いクラスの確率過程を表現することのできる（初期時刻を　T_0　とする）KM_2O-ランジュバン方程式 (Kubo, Mori, Miyoshi,Okabe の Langevin 方程式) [14][15][16]

$$\eta_k^{ij}(T_0 + t) = -\sum_{s=0}^{t-1} \gamma_+^{ij}(t, s)\eta_k^{ij}(T_0 + s) + \nu_{+k}^{ij}(T_0 + t) \tag{4.13}$$

$$\gamma_+^{ij}(t, s) : \ KM_2O\text{-ランジュヴァンデータ}$$
$$\nu_{+k}^{ij}(t) : \ KM_2O\text{-ランジュヴァン揺動力}$$

を採用する。ここで T_0 は初期時刻であり、その値は場合によって $0,1,2,\cdots,T$ のうちの何らかの値である。$\gamma_+^{ij}(t,s)$ は時刻 t,s とともに変化する定数であり、初期時刻 T_0 に依存せず、時刻 t,s のみによって決まる [16] [18]。それは観測データ $x_k^{ij}(t)$ から推定されるべきものである。$\nu_{+k}^{ij}(T_0+t)$ は白色雑音であり、その確率分布は必ずしもガウス型でなくてもよく、その分散 $V_{+k}^{ij}(t)$ は時刻 t とともに変化してもよい。$\nu_{+k}^{ij}(t)$ の平均と共分散は以下のようになる。ただし E は期待値を表す。

$$E\left[\nu_{+k}^{ij}(T_0+t)\right]=0 \tag{4.14}$$

$$E\left[\nu_{+k}^{ij}(T_0+s)\nu_{+k}^{ij}(T_0+t)\right]=V^{ij}(t)\delta_{st} \tag{4.15}$$

ここで

$$\delta_{st}=\begin{cases}1 & \text{if } s=t\\ 0 & \text{if } s\neq t\end{cases} \tag{4.16}$$

である。$V_+^{ij}(t)$ も $\gamma_+^{ij}(t,s)$ と同様に、観測データ $x_k^{ij}(t)$ から推定されるべきものである。

機能の要素 $y^i(t)$ に対して、機能の入力 $u^i(t)$ を

$$u^i(T_0+t)=y^i(T_0+t)+\sum_{s=0}^{t-1}\gamma_+^{ij}(t,s)y_{(pq)}^i(T_0+s) \tag{4.17}$$

と定義し、情緒の要素 $z^{ij}(t)$ に対して、情緒の入力 $v^{ij}(t)$ を

$$v^{ij}(T_0+t)=z^{ij}(T_0+t)+\sum_{s=0}^{t-1}\gamma_+^{ij}(t,s)z^{ij}(T_0+s) \tag{4.18}$$

と定義する。これらの入力は未知であり、観測データから推定されるべきものである。これらを移項して、

$$y^i(T_0+t)=-\sum_{s=0}^{t-1}\gamma_+^{ij}(t,s)y^i(T_0+s)+u^i(T_0+t) \tag{4.19}$$

$$z^{ij}(T_0+t)=-\sum_{s=0}^{t-1}\gamma_+^{ij}(t,s)z^{ij}(T_0+s)+v^{ij}(T_0+t) \tag{4.20}$$

を得る。これらと 式 (4.13) の 3 個の式を辺々加え合わせて、式 (4.12) に注目すると、T_0 を初期時刻とし外生入力 $u^i(t)$, $v^{ij}(t)$ を有する KM_2O-ランジュバン方程式

$$\begin{aligned}&x_k^{ij}(T_0+t)=\\&-\sum_{s=0}^{t-1}\gamma_+^{ij}(t,s)x_k^{ij}(T_0+s)+u^i(T_0+t)+v^{ij}(T_0+t)+\nu_{+k}^{ij}(T_0+t)\end{aligned} \tag{4.21}$$

で動作時系列 $x_k^{ij}(t)$ が表される。

4.4.2 動作時系列の各要素の抽出

動作時系列の各要素の抽出のために、一つずつ未知のパラメータを同定していかなければならない。まず、文楽人形の動作時系列 $x_k^{ij}(t)$ $(t = 0, 1, \cdots, T)$ を表した式 (4.21) のパラメータ $\gamma_+^{ij}(t,s)$ $(0 \leq s < t \leq T)$ および共分散行列 $V_{+k}^{ij}(t)$ $(t = 0, 1, \cdots, T)$ を同定する場合に、式 (4.21) から外生入力 $u^i(t)$ や $v^{ij}(t)$ を消去した $x_{(diff)k}^{ij}(t)$ $(t = 0, 1, \cdots, T)$ を以下のように作り出す。

まず、$x_k^{ij}(t)$ の k についての平均をとり、$x_{(mean)}^{ij}(t)$ とする。

$$x_{(mean)}^{ij}(t) = \frac{1}{K}\sum_{k=1}^{K} x_k^{ij}(t) \tag{4.22}$$

これを式 (4.21) にあてはめると、

$$\begin{aligned} &x_{(mean)}^{ij}(T_0 + t) = \\ &-\sum_{s=0}^{t-1}\gamma_+^{ij}(t,s)x_{(mean)}^{ij}(T_0+s) + u^i(T_0+t) + v^{ij}(T_0+t) + \frac{1}{K}\sum_{k'=1}^{K}\nu_{+k'}^{ij}(T_0+t) \end{aligned} \tag{4.23}$$

$x_{(diff)k}^{ij}(t)$ を以下のように定めると

$$x_{(diff)k}^{ij}(t) = x_k^{ij}(t) - x_{(mean)}^{ij}(t) \tag{4.24}$$

この $x_{(diff)k}^{ij}(t)$ は次式で表される。

$$\begin{aligned} &x_{(diff)k}^{ij}(T_0 + t) = \\ &-\sum_{s=0}^{t-1}\gamma_+^{ij}(t,s)x_{(diff)k}^{ij}(T_0+s) + \nu_{+k}^{ij}(T_0+t) - \frac{1}{K}\sum_{k'=1}^{K}\nu_{+k'}^{ij}(T_0+t) \end{aligned} \tag{4.25}$$

$x_{(diff)k}^{ij}(t)$ は白色雑音だけによって駆動される確率過程であるので、式 (4.25) における時変定数 $\gamma_+^{ij}(t,s)$ と白色雑音 $\nu_{+k}^{ij}(T_0+t) - \frac{1}{K}\sum_{k'=1}^{K}\nu_{+k'}^{ij}(T_0+t)$ の分散 $V_{+(diff)}^{ij}(t)$ は、$x_{(diff)k}^{ij}(t)$ のサンプル自己相関関数

$$R_{(diff)k}^{ij}(\tau) = \frac{1}{1+T}\sum_{s=0}^{T-\tau} x_{(diff)k}^{ij}(s+\tau)x_{(diff)k}^{ij}(s) \tag{4.26}$$

から、[14] [15][16] で提案されている揺動散逸定理を用いて推定することができる。

ここで以下に述べる理由により、時変定数 $\gamma_+^{ij}(t,s)$ と白色雑音 $\nu_{+k}^{ij}(t) - \frac{1}{K}\sum_{k'=1}^{K}\nu_{+k'}^{ij}(t)$ の分散 $V_{+(diff)}^{ij}(t)$ の推定値は、

$$t = 0, 1, 2, \cdots, M \tag{4.27}$$

$$s = 0, 1, \cdots, t-1 \tag{4.28}$$

に対してしか求められない。但し

$$M \leq M_{max} \tag{4.29}$$

$$M_{max} = \left[3\sqrt{T_f + 1}\right] - 1 \tag{4.30}$$

とする。

[17] p.36 や [16] などで指摘されているように、式 (4.26) に従って計算するので、t が大きくなると自己相関関数を $R^{ij}_{(diff)k}(t)$ の値は不正確になる。[16] によると M_{max} より大きな t に対する $R^{ij}_{(diff)k}(t)$ の値は信頼できないので、[15][16] で指摘されているように、M_{max} 以下の M を適当に選んで、時間の長さ M の自己相関関数

$$R^{ij}_{(diff)k}(0), R^{ij}_{(diff)k}(1), \cdots, R^{ij}_{(diff)k}(M)$$

を揺動散逸定理に適用して得られる時変定数　$\gamma^{ij}_{+}(t,s)$ と分散 $V^{ij}_{+(diff)}(t)$ の推定値は、式 (4.27)(4.28)　なる t, s に対してだけ求められる。

このようにして白色雑音
$\nu^{ij}_{+k}(T_0+t) - \frac{1}{K}\sum_{k'=1}^{K}\nu^{ij}_{+k'}(T_0+t)$　の分散　$V^{ij}_{+(diff)}(t)$　が求まると、それを用いて、白色雑音　$\nu^{ij}_{+k}(T_0+t)$ の分散 $V^{ij}_{+}(t)$ を次の式から求めることができる。

$$\begin{aligned}
& V^{ij}_{+(diff)}(t) && (4.31)\\
= \; & Var\left[\nu^{ij}_{+k}(T_0+t) - \frac{1}{K}\sum_{k'=1}^{K}\nu^{ij}_{+k'}(T_0+t)\right] && (4.32)\\
= \; & Var\left[\left(1-\frac{1}{K}\right)\nu^{ij}_{+k}(T_0+t)\right] && (4.33)\\
& +Var\left[\frac{1}{K}\sum_{k'\neq k}\nu^{ij}_{+k'}(T_0+t)\right] \\
= \; & \left(1-\frac{1}{K}\right)^2 V^{ij}_{+}(T_0+t) + \frac{K-1}{K^2}V^{ij}_{+}(T_0+t) && (4.34)\\
= \; & \frac{K-1}{K}V^{ij}_{+}(t) && (4.35)
\end{aligned}$$

上記で求まった $\nu^{ij}_{+k}(t)$ の分散 $V^{ij}_{+}(t)$ と時変定数 $\gamma^{ij}_{+}(t,s)$ と観測された時系列のサンプル平均 $x^{ij}_{(mean)}(t)$ を

$$\begin{aligned}
& x^{ij}_{(mean)}(T_0+t) + \sum_{s=0}^{t-1}\gamma^{ij}_{+}(t,s)x^{ij}_{(mean)}(T_0+s) \\
= \; & u^{i}(T_0+t) + v^{ij}(T_0+t) + \frac{1}{K}\sum_{k'=1}^{K}\nu^{ij}_{+k'}(T_0+t) && (4.36)
\end{aligned}$$

for $i = 1, 2, \cdots, I$; $j = 1, 2, \cdots, J$
と考え合わせることによって、機能の入力 $u^i(T_0+t)$ と　情緒の入力 $v^{ij}(T_0+t)$ のマルコフ推定値を求めることができる　[17]。

推定された機能の入力 $u^i(T_0+t)$ と　情緒の入力 $v^{ij}(T_0+t)$ を、式 (4.19),(4.20) に代入することによって、機能の要素 $y^i(T_0+t)$ と　情緒の要素 $z^{ij}(T_0+t)$ を求めることができる。

自己相関関数の長さ M と初期時刻　T_0　を様々に変化させて上記の推定を行う事によって、T_0+t　が同一である機能の要素と情緒の要素の推定値を複数個もとめる事ができる。但し、一つの初期時刻　T_0　に対して推定できる　機能の要素 $y^i(T_0+t)$ と　情緒の要素 $z^{ij}(T_0+t)$ は、式 (4.27) で制限される t に対してだけである事に注意する。これら同一時刻の複数の推定値の平均を計算することにより、あらゆる時刻 t に対して、機能の要素 $y^i(t)$ と　情緒の要素 $z^{ij}(t)$ の、より精度の高い推定値が、得られる。

4.4.3　数値シミュレーションによる同定方法の検証

提案した外生入力の同定手法の有効性を示すために、数値シミュレーションによる検証を行なう。シミュレーションで用いる時系列には、式 (4.21) で表される外生入力を伴った KM_2O-ランジュバン方程式を ARX モデルの時系列に制限する。KM_2O-ランジュヴァンデータ $\gamma_+^{ij}(t,s)$ は過去の時刻の状態への依存の仕方が各時刻で変化している。しかし、シミュレーションを簡単化するために、これを変化しないとして、$\gamma_+^{ij}(t,s)$ が $(t-s)$ に依存し、またある一定以上過去の時刻の影響は受けない、すなわち $(t-s)$ が大きい時 $\gamma_+^{ij}(t,s) = 0$ とする。結局、以下の式を満たすような $a^{ij}(p)$ $(p = 1, 2, \cdots, N_a)$ が存在する。

$$a^{ij}(t-s) = \gamma_+^{ij}(t,s) \quad \text{for} \quad t > s \tag{4.37}$$

$$a^{ij}(p) = 0 \quad \text{for} \quad p > N_a \tag{4.38}$$

$$
\begin{cases}
-\sum_{p=1}^{2} a_p^1 x_k^{11}(t-p) + u^1(t) & & & & +c^{11}w_k(t) = x_k^{11}(t) \\
-\sum_{p=1}^{2} a_p^1 x_k^{12}(t-p) + u^1(t) & +v^{12}(t) & & & +c^{12}w_k(t) = x_k^{12}(t) \\
-\sum_{p=1}^{2} a_p^1 x_k^{13}(t-p) + u^1(t) & & +v^{13}(t) & & +c^{13}w_k(t) = x_k^{13}(t) \\
-\sum_{p=1}^{2} a_p^2 x_k^{21}(t-p) & +u^2(t) & & & +c^{21}w_k(t) = x_k^{21}(t) \\
-\sum_{p=1}^{2} a_p^2 x_k^{22}(t-p) & +u^2(t) & & +v^{22}(t) & +c^{22}w_k(t) = x_k^{22}(t) \\
-\sum_{p=1}^{2} a_p^2 x_k^{23}(t-p) & +u^2(t) & & & +v^{23}(t) \quad +c^{23}w_k(t) = x_k^{23}(t)
\end{cases}
\tag{4.39}
$$

$$i = 1, 2;\ j = 1, 2, 3;\ k = 1, 2, \cdots, 5;\ t = 0, 1, \cdots, T$$

$w_k(t)$: 標準白色雑音

式 (4.39) で生成した時系列に対して、提案された手法で外生入力を同定した結果の例を示す。

まず、時系列の最終時刻を $T = 300$ として、Table 4.2 のような外生入力とパラメータで生成した時系列を同定した結果を、Fig.4.9 から Fig.4.14 に示す。実線が真の値であり、点線が推定値である。

Table 4.2: Parameters of ARX model.

	i=1		i=2	
	j=2	j=3	j=2	j=3
a_1^i	$\frac{1}{2}$		$\frac{1}{4}$	
a_2^i	$\frac{1}{3}$		$\frac{1}{5}$	
$u^i(t)$	$2e^{0.005t}sin(\frac{10}{140}t)$		$3e^{0.005t}sin(\frac{10}{140}t)$	
$v^{ij}(t)$	$-2e^{0.005t}cos(\frac{10}{120}t)$	$-2e^{0.005t}cos(\frac{10}{100}t)$	$-3e^{0.005t}cos(\frac{10}{120}t)$	$-3e^{0.005t}cos(\frac{10}{100}t)$

	i=1			i=2		
	j=1	j=2	j=3	j=1	j=2	j=3
$x_k^{ij}(0)$	-5.83					
$x_k^{ij}(1)$	1.39					
c^{ij}	1.4×10^{-3}					

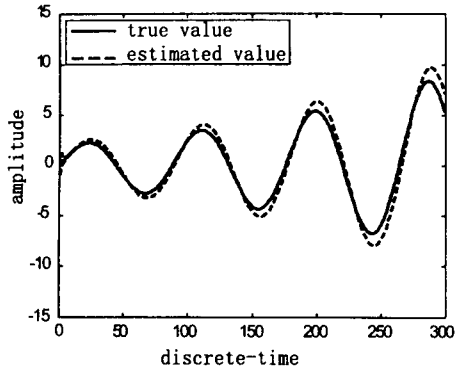


Figure 4.9: The estimate of $u^1(t)$.

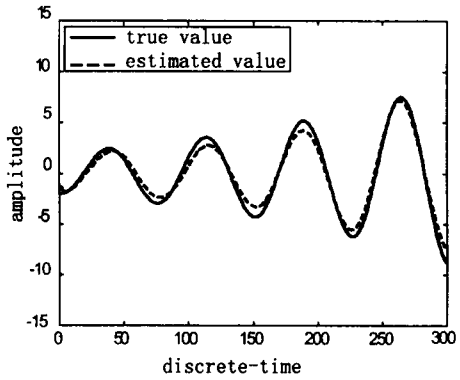


Figure 4.10: The estimate of $v^{12}(t)$.

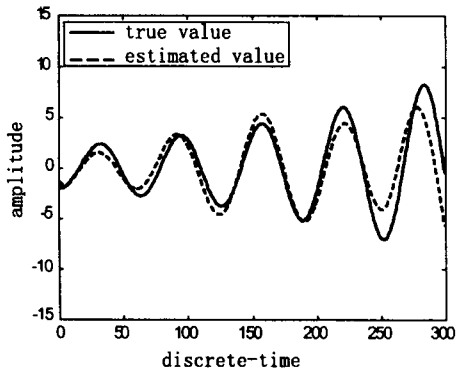


Figure 4.11: The estimate of $v^{13}(t)$.

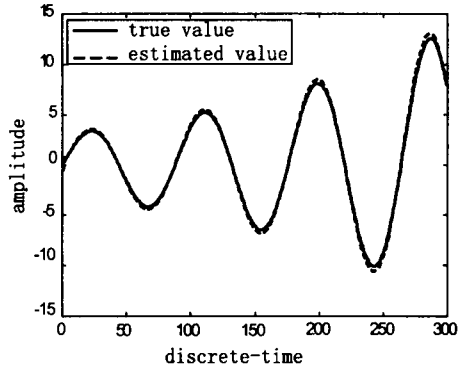


Figure 4.12: The estimate of $u^2(t)$.

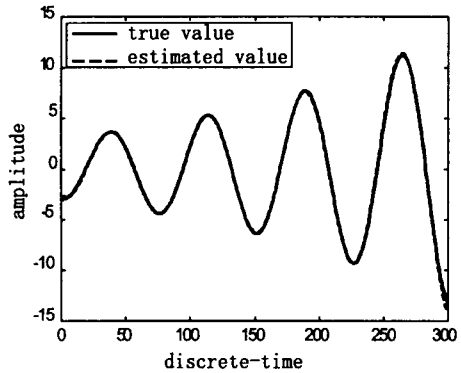


Figure 4.13: The estimate of $v^{22}(t)$.

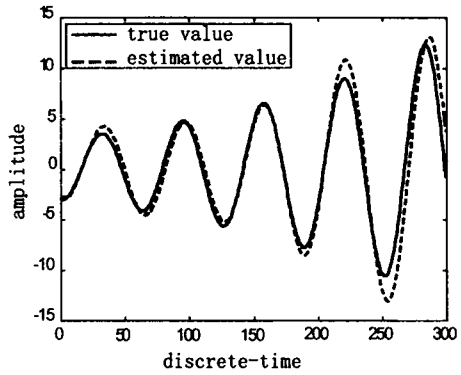


Figure 4.14: The estimate of $v^{23}(t)$.

次に 時系列の最終時刻を $T = 1000$ として Table 4.3 のような外生入力とパラメータで生成した時系列を同定した結果を、Fig. 4.15 から Fig. 4.20 に示す。点線の推定値が実線の真の値と、ほとんど重なっている様子が見られる。

Table 4.3: Parameters of ARX model.

	i=1		i=2	
	j=2	j=3	j=2	j=3
a_1^i	$\frac{1}{2}$		$\frac{1}{4}$	
a_2^i	$\frac{1}{3}$		$\frac{1}{5}$	
$u^i(t)$	$5e^{0.005t}sin(\frac{10}{220}t)$		$4e^{0.005t}sin(\frac{10}{220}t)$	
$v^{ij}(t)$	$-5e^{0.005t}cos(\frac{10}{180}t)$	$-5e^{0.005t}cos(\frac{10}{200}t)$	$-4e^{0.005t}cos(\frac{10}{180}t)$	$-4e^{0.005t}cos(\frac{10}{200}t)$

	i=1			i=2		
	j=1	j=2	j=3	j=1	j=2	j=3
$x_k^{ij}(0)$	10					
$x_k^{ij}(1)$	5					
c^{ij}	$1.0 \times 10^{-0.3}$					

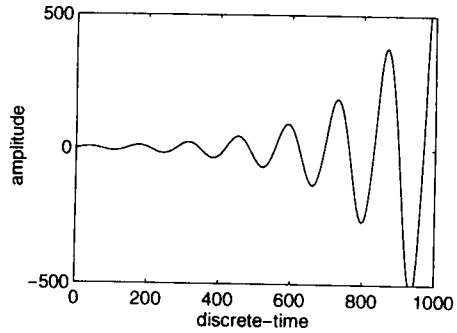


Figure 4.15: The estimate of $u^1(t)$.

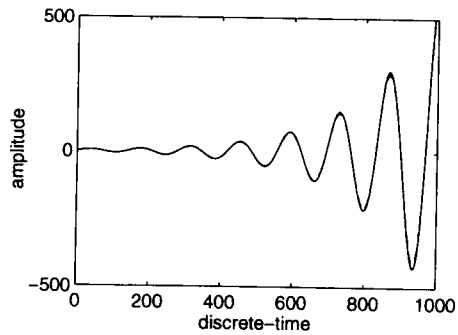


Figure 4.18: The estimate of $u^2(t)$.

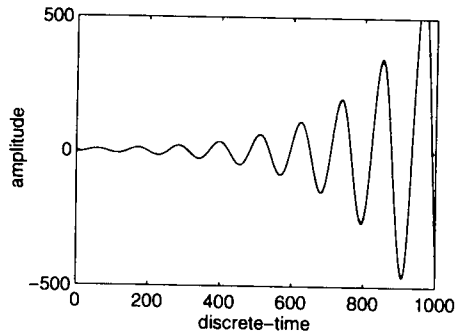


Figure 4.16: The estimate of $v^{12}(t)$.

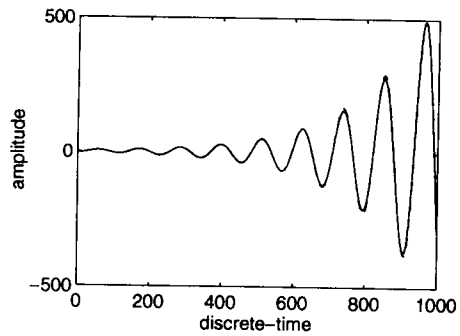


Figure 4.19: The estimate of $v^{22}(t)$.

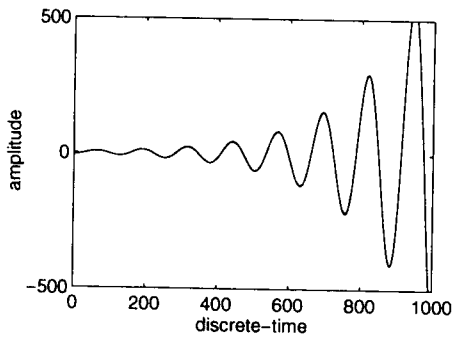


Figure 4.17: The estimate of $v^{13}(t)$.

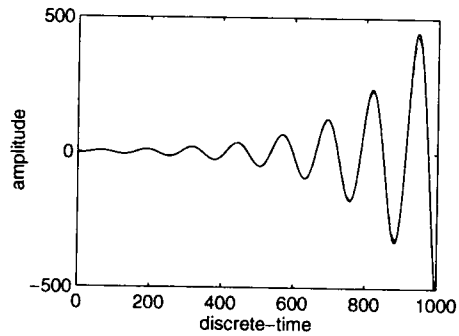


Figure 4.20: The estimate of $v^{23}(t)$.

自己相関関数の長さ M を変えることで推定結果が異なることが分かった。そこで、真の値

$$P = \frac{\sum_{t=0}^{T}\{u_{estim}(t) - u_{true}(t)\}^2}{\sum_{t=0}^{T}\{u_{true}(t)\}^2} \tag{4.40}$$

の評価を行った。外生入力 $u^i(t)$ と $v^{ij}(t)$ のそれぞれの推定値の相対誤差 P を求め、その平均を求めた。T=300 と T=1000 の場合を、Fig.4.21 に示す。ただし、ARX モデルの次数を 2 にしているので、自己相関関数の長さ M は 3 以上をとる。

ここで、時系列の長さ $T = 300$ のとき、P の単位は [%] であるが、$T = 1000$ の時の単位が [0.1 × %] である事に注意する。

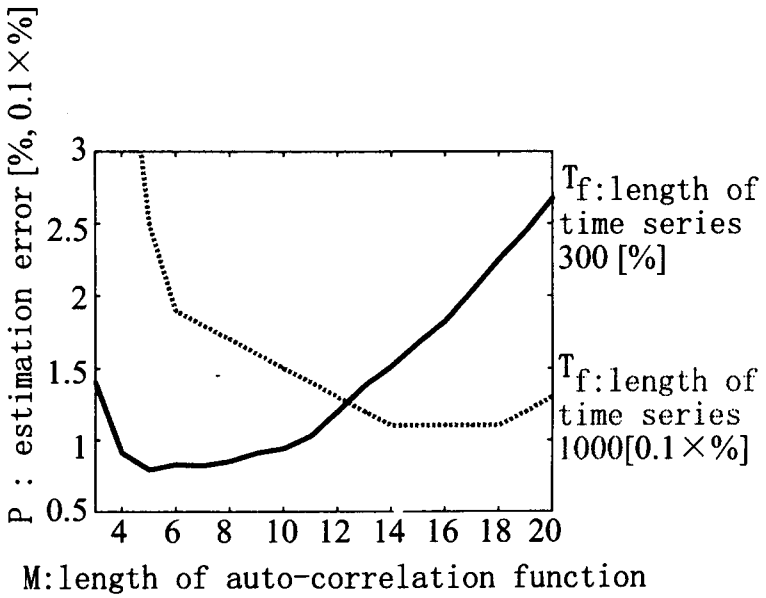


Figure 4.21: The errors of the estimated inputs.

Table 4.4: Relative errors P[%].

	P								
M	3	4	5	6	7	8	9	10	11
T=300	1.40	0.91	0.79	0.83	0.82	0.85	0.91	0.94	1.03
T=1000	1.00	0.43	0.25	0.19	0.18	0.17	0.16	0.15	0.14

	P								
M	12	13	14	15	16	17	18	19	20
T=300	1.20	1.38	1.51	1.68	1.83	2.04	2.25	2.45	2.68
T=1000	0.13	0.12	0.11	0.11	0.11	0.11	0.11	0.12	0.13

Table 4.4より、データ数Tが多いほど自己相関関数に信頼性が増し、推定精度が上がることがわかった。T=300の時では、信頼値Mは最小値3から最大値50までの範囲をとるが、その範囲内の全てのMから求まる、時系列のパーセント表示Pを求めなければ、最適な信頼値Mを見出せない。また、Tが異なると最適な信頼値Mが変化してしまうこともわかった。

4.4.4 動作時系列の各要素の抽出

ここでは、文楽人形の動作時系列から、外生入力$u^i(t)$と$v^{ij}(t)$を先程の同定手法を用いて推定し、外生入力の動作時系列を用いて、動作時系列の機能的要素と情緒的要素を求める。シミュレーションの結果、外生入力の同定手法の有効性を示すことができたが、測定実験で得られた文楽人形の動作時系列に外生入力の同定手法を当てはめると、どのようになるかを考える。先程のシミュレーションでは、外生入力の時系列を決めていたので、評価を行なえたが、文楽人形の動作時系列の外生入力は未知なので、評価を行なうことができない。つまり、最適な信頼値Mを求められない。そこで信頼値Mを変えて、その都度求まる外生入力から、機能的要素と情緒的要素を求め、それらの要素を組み合わせて動作を生成した動作時系列と、文楽人形の動作時系列を比較した相対誤差Pの評価を行なって、最適な信頼値Mを求める。以下に機能的要素と情緒的要素の抽出方法を述べる。

文楽人形の動作時系列は機能的要素、情緒的要素、確率的要素から成り立っていることはすでに述べている。式(4.12)をt=0とすると、

$$x_k^{ij}(0) = y^i(0) + z^{ij}(0) + \eta_k^{ij}(0) \tag{4.41}$$

と表すことができ、機能的要素$y^i(t)$と情緒的要素$z^{ij}(t)$の初期値$y^i(0)$と$z^{ij}(0)$を外生入力$u^i(t)$と$v^{ij}(t)$を求める時と同じ方法を用いて求める。$\eta_k^{ij}(0)$の分散は、式

(4.35) から求まる $V_{+k}^{ij}(t)$ の初期値 $V_{+k}^{ij}(0)$ を用いる。

ある時刻 T_0 を初期時刻とすると、$y^i(T_0+t)$ と $z^{ij}(T_0+t)$ は式 (4.19) と式 (4.20) から

$$y^i(T_0+t) = -\sum_{s=0}^{t-1} \gamma_+^{ij}(t,s) y^i(T_0+s) + u^i(T_0+t) \tag{4.42}$$

$$z^{ij}(T_0+t) = -\sum_{s=0}^{t-1} \gamma_+^{ij}(t,s) z^{ij}(T_0+s) + v^{ij}(T_0+t) \tag{4.43}$$

$$M : \text{相関関数の長さ}$$
$$T_0 = 0, 1, \cdots, T-M-1$$
$$t = 0, 1, \cdots, M$$

と表し、$y^i(T_0+t)$ と $z^{ij}(T_0+t)$ を求め、同一時刻のそれぞれの要素の平均を計算して機能的要素 $y^i(t)$ と情緒的要素 $z^{ij}(t)$ を求めた。
推定した結果、相関関数の長さ M は 2 が最適であった。M=2 で推定された機能的要素 $y^i(t)$ と情緒的要素 $z^{ij}(t)$ から、6 種類の動作を生成して、式 (4.40) のような推定値の真の値に対する相対誤差 P^{ij} $(i = 1, \cdots, I;\ j = 1, \cdots, J)$ を求めた。各動作での回数 k の動作時系列の平均を基準として、生成された動作時系列との評価を行なう。24(自由度)×6(動作)=144 自由度の動作時系列全ての P^{ij} の平均は 2.32%となった。

最後に、この手法では、各動作の同じ自由度、つまり R_1(頭) の x 座標なら各動作の R_1(頭) の x 座標を同時に推定している。文楽人形の動作は 24 自由度を合わせ持って、1 つの動作を生成している。しかし、今回は各自由度の関連性を考えず、1 自由度ずつ推定を行なっているので、推定された各自由度の動作時系列から、Computer Graphics(CG) で文楽人形の動作を生成してみると、違和感なく動いているのがわかった。

Table 4.5: Relative errors P^{ij}[%].

	P^{11}	P^{12}	P^{13}	P^{21}	P^{22}	P^{23}
head - x	0.05	0.22	5.32	0.01	2.77	0.95
head - y	1.09	2.16	0.57	0.10	0.67	0.44
head - z	0.05	1.06	49.91	0.31	0.36	0.11
head - yaw	0.64	0.05	0.20	6.53	0.72	0.05
head - pitch	0.02	0.19	0.11	3.99	0.22	0.25
head - roll	1.56	0.91	0.73	1.33	0.35	0.38
shoulder - x	0.31	0.12	2.18	0.04	2.77	0.76
shoulder - y	0.46	0.70	0.31	0.08	1.49	0.37
shoulder - z	0.09	2.57	31.70	0.25	0.48	0.12
shoulder - yaw	8.94	0.17	0.61	1.25	0.72	0.33
shoulder - pitch	0.62	1.74	9.92	2.34	1.75	0.20
shoulder - roll	0.12	1.94	6.46	0.48	0.31	1.19
arm - x	0.43	0.14	1.56	0.03	2.86	0.08
arm - y	0.08	0.27	0.65	0.21	0.99	0.09
arm - z	0.07	1.21	15.91	0.89	0.04	0.32
arm - yaw	0.19	3.89	0.70	2.52	1.01	0.21
arm - pitch	0.58	0.47	7.27	0.23	1.01	1.08
arm - roll	2.05	1.17	8.12	0.13	0.88	0.41
hand - x	0.49	0.14	1.70	0.04	3.16	0.07
hand - y	0.07	0.49	0.53	0.20	0.85	0.11
hand - z	0.09	1.75	15.47	0.70	0.08	0.20
hand - yaw	0.33	1.65	1.03	1.99	5.31	17.52
hand - pitch	2.50	1.90	1.34	4.89	9.91	14.16
hand - roll	2.36	1.42	14.43	0.10	0.40	0.50

4.4.5 結果と考察

文楽人形の動作時系列は、外生入力を持った KM_2O-ランジュバン方程式で表せることを示した。その動作時系列に含まれた全てのパラメータと外生入力を求め、それらを用いて機能的要素と情緒的要素を推定した。

まず、外生入力の同定手法の有効性をシミュレーションによって示し、有効な自己相関関数の長さを定め、データ数が多ければ推定精度が増し、また、データ数(時系列の長さ)によって自己相関関数の最適な長さが存在することがわかった。

そこで、文楽人形の動作時系列における外生入力が未知であることから、自己相関関数の長さを変えることによって、その都度求まる機能的要素と情緒的要素から動作を生成させ、測定で得られた動作時系列と比較した評価を行なうことで、自己相関関数の最適な長さを決定した。

生成された6種類の動作の推定精度の評価を行なった結果、文楽人形の動作時系列が機能的要素と情緒的要素に分配されており、またCGにより文楽人形の動作を生成した結果、人形の動作として判断できることから、機能・情緒の外生入力を同定するアルゴリズムの有効性を示した。この測定された動作時系列から同定される機能・情緒の外生入力から、機能・情緒の要素を求めることができる。

4.5 機能的要素と情緒的要素による情緒的動作の生成

文楽人形の動作時系列は仮説より、機能的要素、情緒的要素、確率的要素から成り立ち、確率的要素は同一の機能、情緒のもとで数回同じ動作を繰り返す度に変わる残差であるということは既に述べた。情緒的要素とは、機能に依存した情緒的要素を表す。よって、動作とは機能的要素、情緒的要素のみで生成することができ、式 (4.12) から

$$x^{ij}(t) = y^i(t) + z^{ij}(t) \tag{4.44}$$

$y^i(t)$:機能的要素
$z^{ij}(t)$:情緒的要素

と表せる。

式 (4.44) から、「感情なし」の動作を生成しようとすれば、機能的要素 $y^i(t)$ のみを用いて、「感情なし」の動作を生成し、機能的要素 $y^i(t)$ に情緒的要素 $z^{ij}(t)$ を足すことで、「好意」「敵意」の情緒を伴った動作を生成することができる。「好意」「敵意」の感情をより大きく表に出して動作を行ないたい場合や、また、「好意」「敵意」はあるのだが、それほど動作の中に反映されない動作を生成するにはどのようにすればよいか。そこで、情緒的要素である $z^{ij}(t)$ の性質を強めたり、また弱めたりすれば「好意」「敵意」の感情をコントロールして、情緒的動作の生成を行なえると考えた。式 (4.44) に感情の重み a を設けるとすると、

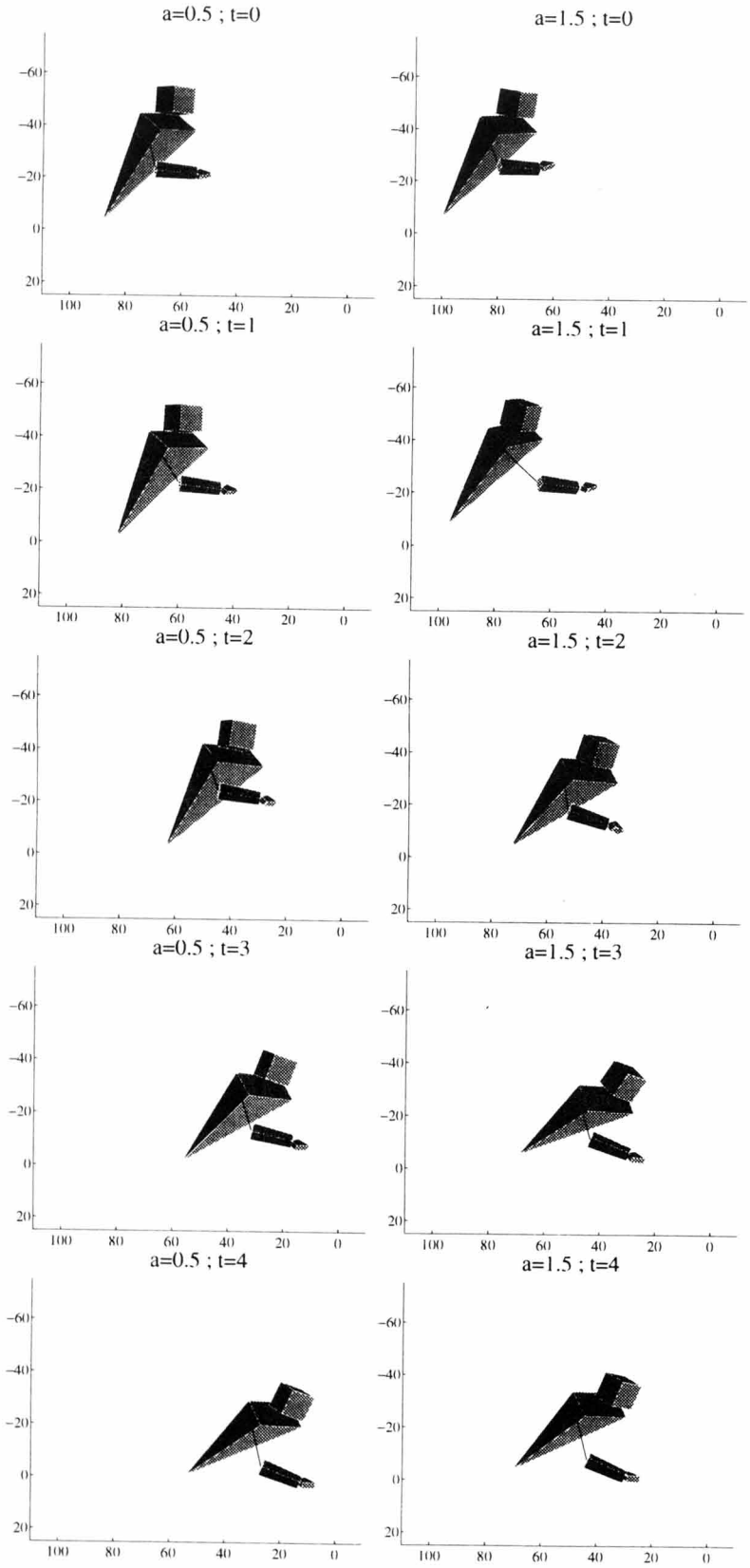
$$x^{ij}(t) = y^i(t) + az^{ij}(t) \tag{4.45}$$

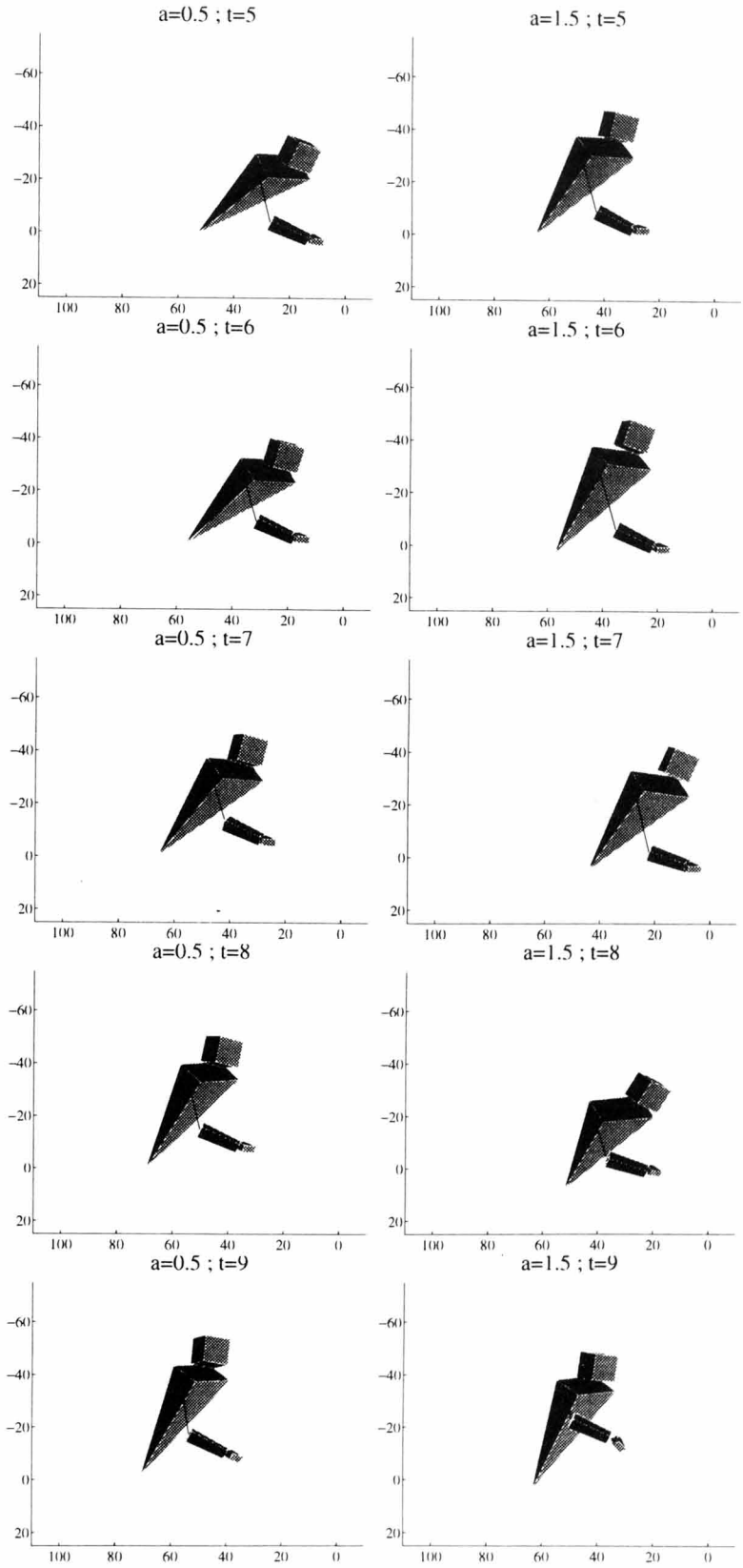
と表せる。ここで重み a を 1 より大きくすれば、j という情緒を強調した機能 i の動作を設計できる。逆に $0 \leq a < 1$ とすれば、情緒を減衰させた動作を設計することができる。ここで、文楽人形の動作生成に用いる動作時系列は、外生入力を介して推定された機能的、情緒的要素の時系列を用いる。

好意を持って ($j = 2$) お茶を差し出す ($i = 2$) という動作に対して、Fig. 4.22 に、$a = 0.5$ として情緒を減衰させた動作 (左側) と、$a = 1.5$ として情緒を強調させた動作 (右側) の例を示す。「好意を持ってお茶を渡す」動作について、サンプリング時間を $1[s]$ とした時の、時刻 t の文楽人形の状態を表している。動作の変移がわかりやすいように両軸の単位を表示させていないが、両軸とも cm を表している。

「お茶を渡す」とは、対象相手に向かって、お茶を持った手を前に差し出す機能である。相手への「好意」の感情が大きくなる、つまり a が大きくなると、

1. 初期時刻 (t=0) における文楽人形の位置から対象相手にお茶を差し出すまでの移動距離が長くなる。
2. 人形の上下運動が大きくなる。
3. お茶を差し出した後も、文楽人形の姿勢が低い。


a=0.5 ; t=0
a=1.5 ; t=0
a=0.5 ; t=1
a=1.5 ; t=1
a=0.5 ; t=2
a=1.5 ; t=2
a=0.5 ; t=3
a=1.5 ; t=3
a=0.5 ; t=4
a=1.5 ; t=4
-60
-40
-20
0
20
100
80
60
40
20
0


a=0.5 ; t=5
a=1.5 ; t=5
a=0.5 ; t=6
a=1.5 ; t=6
a=0.5 ; t=7
a=1.5 ; t=7
a=0.5 ; t=8
a=1.5 ; t=8
a=0.5 ; t=9
a=1.5 ; t=9
-60
-40
-20
0
20
100
80
60
40
20
0

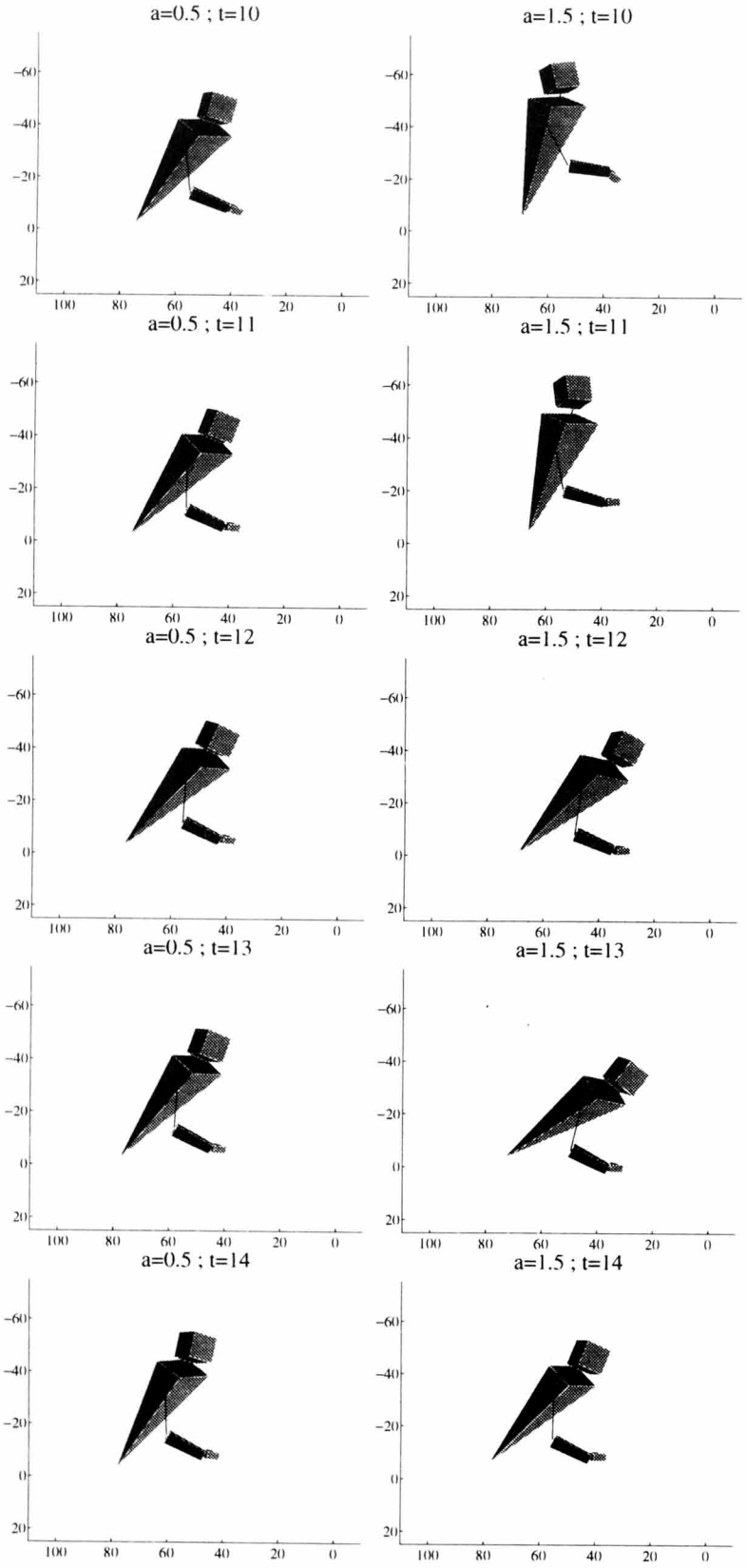


Figure 4.22: Motion(serving tea - friendly).

4.6 結言

本章では、機能・情緒・確率的揺らぎの線形構造に基づいて、文楽人形の演技動作を解析し、人間型のアミューズメント・ロボットの動作生成に使用する動作時系列の情緒を、強調させたり減衰させたりして加工するシステマティックな手法を提案した。様々な機能と様々な情緒の文楽人形の動作動作時系列からから、機能の要素と情緒の要素を推定・抽出し、情緒の要素を強調させたり減衰させたりして、新しい動作を計算することができる。

動作時系列から機能・情緒の要素を推定するために、次のような時系列解析の手法を確立し、その有効性を数値シミュレーションで確認した。

- 様々な要素で構成される複数の時系列を、それらの要素に対応する複数の入力を持つ KM_2O-ランジュバン方程式で表現する事を提案した。
- KM_2O-ランジュバン方程式の複数の入力を推定する方法を考案した。全ての時刻に対して入力を推定するにあたっては、自己相関関数の長さを適切に選んで KM_2O-ランジュバン方程式の初期時刻を変化させながら推定を行う事が、鍵となるアイディアである。
- 推定された複数の入力から、時系列を構成する複数の要素を計算することができる。

第 5 章　文楽人形の動作軸の位相・振幅の解析

5.1　緒言

文楽人形の動作は、頭・肩・腕・手などの各部位の動きから構成される多自由度の時系列である。そのような動作を解析したり設計したりするにあたって、各部位の動きをばらばらに考察していては、見通しが悪く、人形一体としての動作を扱うことはできない。

そこで本章では、動作の主要な部分を表す物を用いて解析を行う。この動作の主要な部分、すなわち動作の本質的な部分を表す物として動作軸を提案する。文楽人形の動きのうち、動作軸の動きを動作の主要な部分であると考え、動作軸に対する人形の各部位の相対的な動きを、装飾的・副次的なものであると考える。従って、文楽人形の演技動作の情緒表現のなかでも、特にこの動作軸に表れている部分に注目する。

第 2 章で測定された様々な機能、情緒の動作の動作軸を計算し、情緒表現を伴った動作 (情緒動作) と情緒表現を伴わない動作 (機能動作) を比較した。その結果、動作に情緒表現が加わることにより動作の位相 (動作時間の長さ) と振幅 (動作の大きさ) が変化していると判断できた。そこで、情緒動作からこれらの変化を取り除き、情緒表現における位相要素、振幅要素を抽出する手法を提案する。

まず、情緒表現における位相要素を抽出するために、位相の変化は、動作の最小単位である事象の実現時刻の変化であるという仮説を提案する。情緒動作と機能動作間の対応する事象を動作時系列のグラフの局所的な形を表すウェーブレット係数によって決定し、事象の実現時刻を調整することで情緒表現における位相要素を抽出する。

次に、情緒表現における振幅要素を抽出するために、機能動作と情緒動作を、共

に対応する幾つかのフレーズ (phrase) に分割する。対応するフレーズの振幅の違いを、機能動作と情緒動作との間で比較して、その違いを説明するゲインを導入する。これらの各フレーズ毎のゲイン達によって、情緒表現における振幅要素を説明する。

最後に、動作軸に注目した動作生成手法を提案し、抽出した情緒表現における位相要素・振幅要素を機能動作の動作軸に適用することで、実際に情緒表現を伴った動作を設計・生成することを試みる。

5.2 人形の動きを代表する動作軸

5.2.1 動作軸の提案

文楽人形には頭、肩、右腕、右手と4つの部位があり、それぞれの部位は3自由度の位置と3自由度の姿勢で表現される。従って、測定された文楽人形の動作時系列は, 一つの動作につき24の自由度を持つ. これら24のそれぞれの自由度に対し解析を行なうことは非常に見通しが悪い. また、人間型ロボットの動作を設計するという立場からも、頭・肩・腕・手などの身体の各部位の位置と姿勢からなる多自由度を同時に考慮して動作を設計することは、見通しが悪く困難である。

そこで, これらすべての自由度の情報, すなわち動作全体の情報のうち, できるだけ本質的な部分を残して細かい部分を捨て去った, より簡単な要素を用いて解析を行なうことが望ましい. 動作から、主要な部分を抽出して、それで表現しきれない装飾的な部分から区別する事を考える。そして動作の主要な部分から重点的に解析する方針をとる。つまり多自由度の動作を、低自由度の主要な部分と、それでは表現しきれない副次的な部分に分解し、動作の主要な部分を優先的に解析し、その設計手法を整える事を考える。

動作の主要部分を何と考えるかが問題であるが、文楽人形の操作者 (人形遣い) は演技時に人形の「軸」を意識しながら操作をしているという知見に基づいて、人形の軸の動きを動作の主要部分と考える。この動作の軸は動作を代表して表すものであり, 動作の本質的な情報を含んだものであると考えられる.

そこで, このような動作軸を工学的に計算して、

$$(\text{人形の動き}) = (\text{軸の動き}) + (\text{相対運動}) \tag{5.1}$$

という形で、人形の動作を、(1) 動作軸の動きと (2) 軸に対する相対運動 の2つに分解して、動作の主要な部分である (1) 軸の動きが情緒にどのように依存しているかを本章で解析し、その解析結果に基づいた動作の設計方法を考察する。

動作の「軸」について考察した結果、体の各部分はこの「軸」を中心として動かされ、「軸」は各時刻における人形の姿勢を代表してあらわしていると判断された。工学的な見地からするとこれは、

- 「軸」に対する相対運動が最小になる
- 各時刻において人形の体を平均的に貫く

ということであると考えられる。このような条件をみたすような「軸」として、各時刻において人形を一つの剛体と見たときの慣性主軸を考えた。この時刻毎に変化する慣性主軸を動作の動作軸と呼ぶことにする。

動作の軸であると考えられた、人形の慣性主軸を Figure 5.1 に示す。人形の重心の位置（x 座標・y 座標・z 座標）と慣性主軸の z 軸（人形から見て上下の軸）を表す方向ベクトル（2 自由度）および慣性主軸の x 軸（人形の前後の軸）が z 軸を軸として回転した角度（1 自由度）の 6 自由度の時間変化で、動作軸の動きを表すことができる。人形の各部位の動きの総体が２４自由度と比較すると、動作軸の動きの 6 自由度は、かなり低次元化されている。

Figure 5.1: Example of motion axes.

人形の慣性主軸を計算する方法を述べる。まず、文楽人形の各部分の体積を測定し、密度を一様と仮定して質量 M_p を計算する。次に各部分を K_p 個の質点系と考え、その位置と姿勢に基づいて各時刻における質点の座標 $(x_{pk}(t), y_{pk}(t), z_{pk}(t))$ を

計算し、慣性テンソル $I(t)$ を計算する。この慣性テンソルを対角化することで慣性主軸が得られる。添字 k は質点番号、添字 p は Table.5.1 に示される体の各部分であり、K_p は各部分の質点数、M_p は各部分の質量である。

$$I(t) = \begin{bmatrix} Ixx & -Ixy & -Ixz \\ -Ixy & Iyy & -Iyz \\ -Ixz & -Iyz & Izz \end{bmatrix} \tag{5.2}$$

$$Ixx = \sum_{p=1}^{6} M_p \sum_{k=1}^{K_p} (y_{pk}(t)^2 + z_{pk}(t)^2) \tag{5.3}$$

$$Ixy = \sum_{p=1}^{6} M_p \sum_{k=1}^{K_p} x_{pk}(t) y_{pk}(t) \tag{5.4}$$

$$Iyy = \sum_{p=1}^{6} M_p \sum_{k=1}^{K_p} (x_{pk}(t)^2 + z_{pk}(t)^2) \tag{5.5}$$

$$Ixz = \sum_{p=1}^{6} M_p \sum_{k=1}^{K_p} x_{pk}(t) z_{pk}(t) \tag{5.6}$$

$$Izz = \sum_{p=1}^{6} M_p \sum_{k=1}^{K_p} (x_{pk}(t)^2 + y_{pk}(t)^2) \tag{5.7}$$

$$Iyz = \sum_{p=1}^{6} M_p \sum_{k=1}^{K_p} y_{pk}(t) z_{pk}(t) \tag{5.8}$$

Table 5.1: Subscript p.

$p=1$	$p=2$	$p=3$	$p=4$	$p=5$	$p=6$
head	shoulder	right arm	right hand	left arm	left hand

この動作軸の計算において、動作時系列に左腕、左手が含まれていないことが問題となった。本来、文楽人形には左腕、左手も存在しているため、動作軸の計算にはこれらの動作時系列も必要であり、測定された 4 つの部分のみでは、正しい動作軸を得ることはできない。そこで、左腕、左手の動作が右腕、右手に対して追従し、ほぼ対象な動きをしていることをもとに、仮想的に左腕、左手を作成した。将来的には左腕、左手を測定して動作軸を計算することが望まれる。

動作軸は各時刻において一つの座標系であるので、自由度は6となる。時刻tにおける動作軸は

$$X^{ij}(t) = \left(X_q^{ij}(t)\ ;\ q \downarrow 1, \cdots, 6\right) \tag{5.9}$$

$$= \left(X_1^{ij}(t), \cdots, X_6^{ij}(t)\right)^T \tag{5.10}$$

であり、動作軸を表す時系列は

$$X^{ij}(t) \quad (0, 1 \cdots, T) \tag{5.11}$$

となる。

ここで、添字i, jは第2章で示した機能、情緒を表す添字であり、qは、Table5.2で表される動作軸の自由度である。$q = 1, 2, 3$は各時刻における人形の重心、すなわち動作軸の原点のx, y, z座標を表す。また、動作の特徴をより明確に捉えるため、座標系の姿勢表現として、等価回転角-回転軸表現を用いた。回転軸は、慣性テンソルを対角化して得られる固有値、すなわちその軸回りの慣性モーメントが最小となる軸とした。

すると、人形の重心の位置 x 座標 ($q = 1$)・y 座標 ($q = 2$)・z 座標 ($q = 3$) の3自由度と慣性主軸の z 軸（人形から見て上下の軸）を表す方向ベクトル（2自由度 $q = 4, q = 5$）および慣性主軸の x 軸（人形の前後の軸）が z 軸を中心として回転した角度（1自由度 $q = 6$）の6自由度の時間変化で、動作軸の動きを表すことができる。人形の各部位の動きの総体が2 4自由度と比較すると、動作軸の動きの6自由度は、かなり低次元化されている。

Table 5.2: subscript q.

$q = 1$	$q = 2$	$q = 3$	$q = 4$	$q = 5$	$q = 6$
$x-$ $coordinate$	$y-$ $coordinate$	$z-$ $coordinate$	$equivalent-$ $axis - x$	$equivalent-$ $axis - y$	$equivalent-$ $angle$

本章では、情緒表現をより見通しよく捉えるために、動作を動作軸とこれに対する各部分の相対動作にわけて考える。動作に含まれる情緒表現のなかでも、この動

作軸に表れている動作の主要な情緒表現に関して解析を行う。以下、特に断りがない限り、「機能動作」は機能動作の動作軸を、「情緒動作」は情緒動作の動作軸を意味することとする。

5.2.2 動作の構成に関する考察

本節では、人間の動作に関して分析、考察を行う。まず、人間の動作がどのように構成されているのかを考える。

動作というものはいくつかの段階ごとに、より細かい単位へと分割することができる。最初の段階として、動作をいくつかの機能を持った動作単位へ分割することを考える。以下、この機能を持った動作として最小の動作単位をフレーズ (phrase) と呼ぶ。フレーズがいくつか集まったフレーズ列もまた一つの機能を持った動作であり、これを広い意味でのフレーズと言うこともできる。フレーズの例として「おじぎをする」「手を振る」などが挙げられる。

次の段階として、このフレーズをより細かく分割することを考える。フレーズは機能を持った動作としては最小単位であるが、これは更に機能を持たない単純な動きに分割することができる。この機能を持たない単純な動きを要素動作と呼ぶことにする。要素動作の例として「前に動く」「右回りに回転する」などが挙げられる。

また、この要素動作をより細かく分割すると、動作の最小単位として、一つの姿勢を得ることができる。以下、この一つの姿勢を事象と呼ぶ。

具体例として、測定した文楽人形の演技動作のうち、情緒を伴わない「お茶を渡す」動作を分割してみる。

まず、この「お茶を渡す」動作は

- お茶を差し出す
- おじぎをする

という 2 つのフレーズに分割することができる。このうち「お茶を差し出す」というフレーズは

- 前に進む
- 腕を前に出す
- 手を離す

等の要素動作に分割することができ、「前に進む」という要素動作は

- 元の位置から、わずかに前にいる
- 元の位置から、もう少し前にいる
- 対象との中間の位置にいる

といった事象に分割することができる。

同様に、好意をもった「お茶を渡す」動作に関しても分割を考えたところ、分割構造は同じになった。このことから、同一の機能のフレーズであれば動作の分割構造は同一であると考えられる。そこで、事象とフレーズに関して次のような仮説をたてる。

(仮説 1) フレーズの機能は、そのフレーズを構成する事象の種類と実現順序によって一意に決定される

しかし、ある機能のフレーズを考えたとき、そのフレーズを構成する事象の種類と実現順序は決定されるが、実現されるフレーズは必ずしも一通りに決まるわけではない。例えば、測定した演技動作のように、「お茶を渡す」という動作 (フレーズ) であっても情緒表現が加わることにより、機能動作とは変化が生じている。この情緒表現による変化について次節で詳しく述べる。

5.2.3 情緒表現の要素に関する仮説

前節で、同一の機能を持った動作 (フレーズ) においては、その動作を構成する事象の種類と実現順序が同一であるとの仮説を提案した。しかし、情緒表現を伴った動作 (情緒動作) には情緒表現を伴わない動作 (機能動作) と比べて何らかの変化が生じている。この情緒表現による動作の変化について、5.2.1 節で得られた動作軸をもとに考察を行う。

計算された動作軸の一例を Figure5.2に示す。Figure5.2は「お茶を渡す」動作の動作軸の重心の y 座標を表すグラフで、上からそれぞれ、情緒を伴わない動作 (機能動作) 、好意を持った動作、敵意を持った動作 (情緒動作) である。

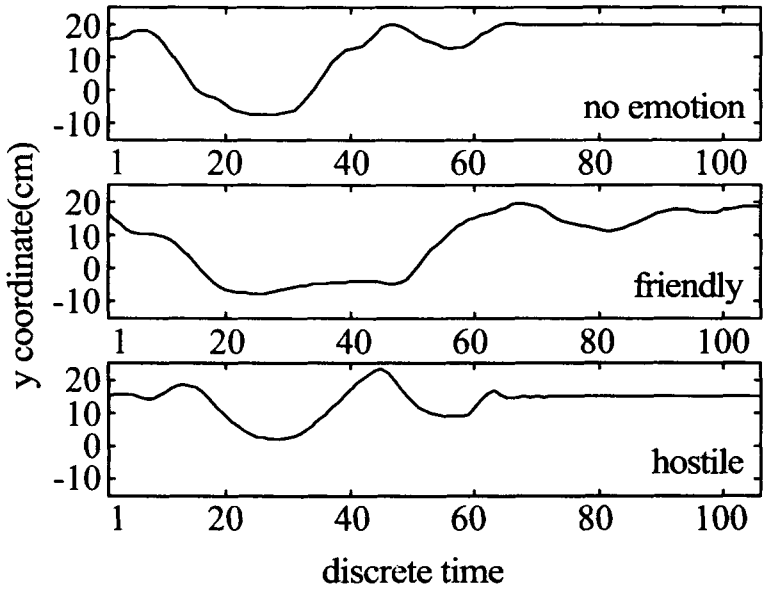


Figure 5.2: Motion axes (serving tea).

動作軸によって比較を行った結果、機能動作と比べて情緒動作には以下のような変化が生じていると感じられた。

- 動作時間の変化
- 動作の大きさの変化

これらを 5.2.2 節で述べた動作の分割構造と併せて考え、次のように言い換える。

- フレーズを構成する事象の実現時刻の変化
- フレーズを構成する事象の実現値の変化

このように、動作時間の変化を事象の実現時刻の変化、動作の大きさの変化を事象の実現値の変化と考え、情緒表現による動作の変化を大きく 2 つに分けることができた。これらをまとめ、次のような仮説を提案する。

(仮説 2) フレーズに情緒表現が加わることにより、そのフレーズを構成する事象の実現時刻および実現値が変化する

事象の実現時刻、実現値の２つの変化によって動作に情緒表現が加わるので、これらの変化を生み出す要素が情緒表現の要素であると言える。以下、前者を情緒表現による位相変化、後者を情緒表現による振幅変化と呼び、それぞれの変化を生み出す要素を、情緒表現における位相要素、情緒表現における振幅要素と呼ぶことにする。

次節以降で、情緒表現における位相変化、振幅変化についてより詳しく考察し、情緒表現における位相要素、振幅要素を抽出する手法を提案する。

5.3 情緒表現における位相要素の抽出

前節で、情緒表現の要素として位相要素と振幅要素の2つの要素を提案した。本節ではこのうち情緒表現における位相要素について詳しく考察し、情緒動作から情緒表現における位相要素を抽出する手法を提案する。逆に抽出された位相要素を機能動作に加えることによって、情緒動作が得られることになる。

情緒表現による動作の位相変化は、フレーズを構成する事象の実現時刻の変化である。この位相変化によってもフレーズを構成する事象の種類と実現順序は変化しないので、機能動作と情緒動作の同一機能のフレーズ間には対応する事象が存在することになる。したがって機能動作と情緒動作の同一機能のフレーズ間の対応する事象を見つけ出し、その実現時刻が揃うように位相を調整して位相変化を取り除くことで、情緒表現における位相要素を抽出することができる。つまり機能動作の任意の時刻の事象に対して、それと同一とみなせる情緒動作の事象を対応させ、対応付けられた情緒動作の事象の時刻を機能動作の事象と同じ時刻に修正して新しい動作軸の時系列を作成した物を、位相要素を抽出した情緒動作と呼ぶ。

何をもって機能動作の事象と情緒動作の事象が対応しているとみなすのか問題であるが、後に述べるように、グラフの局所的な形が似ている事象どうしを、対応しているとみなすのが妥当である。グラフの局所的な形は、グラフで表されている時系列の局所的なウェーブレット係数で表されるので、ウェーブレット係数が似ている事象どうしを、対応する事象とみなすことになる。

以下の節でくわしく論じるように、Figure5.3 の上のグラフ (情緒なし) と中のグラフ (好きという情緒) のように、情緒の異なる動作軸のグラフの間でウェーブレット係数の最も近い時刻どうしを、同一の動作の事象とみなして対応付けることになる。

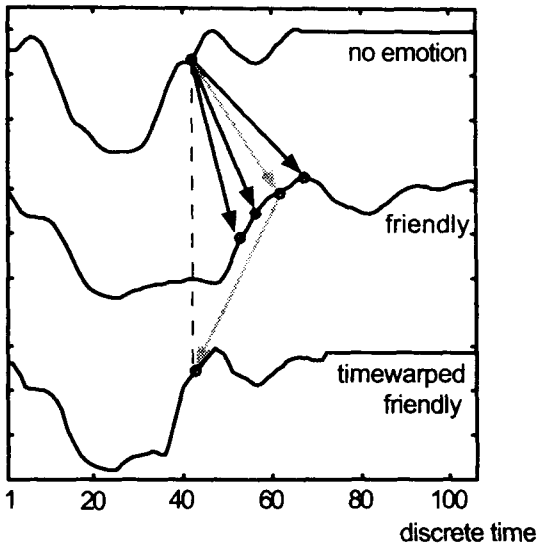


Figure 5.3: Find the same events between 2 motions.

5.3.1 ウェーブレット係数

時系列の任意の時刻における様々な周波数の成分を表すのがウェーブレット係数である。事象を対応付けるにあたって、あらゆるスケール (周波数) のウェーブレット係数を比較するわけにはいかないので、下記のような工夫を行う。ウェーブレット係数においては、周波数そのものではなく、その逆数であるスケールが用いられる。したがって周波数が高いということは、スケールが小さいことであり、周波数が低いということは、スケールが大きいことである。

時系列 $x(t)$ の時刻 t におけるスケール c のウェーブレット係数 $d(t,c)$ は

$$d(t,c) = \int_{-\infty}^{\infty} x(s) \frac{1}{\sqrt{c}} \Psi\left(\frac{s-t}{c}\right) ds \tag{5.12}$$

で与えられる。

そして、ウェーブレット係数 $d(t,c)$ から元の時系列 $x(t)$ を再構成するウェーブレット逆変換は次式で与えられる。

$$x(t) = \frac{2}{C_\psi} \int_0^{\infty} \int_{-\infty}^{\infty} d(s,c) \frac{1}{\sqrt{c}} \Psi\left(\frac{t-s}{c}\right) \frac{dsdc}{a^2} \tag{5.13}$$

$$\frac{C_\psi}{2} = \int_0^\infty \frac{|\hat{\Psi}(\omega)|}{\omega} d\omega \tag{5.14}$$

ここで、$\Psi(t)$ はアナライジングウェーブレットと呼ばれる信号を切り出す時の単位である。C_ψ は、アナライジングウェーブレットによって決まる定数であり、$\hat{\Psi}(\omega)$ はアナライジングウェーブレットのフーリエ変換である。

この $\Psi(t)$ を引き伸ばし (もしくは圧縮) することにより取り出される周波数成分が変化する。この伸縮の度合いを表すのがスケール　a である。このスケールが大きいほど $\Psi(t)$ は引き伸ばされ、低い周波数成分を取り出すことになる。このアナライジングウェーブレットを信号全体にわたってシフトすることにより、各時刻における各周波数成分を取り出すことができる。ウェーブレット係数 $d(t,c)$ の t が、そのシフトを表している。

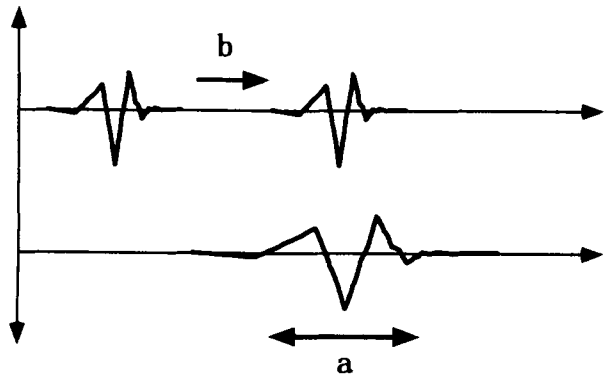


Figure 5.4: Scaling and Shift of analyzing wavelet.

本研究ではウェーブレット解析用のツールとして MATLAB を用いた。また、アナライジングウェーブレットとして symlet3 を用いた。Figure5.5にその波形を示す [19]。

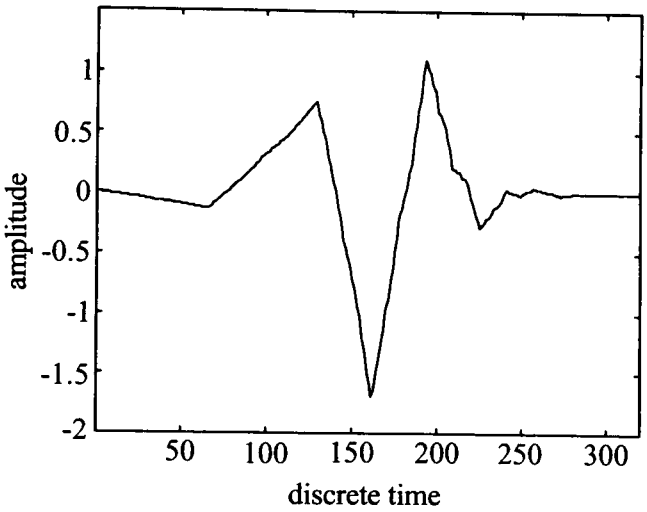


Figure 5.5: Analyzing wavelet 'symlet3'.

5.3.2 スケールと approximation

ウェーブレット変換を用いて動作時系列 (信号) の周波数成分をすべてウェーブレット係数で表現するためには、無限大のスケール a まで計算する必要がある。しかし、実際は有限である動作時系列の長さにより、計算できる最大スケールが決定される。信号を周期的に取り扱うことによって最大スケールを大きくすることも可能であるが、無限大のスケールまで計算することは数値計算上不可能である。したがって最大スケールに値する周波数以下の周波数成分をウェーブレット係数で表現することはできない。そこで、ウェーブレット係数で表現することができない低い周波数の成分を別途考える必要がある。これは approximation と呼ばれる [19]。

アナライジングウェーブレット symlet3 におけるスケールと周波数の関係を調べるため、スケール 1 の symlet3 のフーリエ変換を計算した。Figure5.6にその結果を示す。

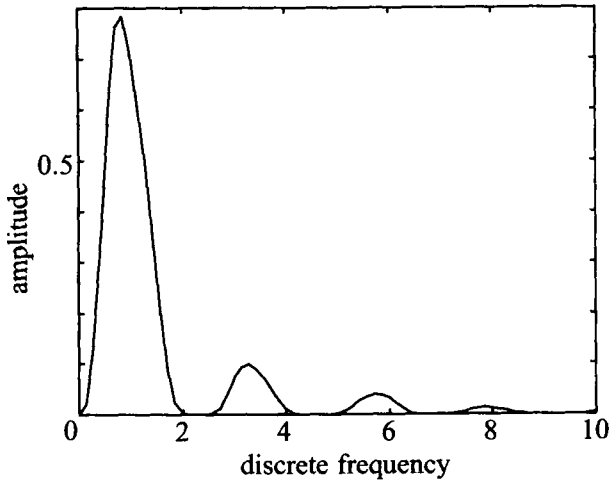


Figure 5.6: Frequency-domain representation of 'symlet3'.

文楽人形の動作時系列を離散時系列と考えたとき、動作時系列に含まれる最高離散周波数は 1 となる。Figure5.6に見られるように、スケール 1 の symlet3 には離散周波数 1 の周波数成分が十分含まれているので、動作時系列の最高周波数成分はウェーブレット係数で表現されている。また、文楽人形の動作時系列長は 106 であり、アナライジングウェーブレット symlet3 を用いた場合、時系列長によって決定する最大スケールは 21 となる。このとき離散周波数 0.0408 以下の周波数成分はウェーブレット係数として表現できないことになるので、この周波数成分を approximation として考える必要があるが、MATLAB のウェーブレットツールボックスには、連続ウェーブレット変換の逆変換、approximation の計算プログラムが用意されていない。そこで、定義にしたがいそれらの計算プログラムを作成した。このウェーブレット逆変換と approximation の計算プログラムの確認のため、簡単なサンプル信号を作成し、ウェーブレット係数の逆変換の結果である再構成信号 (reconstruction signal) と approximation との比較を行った。

サンプル信号として、以下のサンプル信号 A とサンプル信号 B を足しあわせたものを用いた (Figure5.7)。このサンプル信号について、いくつかの範囲のスケールで再構成信号と approximation を計算した。その結果を Figure5.8から Figure5.11に示す。

- サンプル信号 A : 離散周波数約 0.110 の正弦波→スケール 8 に相当
- サンプル信号 B : 離散周波数約 0.027 の正弦波→スケール 32 に相当

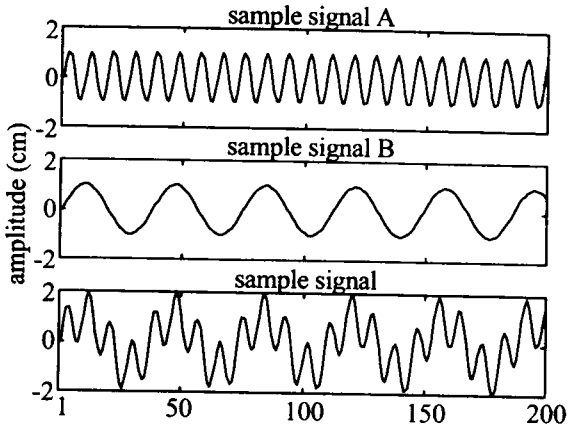


Figure 5.7: Sample signal.

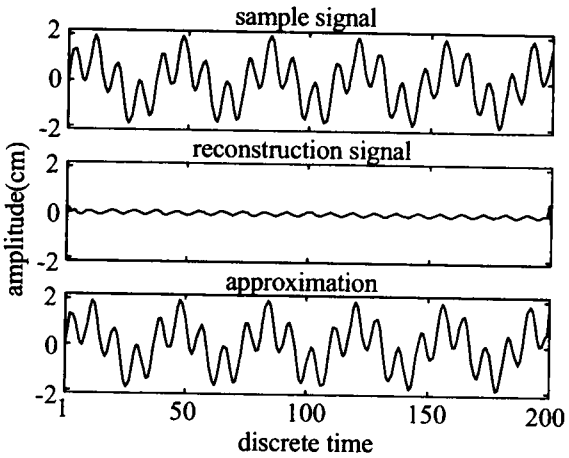


Figure 5.8: Scale 1 to 4.

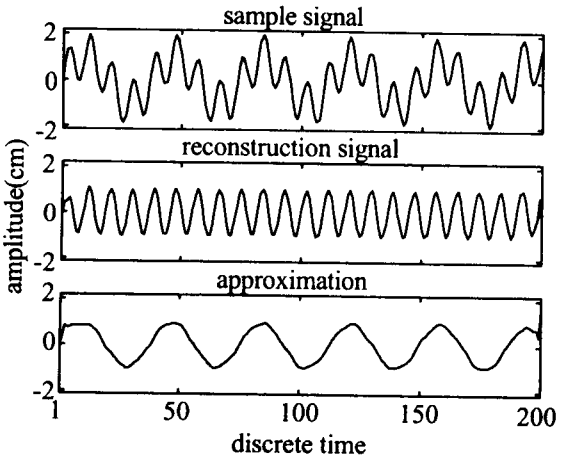


Figure 5.10: Scale 1 to 12.

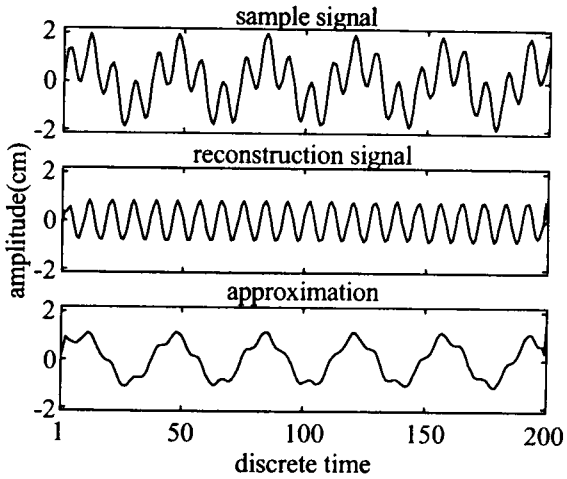


Figure 5.9: Scale 1 to 8.

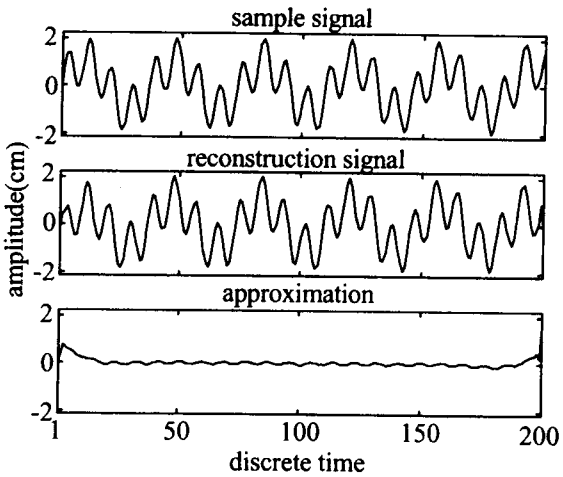


Figure 5.11: Scale 1 to 32.

Figure5.8においては、ウェーブレット係数で表現される周波数成分がサンプル信号に含まれていないため、サンプル信号がapproximationに残っている。Figure5.9においては、サンプル信号に含まれるサンプル信号Aの成分がウェーブレット係数として表現されるはずであるが、この成分は完全には取り出されておらず、approximationにいくらか残っている。これは、アナライジングウェーブレットによって切り出される周波数成分には幅があるためと考えられる。しかし、Figure5.10に見られるように、スケール12までウェーブレット係数を計算したとき、サンプル信号Aの成分はすべてウェーブレット係数で表現されている。また、Figure5.11においては、スケール32まで計算すると、サンプル信号AおよびBがほぼ完全にウェーブレット係数で表現されていることがわかる。

結果として、わずかな誤差は残るがほぼ正確なウェーブレット逆変換とapproximationの計算が行われていることが確認された。

5.3.3 位相調整手法の提案

機能動作 $X_{ref}(t)$ $(t = 0, 1, 2, \cdots, T)$ の各事象に対応する情緒動作 $X_{smp}(t)$ $(t = 0, 1, 2, \cdots, T)$ の事象を検出して、情緒の位相要素を抽出する方法は、下記のように、機能動作 $X_{ref}(t)$ $(t = 0, 1, 2, \cdots, T)$ を参照信号 (reference signal) とし、情緒動作 $X_{smp}(t)$ $(t = 0, 1, 2, \cdots, T)$ をサンプル信号 (sample signal) として、第 3 章で提案された方法で標準化する方法を改良した物である。

$d_{ref}(t, c)$ ： 参照信号の時刻 t スケール c のウェーブレット係数

$d_{smp}(t, c)$ ： サンプル信号の時刻 t スケール c のウェーブレット係数

とする。

参照信号の任意の時刻 t_{ref} におけるウェーブレット係数と最も近いウェーブレット係数を持つサンプル信号の時刻 t_{smp} を対応する時刻と決定する。このとき各時刻の信号としての特徴を考える上で、その時刻の前後の特徴もその時刻と関連があると考え、前後 τ 時刻のウェーブレット係数も考慮にいれる。この参照する前後時刻の範囲 τ を参照範囲と呼ぶ。従って、以下のような評価式 $V(t_{smp})$ が最小となる t_{smp} を t_{ref} に対応する時刻とする。

$$V(t_{smp}) = \sum_{s=-\tau}^{\tau} \sum_{c=1}^{c_{max}} |d_{ref}(t_{ref} + \tau, c) - d_{smp}(t_{smp} + \tau, c)|^2 \tag{5.15}$$

最適な対応が得られるように、最大スケール c_{max}、参照範囲 τ を調整し、これを $t_{ref} = 0, 1, \cdots, T$ まで繰り返すことによって、基準信号とサンプル信号の対応が決定される。

第 3 章で確認したように、同一の機能、情緒である 5 回の動作の標準化においては、これらがほぼ同様の特徴を持つため、従来のウェーブレット係数のみの比較で適切な対応関係が得られた。しかし、これを異なる情緒の動作時系列に適用した結果、不適切な対応をとる時刻が存在した。これは、ウェーブレット係数のみの比較では不十分であるからだと考えられる。そこで本章では、異なる情緒の動作時系列に適用できるように上記の手法を改良した。以下に改良した手法を示す。

まず、基準信号とサンプル信号をそれぞれウェーブレット変換し、それぞれのウェーブレット係数および approximation を計算する。

$$
\begin{aligned}
d_{ref}(t,c) &: \text{参照信号のウェーブレット係数} \\
d_{smp}(t,c) &: \text{サンプル信号のウェーブレット係数} \\
A_{ref}(t) &: \text{基準信号の } approximation \\
A_{smp}(t) &: \text{サンプル信号の } approximation
\end{aligned}
$$

また、最小化評価式として、approximation も考慮した以下の $V'(t_{smp})$ を用いる。

$$
\begin{aligned}
V'(t_{smp}) &= \sum_{s=-\tau}^{\tau} \sum_{c=1}^{c_{max}} |d_{ref}(t_{ref}+s,c) - d_{smp}(t_{smp}+s,c)|^2 \\
&+ M \sum_{s=-\tau}^{\tau} |A_{ref}(t_{ref}+s) - A_{smp}(t_{smp}+s)|^2
\end{aligned} \tag{5.16}
$$

最適な対応が得られるように、最大スケール c_{max}、参照範囲 τ、approximation の重みづけ M を調整し、これを $t_{ref} = 0, 1, \cdots, T$ まで繰り返すことによって、基準信号とサンプル信号の対応が決定される。この対応関係に基づいて、対応する時刻が揃うようにサンプル信号を局所的に伸縮させて位相を調整する。これを位相調整と呼ぶ。この概念図は、5.3 節の Figure5.3 で示されている。

この手法の検証のため、簡単な基準信号、サンプル信号を作成し、その対応を決定した。この対応関係を Figure5.12左図のように、横軸に基準信号の時刻、縦軸にサンプル信号の時刻をとったグラフで表す。また決定した対応関係を元に、対応する時刻が揃うように信号を局所的に伸縮して位相を調整した結果を Figure5.12右図に示す。Figure5.12右図は上から順に基準信号、サンプル信号、位相を調整したサンプル信号 (timewarped signal) である。この結果より、適切な対応が検出されていることが確認された。

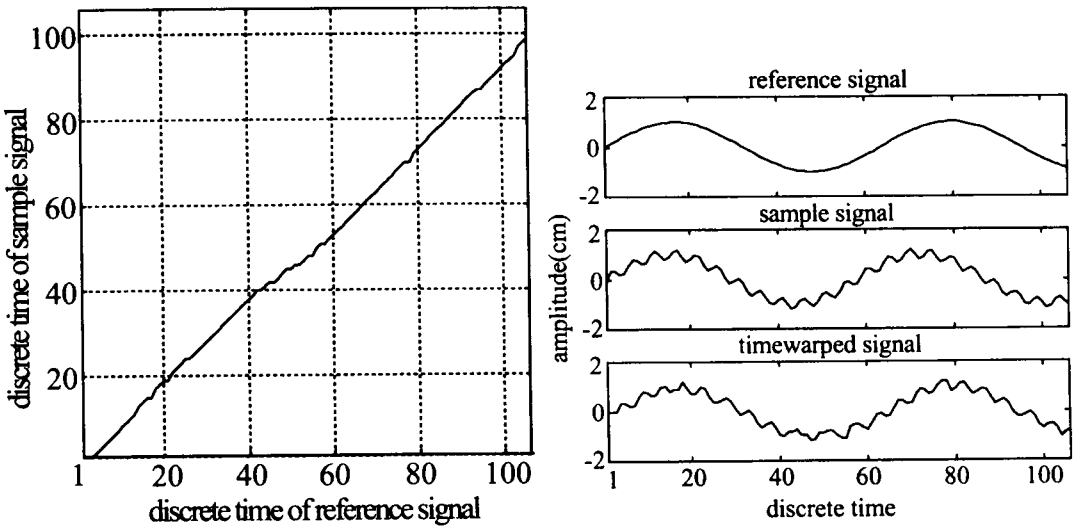


Figure 5.12: Time-correspondence and Timewarp of sample signal.

5.3.4 位相要素の抽出

本節では、前節のアルゴリズムを動作軸の動作時系列に適用し、情緒表現における位相要素の抽出を行う。

まず、機能動作を基準信号、情緒動作をサンプル信号とし、各時刻における機能動作の事象に対応する情緒動作の事象の時刻を、動作軸の 6 自由度それぞれに対して求めた。このとき、スケール範囲を 1 から 16、approximation の重みづけは一つのスケールが 1 に対して 8 とした。

この結果より、動作軸の 6 自由度それぞれにおいての対応関係は非常によく似ていることがわかった。そこで、動作軸の 6 自由度の評価式 (5.16) を足しあわせた

$$\sum_{q=1}^{6} V'_q(t_{smp}) \rightarrow min \tag{5.17}$$

を最小化評価関数として一つの動作全体に対する対応を計算した。「お茶を渡す」の機能動作と情緒動作における対応関係の結果を Figure5.13に示す。Figure5.13左図が 6 自由度それぞれにおける対応関係、右図が 6 自由度同時の評価式を用いたときの対応関係である。

この対応を用いて動作軸の 6 自由度すべての位相を調整した結果、情緒表現によ

る位相変化を取り除き、情緒表現における位相要素を抽出することができた。この結果の一部をを Figure5.14に示す。位相調整された情緒 (好意) の動作軸のグラフ (一番下) が情緒なしの動作軸 (一番上) に、かなり近い物となっている。

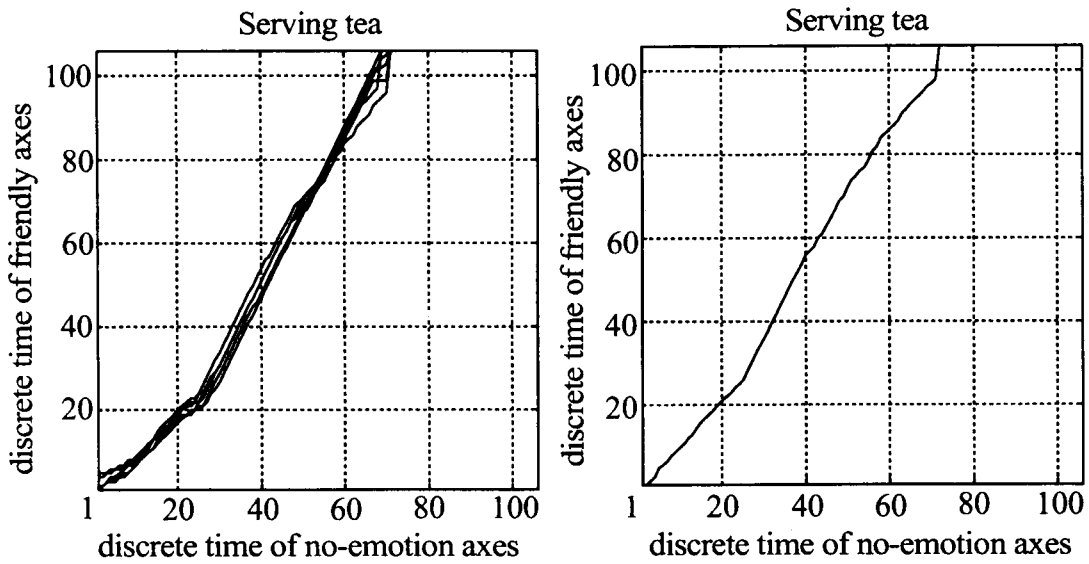


Figure 5.13: Time-correspondent of motion axes.

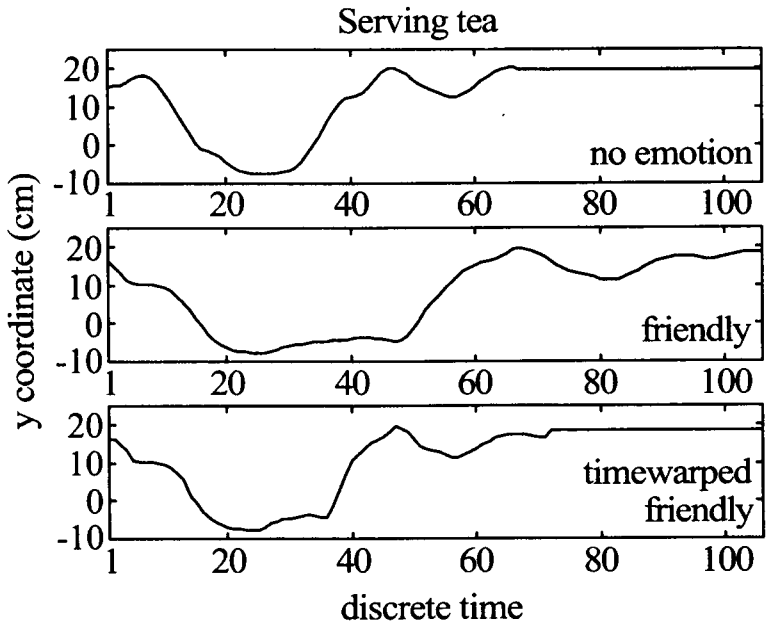


Figure 5.14: Timewarp of friendly axes.

5.3.5 結果と考察

本節では、情緒表現における位相要素について考察を行い、これを抽出する手法を提案した。

5.2 節で提案した仮説から、機能動作と情緒動作の同一機能のフレーズ間には対応する事象が存在すると言えた。本章では、この対応する事象を検出し、位相を調整することで、情緒表現における位相要素を抽出することを試みた。

異なる情緒のフレーズ間の対応する事象を検出するため、第 3 章で考案された動作時系列の標準化手法を改良した。これを機能動作と情緒動作の動作軸に適用することにより、適切な対応関係が得られた。この対応関係を用いて情緒動作の位相を調整した結果、情緒動作の位相を機能動作にあわせることができた (Figure5.14 など)。よって、情緒表現による位相変化を取り除き、位相要素を抽出することができたといえる。

また、これを用いて情緒動作全体、すなわち 24 の自由度すべての位相を調整した。この動作を CG アニメーションによって確認した結果、動作の機能を失うことはなく、自然な動作として機能動作と同じ位相の動作を実現することができた。

この位相要素を同一の情緒、同一の機能の 5 回のフレーズについて比較した結果、これらがほぼ同様の対応関係をとっていることがわかった。

5.4 動作軸の振幅要素の抽出

5.4.1 位相抽出の結果の検討

前節において、情緒表現の2つの要素のうち、位相要素を抽出することができた。位相要素を抽出した情緒動作の一例を、次ページの Figure 5.15 に示して、情緒なしの機能だけの動作と比較する。位相抽出された好意の動作軸と情緒なしの動作軸を比較すると、位相はそろっているが、振幅の大きさに違いがみられる。そこで、本節では情緒表現における振幅要素について詳しく考察し、位相要素を抽出した情緒動作から振幅要素を抽出する手法を提案する。

5.2節において、情緒表現による振幅変化はフレーズを構成する事象の実現値の変化であるとの仮説を提案した。この振幅変化を生み出す振幅要素を抽出することを試みて、信号の局所的な特徴をあらわすウェーブレット係数を用いて機能動作と情緒動作の対応する事象の特徴を比較する。しかし 5.3.2節で述べたように、信号に含まれる周波数成分をすべてウェーブレット係数で表現するためには、無限大のスケールまでウェーブレット係数を計算する必要があり、これは有限長の時系列では不可能である。

そこでまず、各スケール毎に機能動作と情緒動作のウェーブレット係数、approximation を計算し、動作軸の各自由度毎に比較した。その結果、スケールが大きくなるにつれて機能動作と情緒動作の approximation が類似してくることが感じられた。あるスケールにおける approximation はそのスケールに対応する周波数以下の周波数成分であり、approximation の類似度が十分高ければ、そのスケールに対応する周波数以下の周波数成分には機能動作と情緒動作の変化が表れていないと言うことができるので、その周波数以下の周波数成分は無視することができる。

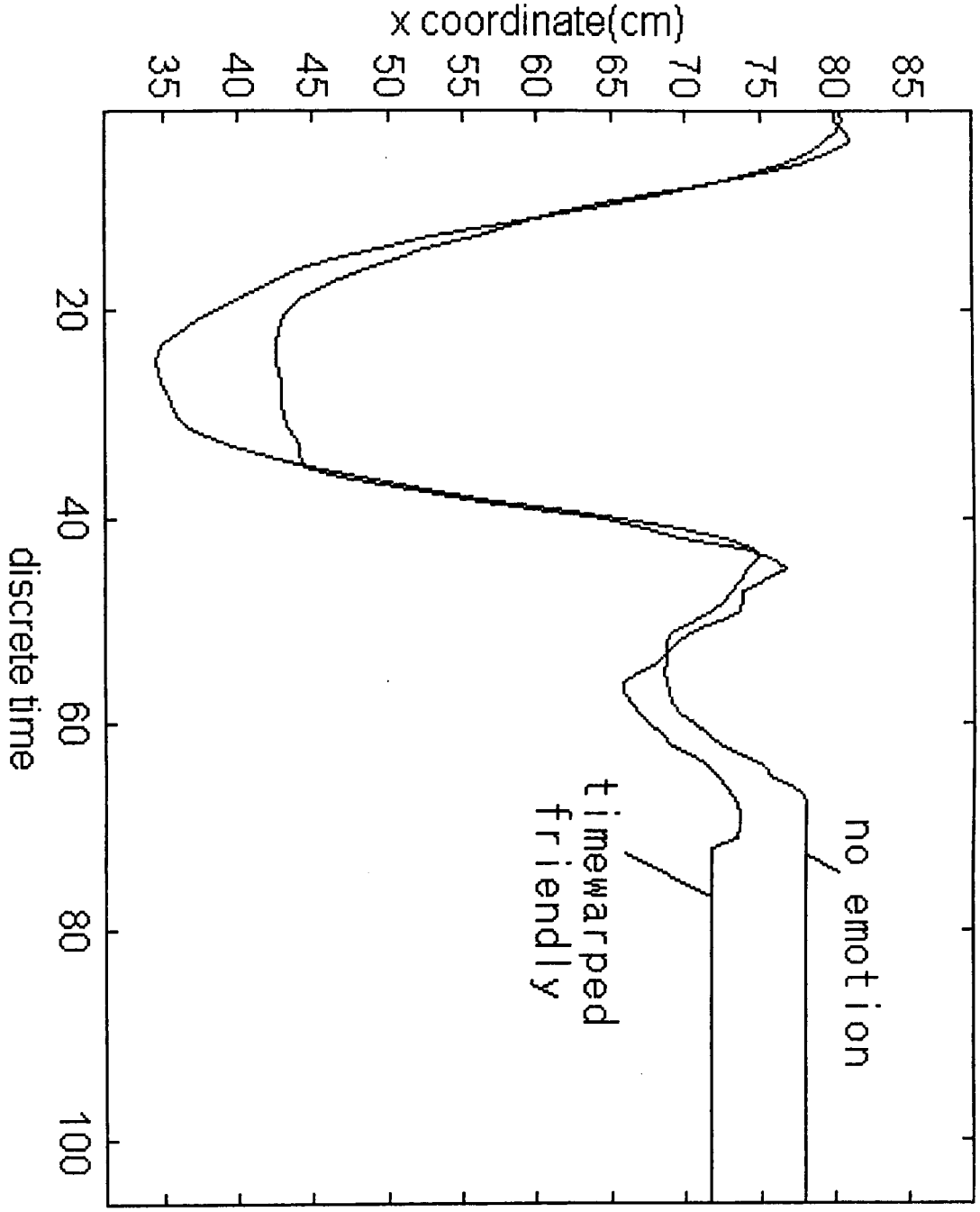


Figure 5.15: Timewarped friendly motion axes.

類似度について調べるため、機能動作と情緒動作の approximation の相対誤差を各スケール毎に計算した。この相対誤差が小さければ、類似度が高いことになる。また、approximation が類似していなくとも、その値が動作と比べ十分に小さければ誤差の影響も小さいので、相対誤差の計算式として以下の式を用いる。また、この計算結果を Figure5.16に示す。

$$\frac{\sqrt{\sum_{t=1}^{T}|A_e(t)-A_f(t)|^2}}{\sqrt{\sum_{t=1}^{T}|X_f(t)|^2}} \tag{5.18}$$

$X_f(t)$: 機能動作

$A_f(t)$: 機能動作の *approximation*

$A_e(t)$: 情緒動作の *approximation*

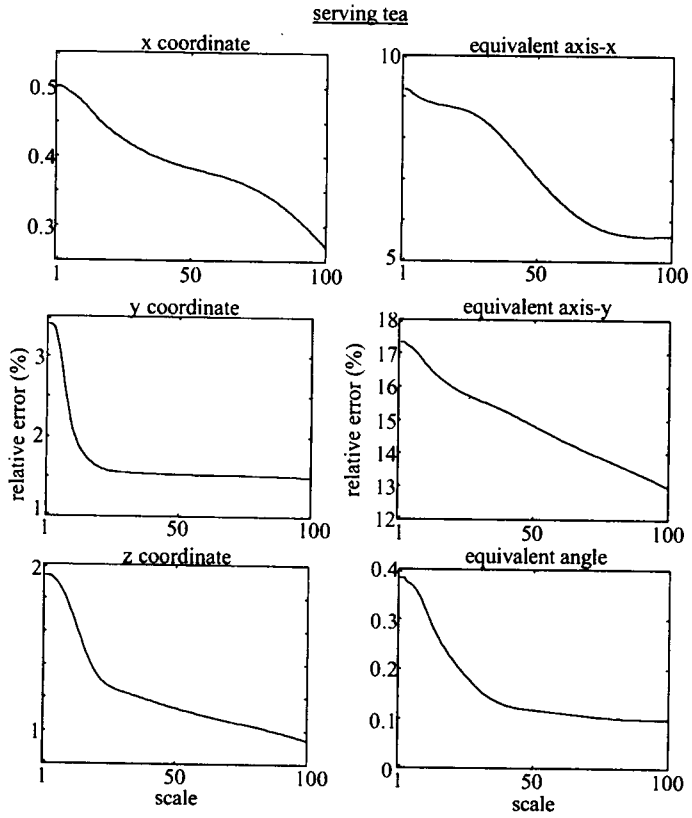


Figure 5.16: Relative error (no emotion and friendly).

この結果より、スケールが大きくなるにつれて機能動作と情緒動作の approximation が類似してくることが確認された。しかし、相対誤差の値は動作軸の各自由度毎にばらつきがあり、自由度によっては、十分大きなスケールにおいても相対誤差

が無視できないほどに大きいことがわかった。この事より、信号の高い周波数に注目しているウェーブレット係数を用いて振幅要素の抽出を試みる事は適切ではないと判断される。

5.4.2 情緒動作に対する機能動作のゲイン

そこで 次ページの Figure 5.17 のように動作軸を幾つかのフレーズ (phrase) に分割して、各フレーズにおける機能動作と情緒動作の振幅の違いを、信号の低周波成分に注目するフーリエ級数を用いて説明する。

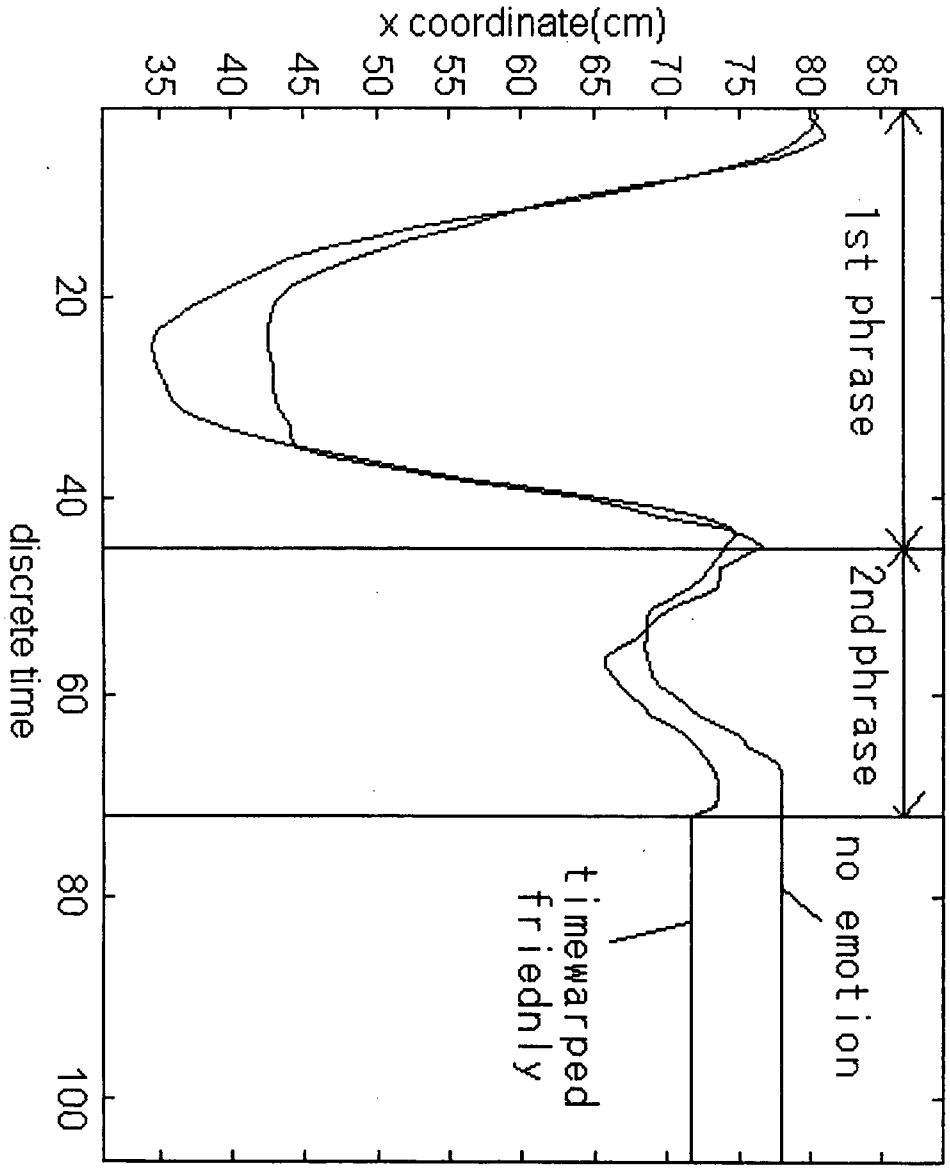


Figure 5.17: Phrases of motion axes.

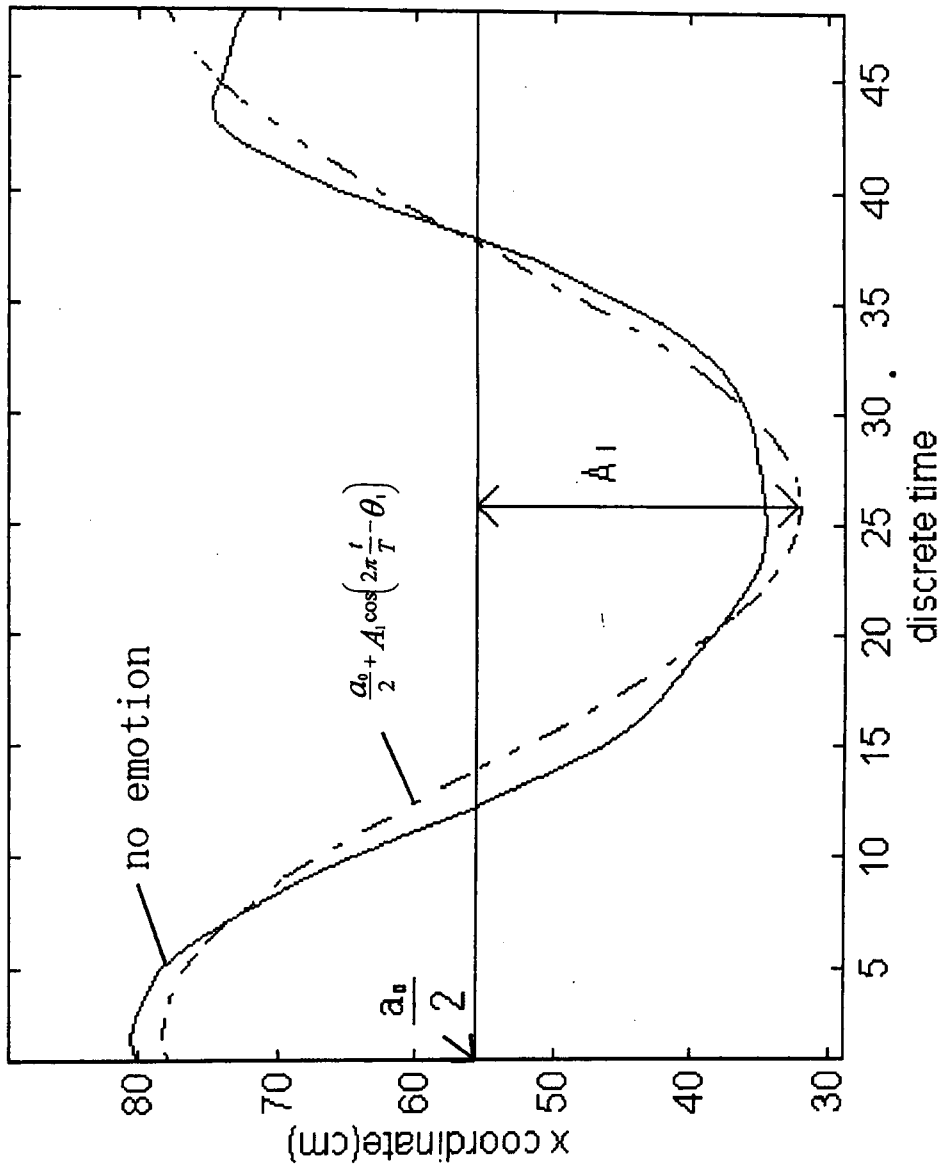


Figure 5.18: Fourier series.

各フレーズの動作軸 $X(t)$ はフレーズの長さを T とすると、次のようにフーリエ級数を用いて表現できる。

$$x(t) = \frac{a_0}{2} + A_1 \cos\left(2\pi\frac{t}{T} - \theta_1\right) + A_2 \cos\left(2\pi\frac{2t}{T} - \theta_2\right) + \ldots \tag{5.19}$$

ここで、動作軸の大まかな形を表すために、第 0 次高調波成分 (直流成分) と第 1 次高調波成分に注目する。動作軸を第 0 次成分と第 1 次成分であらわしたものを、前ページの Figure 5.18 に示す。元々の情緒なしの動作軸を実線で表し、その 第 0 次成分と第 1 次成分による近似を破線で表す。

これより、第 0 次と第 1 次のフーリエ係数を用いることによってそれらを用いて、情緒動作に対する機能動作のゲインを定義することができる。情緒なしの動作軸 $X^{no}(t)$ を

$$X^{no}(t) = \frac{a_0^{no}}{2} + A_1^{no} \cos\left(2\pi\frac{t}{T} - \theta_1^{no}\right) + A_2^{no} \cos\left(2\pi\frac{2t}{T} - \theta_2^{no}\right) + \ldots \tag{5.20}$$

と近似して、位相調整された情緒動作 $X^{em}(t)$ を

$$X^{em}(t) = \frac{a_0^{em}}{2} + A_1^{em} \cos\left(2\pi\frac{t}{T} - \theta_1^{em}\right) + A_2^{em} \cos\left(2\pi\frac{2t}{T} - \theta_2^{em}\right) + \ldots \tag{5.21}$$

と近似することができる。

情緒動作に対する機能動作のゲインは、第 0 次の項については

$$g_0 \stackrel{\text{def}}{=} \frac{a_0^{no}}{a_0^{em}} \tag{5.22}$$

と定義され、第 1 次の項については、

$$g_1 \stackrel{\text{def}}{=} \frac{A_1^{no}}{A_1^{em}} \tag{5.23}$$

と定義される。

それぞれのフレーズにおいて、情緒動作 $X^{em}(t)$ に上記のゲインを施して生成された動作軸 $X^{em}_{\text{抽出}}(t)$ は

$$\theta_1^{no} \doteq \theta_1^{em} \tag{5.24}$$

$$A_1 \gg A_2, A_3, \ldots \tag{5.25}$$

より

$$\begin{aligned}
&X^{em}_{\text{抽出}}(t) \\
&= g_0\frac{a_0^{em}}{2} + g_1 A_1^{em} \cos\left(2\pi\frac{t}{T} - \theta_1^{em}\right) + g_1 A_2^{em} \cos\left(2\pi\frac{2t}{T} - \theta_2^{em}\right) + \ldots && (5.26)\\
&= g_0\frac{a_0^{em}}{2} + g_1 \left[A_1^{em} \cos\left(2\pi\frac{t}{T} - \theta_1^{em}\right) + A_2^{em} \cos\left(2\pi\frac{2t}{T} - \theta_2^{em}\right) + \ldots\right] && (5.27)\\
&= g_0\frac{a_0^{em}}{2} + g_1 \left(x^{em}(t) - \frac{a_0^{em}}{2}\right) && (5.28)
\end{aligned}$$

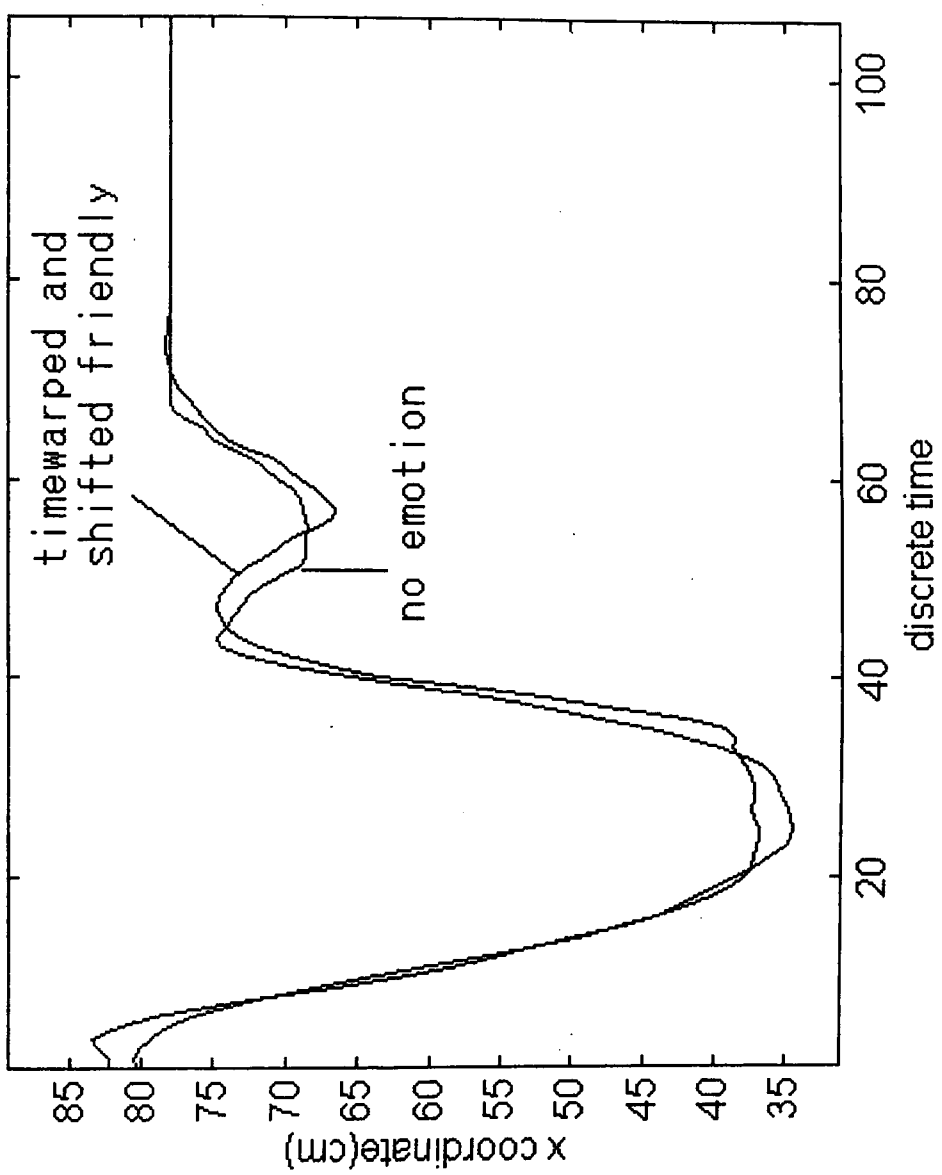


Figure 5.19: Timewarped and shifted motion axes.

で求められる。このように各フレーズにおいて、情緒要素から振幅要素を抽出した得られた $X^{em}_{抽出}(t)$ が得られるので、それらをスプライン補間によってフレーズ間で滑らかに接合することによって、振幅要素を抽出した動作軸が求められる。

このように、位相調整された情緒動作にゲインを施して (情緒動作の振幅要素を抽出して) 得られる動作軸の例を、前ページの Figure 5.19 に示して、情緒なしの動作軸 (機能動作) と比較する。振幅要素を抽出した結果、情緒動作は機能動作にかなり近くなっている。

上記に提案した振幅要素の抽出方法を、「お茶を渡す」の機能動作と情緒動作に適用した結果の一部を Figure5.20に示す。その結果、機能動作とほぼ一致した情緒動作を得ることができた。よって、情緒表現における振幅要素を抽出することができたといえる。

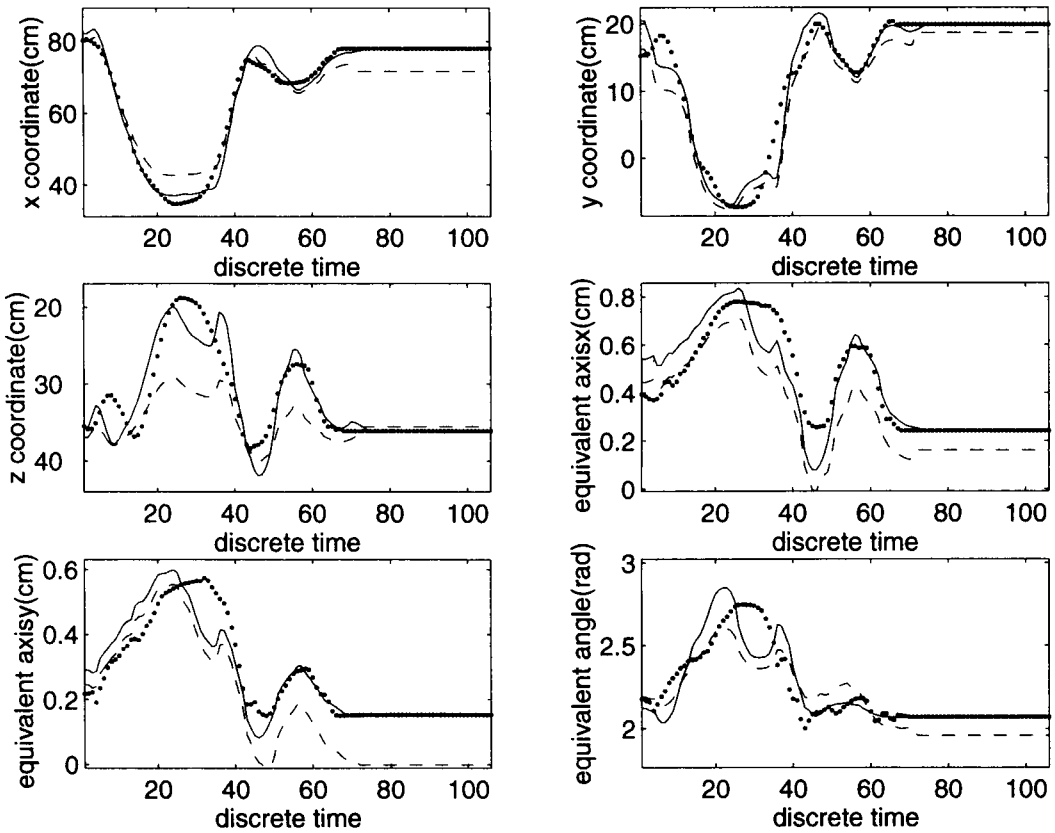


Figure 5.20: Timewarped friendly motion axes.

5.4.3 結果と考察

本節では、情緒表現における振幅要素について考察を行い、これを抽出する手法を提案した。

5.3 節において、機能動作と情緒動作の同一フレーズ間の対応を検出することにより、情緒表現における位相要素を抽出した。この位相要素を取り除いた情緒動作から振幅要素を抽出するために、いくつかのフレーズに分割して、各フレーズごとに機能動作と情緒動作をフーリエ級数に展開して、第 0 次の項 (直流成分) と第 1 次の項について情緒動作を機能動作に近づけるためのゲインを求めた。

各フレーズごとに上記のゲインを情緒動作に施して、フレーズどうしをスプライン補間して つなぎ合わせると、情緒動作と機能動作の動作軸がほぼ一致した。よって、情緒表現における振幅要素を抽出することができたと言える。

5.5 動作軸の位相・振幅成分の調整による動作設計

動作生成における問題点の一つとして、動作全体の動きのバランスをとることが困難であることが挙げられる。人間のような多関節を有する対象の動作の場合、その動作が自然に見えるためには各部分の動きのバランスが非常に重要である。キャプチャーされた動作に処理を行う場合でも、各部分、各自由度ごとに処理を行うと、全体のバランスが崩れてしまうことがおこり得る。

そこで本章では、動作生成の手法として、動作軸に注目した生成手法を提案する。動作軸は動作の主要な部分をあらわす要素であり、この動作軸のみに処理を加えることで、動作全体のバランスを崩すことなく新たな動作を生成することが可能であると考えた。

また、5.3 節 5.4 節で情緒表現における位相要素、振幅要素を抽出する手法を提案したが、この手法を用いて各要素を取り除くことで情緒動作から機能動作を作り出すことができた。従って、逆に各要素を機能動作に加えることで、情緒動作、すなわち情緒表現を伴った動作を生成することができる。各要素を機能動作に加えるためには、各要素の抽出手法の逆手法を用いれば良い。このとき、抽出された情緒表現の各要素はそれぞれ、位相要素が動作時間の長さを、振幅要素は、各フレーズのゲインが動きの大きさを、あらわすものであると考えられる。これらの要素はいずれも簡単な時刻対応のグラフ (Figure 5.13 のような) や数値としてあらわされているので、これらを動作生成における設計パラメータと考え、動作生成において、必要に応じて値を調整することで、新たな特徴を持った動作を生成することができる。

以下に、動作軸に注目した情緒動作の生成手法を示す。

1. 動作軸と相対動作の計算

   生成する動作の元となる機能動作について、動作軸と動作軸に対する相対動作を計算する。

2. 各フレーズの振幅の増幅・減衰のゲインの調整

   動作を構成する各フレーズごとのゲインを増減し、増幅・減衰を調整する。この増幅・減衰の値を変化することで、動きの大きさを変化させることができる。

3. 動作の再構成

各調整を行った動作軸と相対動作を組み合わせることで、動作を再構成する。

4. 位相の調整

位相変化の要素を用いて、動作を局所的に伸縮させることで位相を調整する。Figure 5.13 のような、時刻対応のグラフを適切に設計することによって、動作時間を変化させることができる。

この手法を用いて、実際に情緒表現を伴った動作を生成することを試みた。生成手法に従って情緒表現を加えた「お茶を渡す」動作の動作軸の一例を Figure5.21に示す。生成した情緒表現を伴った動作軸と、測定した情緒表現を伴った動作軸 (人形遣いが操作したままの物) の比較を行った結果から、測定した動作軸とほぼ同様の動作軸を生成できたことが確認された。また、生成した動作軸と相対動作を組み合わせることで、文楽人形の動きを生成し CG アニメーションで見ることができる。生成された情緒動作と人形遣いが操作したままの情緒動作を CG アニメーションで比較した結果、本手法で生成した情緒動作が人形遣いが操作したままの情緒動作とよく似ていることが確認できた。位相要素と振幅要素だけを用いた設計で、人形遣いが操作した情緒動作に近い動作を、機能動作から生成できた事になるので、位相要素と振幅要素の設計だけでも、様々な情緒動作を生成できることが期待できる。

更に、動作生成におけるパラメータ、すなわち情緒表現の各要素を変化させ、「ゆっくりと動く」「より深くお辞儀をする」などの新たな特徴を持った動作を生成することを試みた。生成した動作を CG アニメーションで確認した結果、パラメータを変化させても自然な動作が生成されることが確認でき、この手法の有効性を示すことができた。

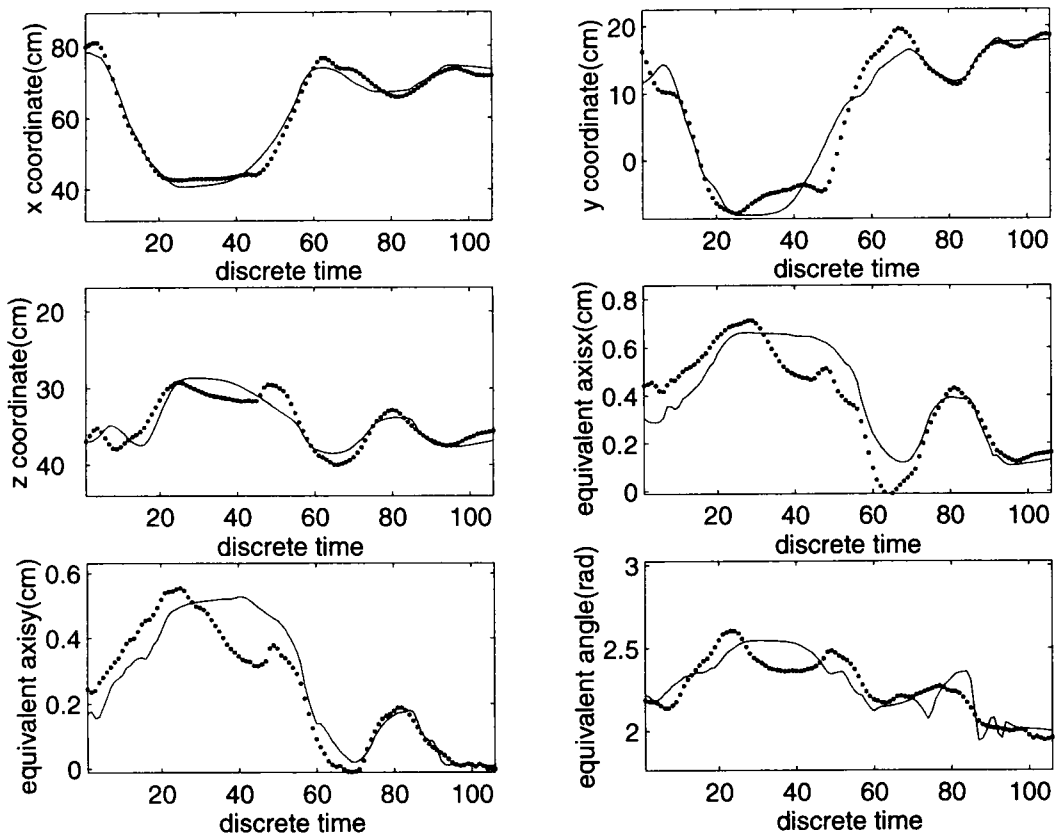


Figure 5.21: Generated motion axes.

5.6 結言

本章では、人間型アミューズメント・ロボットの動作を設計するにあたって、身体の各部位の動きを　ばらばらに設計するのではなく、それらの総体としての整合性のある動作を設計する技法を求めて、文楽人形の動作時系列を解析した。

測定された文楽人形の動作時系列を、（1）人形の慣性主軸の動きという主要な部分と（2）慣性主軸に対する相対動作という副次的な部分　の2つに分解する事を考えた。

動作の主要な部分を表している慣性主軸の動きが、動作の情緒の違いによって変化する様子を、時間伸縮と振幅値の増幅・減衰によって説明できる事を示した。

さらに時間伸縮と振幅値の増幅・減衰を適切に設計する事で、情緒のない動作の軸の動きを加工して、情緒のある動作の軸の動きを生成する事を提案した。動作の主要部分を設計する一つの技法を示した事になる。Figure 5.13 のような位相要素の対応を示す簡単なグラフの作成と振幅を増幅・減衰する数個のゲインの設定だけで、動作軸の動きの設計を行える事が、本手法の特色である。

情緒のない動作から　上記の方法で生成された情緒を有する軸の動きに、情緒のない相対動作を加えた動作をコンピュータ・グラフィックス上で確認した。位相要素と振幅要素だけを用いた設計で、人形遣いが操作した情緒動作に近い動作を、情緒なしの機能だけの動作から生成できたので、位相要素と振幅要素の設計だけでも、様々な情緒動作を生成できることが期待できる。

情緒のない相対動作と情緒のある相対動作を比較検討し、情緒のある相対動作の設計技法を整える事が、今後の課題である。

第 6 章　文楽人形のマルチモーダル動作表示システム

6.1　緒言

前章までの研究において、文楽人形の動作を解析したり設計したりする様々な手法を提案した。これらの有効性を検証するために、文楽人形や動作軸などの動きをコンピュータグラフィクス (CG) で確認しながら、同時に それらのグラフも同時に表示できるような、多方面から動作を表示・解析できるシステムを構築したので、本章でその概要を紹介する。

本研究においては、人形遣いの芸の力によって、あたかも感情を持って生きているように見える文楽人形の動きを対象としている。人がものに対して「感情を持っている」と知覚するには、動きのしなやかさだけでなく、容姿も重要であると考え、よりリアルな人形の姿をした CG を製作することを心がけた。

また、動作解析においてデータの変化を一枚のグラフとして見るだけでなく、データの変化がどのように実際の動作と関連しているかを、アニメーションと照らしあわせて見ることにより、データの性質がより明瞭となり、解析が効果的に行えるようになると考える。

そこで本システムでは、CG による人形モデル (人形オブジェクト) を用いて解析結果のシミュレーションを行い、またデータ同士の関連を見やすくすることで解析の支援を行えるシステムを目指した。これを実現するために、一つのコンピュータグラフィクスアニメーションにおいて、人形の動作、グラフ、動作軸、及び人形部位の動作軌跡を同時に表示できるようにする。

まず CG で人形オブジェクト、グラフオブジェクト、動作軸オブジェクトを製作する。データとオプションに基づき、シーン生成プログラムはそれらのオブジェクトに様々な設定をしたシーンを生成する。シーンは LightWave3D で読み込むことができる。LightWave3D はシーンに基づいてアニメーションを生成する。このアニメーションには各オブジェクトのいくつか、または全てを同時に表示できる。また、

異なる二つのデータから、それぞれ異なる設定をされたオブジェクトを二つ同時に表示することができる。

これにより、あるデータにおける人形の動作、グラフ、動作軸及び人形部位の動作軌跡について、アニメーションによりお互いの関連を視覚的にとらえることができ、それらを同時に考慮しながら効果的な解析が行なえる。また、異なる二つのデータについて異なる人形の動作、グラフ、動作軸及び人形部位の動作軌跡の比較がそれぞれ同時におこなえるので、結果のシミュレーションについて評価を行ないやすくなる。

6.2 全体の構成

本システムは、データ、オブジェクト、シーン生成プログラム、シーン、LightWave3D、及びアニメーションにより構成されている。全体図を Figure 6.1に示す。

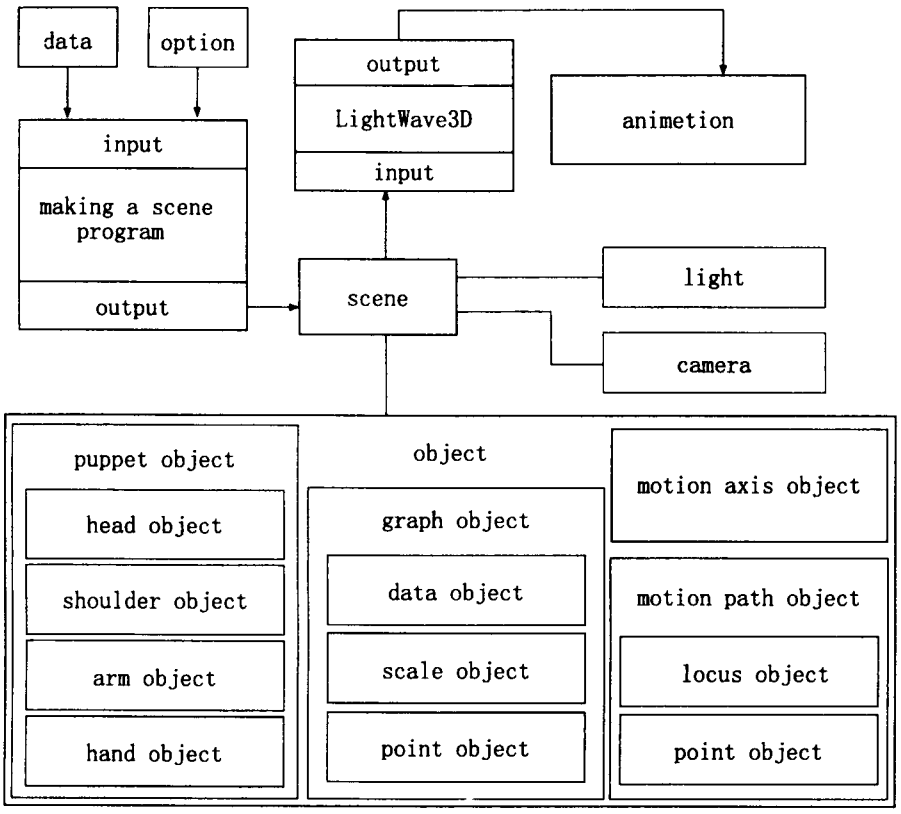


Figure 6.1: Structure of the system.

まずシーン生成プログラムが入力されたデータとオプションに基づいて、シーンを生成する。この時設定するオプションにより、シーンにおける各オブジェクト、カメラ、及びライトについての設定が決まる。これらを適切に設定することにより目的のシーンを得ることができる。また、入力するデータは一つ、もしくは二つである。

生成されたシーンはオブジェクト、ライト、カメラ等から構成される。オブジェクトは人形オブジェクト、グラフオブジェクト、動作軸オブジェクト、及び人形部位の動作軌跡オブジェクトについてそれぞれに表示、非表示の設定ができる。複数のオブジェクトを同時に表示するように設定することにより、それぞれの関連を見ることができる。また、入力したデータが二つのときは、各オブジェクトが二つずつ、それぞれのデータに基づいて設定される。シーンは LightWave3D で読み込むと

ができる。

LightWave3D (NewTek, Inc.) は3次元コンピュータグラフィクスを扱うソフトで、本システムでは読み込んだシーンを元にアニメーションを生成するために使用している。

生成されたアニメーションは実時間での各オブジェクトの動作を表示する。

次節からは、オブジェクト、及びシーンの構成についてそれぞれ詳しく述べる。

6.3 オブジェクトの構成

本システムを構成するオブジェクトを、シーンにおいて表示、非表示の設定ができる単位で分けることにすると、人形オブジェクト、グラフオブジェクト、動作軸オブジェクト、及び人形部位の動作軸オブジェクトに分類できる。本節ではこのように分類されたオブジェクトそれぞれの構成について述べる。

6.3.1 オブジェクトの階層構造

各オブジェクトの具体的な構成を述べる前に、そこで使われているオブジェクトの階層構造について説明する。

オブジェクトには、階層的な関係を設定することができる。これは親子関係と呼ばれ、親の動きは子に影響を及ぼすが、子の動きは親に影響を及ぼさず、その動きは親に対する相対的な動きとして記述されるような関係である。例えば、人間は肩を回せば、腕と手がそれにしたがって動く。腕だけを動かせば手はそれにしたがって動くが、肩はうごかない。また、手だけを動かしても肩や腕はその影響を受けず、動かない。このとき、肩は腕の親であり、腕は手の親であると言える。また、手は腕の子であり、腕は肩の子であるとも言える。

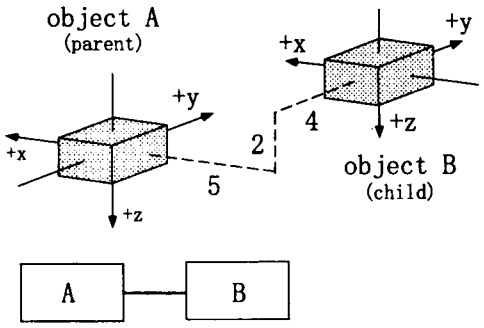


Figure 6.2: A parent and a child.

Figure 6.2では、A が B の親であるような関係をあらわしている。B の位置、姿勢は A に対する相対的な位置、姿勢で表されており、$(x, y, z, roll, pitch, yaw) = (5, 4, -2, 0, 0, 0)$ である。そのため、B が動かなくても A が移動、回転することにより絶対座標系に対しては動くことになる。A が回転する時、B の座標軸が回転する。例えば、B が Figure 6.3の左の図のような座標軸に設定されていた場合、その座標軸に対して A を y 軸について $-90[deg]$ 回転させれば、B は Figure 6.3の右の図のよ

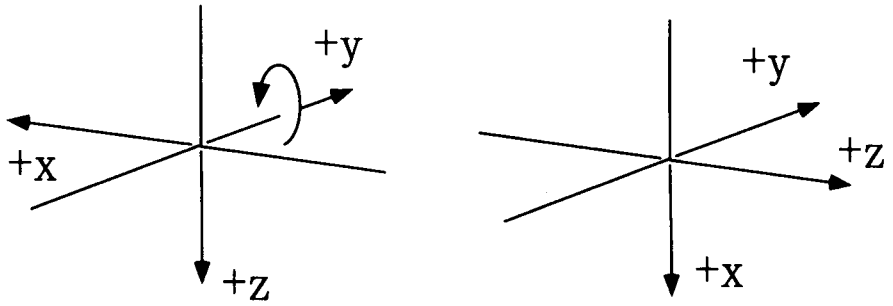


Figure 6.3: Axes of a child object.

うな座標軸に設定されたことになる。

本システムでは、オブジェクト同士に複数の親子関係を設定することにより、目的とするオブジェクトを構成し、また、目的とするシーンの生成を実現している。

6.3.2 人形オブジェクト

一つの人形オブジェクトの構成を、Figure 6.4に示す。上方にそれぞれの位置関係を、下方にオブジェクトの階層構造をを示す。

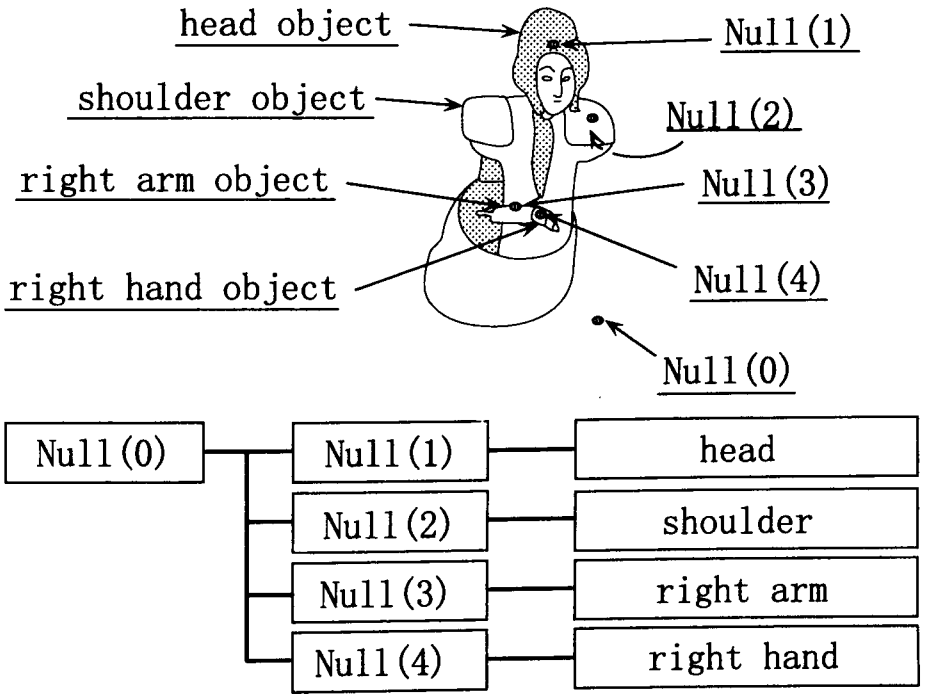


Figure 6.4: The structure of a puppet object.

頭、肩、右腕、右手オブジェクトに、それぞれ一つずつ Null オブジェクトが親として設定されている。Null オブジェクトは実際は目に見えないオブジェクトで、他のオブジェクトを制御するために用いている。また、Null オブジェクト同士を区別するため、Null(1),Null(2), のように番号を付加する。ここでは、頭オブジェクトに Null(1)、肩オブジェクトに Null(2)、右腕オブジェクトに Null(3)、右手オブジェクトに Null(4) をそれぞれ親として設定している。また、人形部位オブジェクトのレシーバの位置とその親である Null オブジェクトの位置は一致している。これにより、Null オブジェクトを回転させることによって、人形部位オブジェクトの位置を変えることなく座標軸の回転を行なうことができる。また、これらの Null オブジェクトの共通の親として Null(0) を設定している。これにより、Null(0) の位置、姿勢を決定することで、人形オブジェクト全体の位置、姿勢を決定することができる。

人形オブジェクトは実際には Figure 6.5のように表示される。

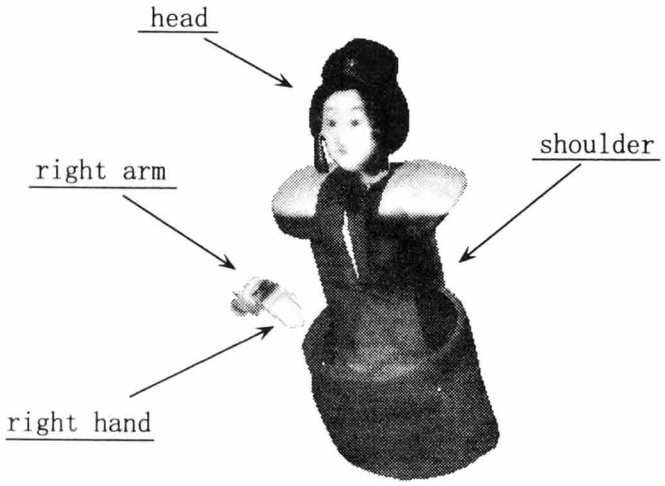


Figure 6.5: Figure of a puppet object.

頭オブジェクト、肩オブジェクト、右腕オブジェクト、右手オブジェクトはそれぞれ文楽人形の頭、肩、右腕、右手に相当するものである。これらの形状データは CyberModeler を用いて人形を計測することによって得られたデータを加工したものである。CyberModeler (三洋電機株式会社) はターンテーブルに乗せた物体を少しづつ回転させ、それを固定カメラで撮影し、得られた複数の静止画像から物体全周の形状を生成することができるシステムである。また、テクスチャデータは、実際の人形の写真を撮り、その画像を加工したものを用いている。

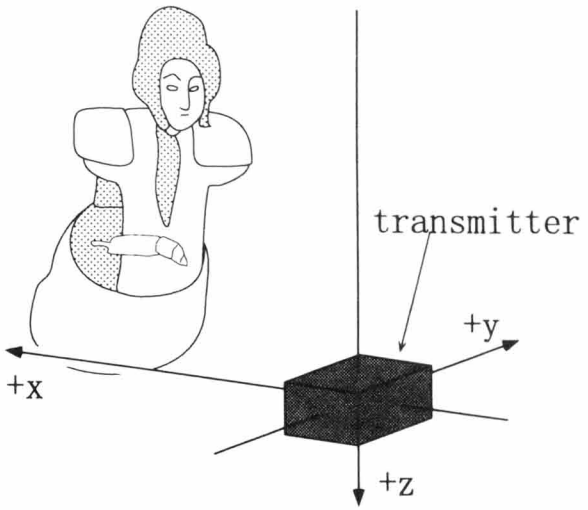


Figure 6.6: Setting of transmitter.

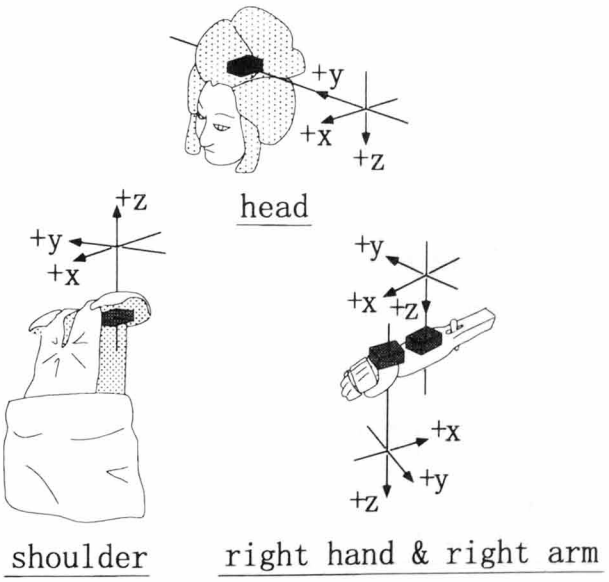


Figure 6.7: Setting of 4 receivers

人形部位オブジェクトの初期状態について考える。使用するデータは Figure 6.6に示した座標系における Figure6.7の Fastrak のレシーバ (第 2 章参照) の動きであるから、オブジェクトの初期状態は取り付けられたレシーバが $(x, y, z, roll, pitch, yaw) = (0, 0, 0, 0, 0, 0)$ となる状態である。各人形部位オブジェクトの初期状態を Figure 6.8 に示す。

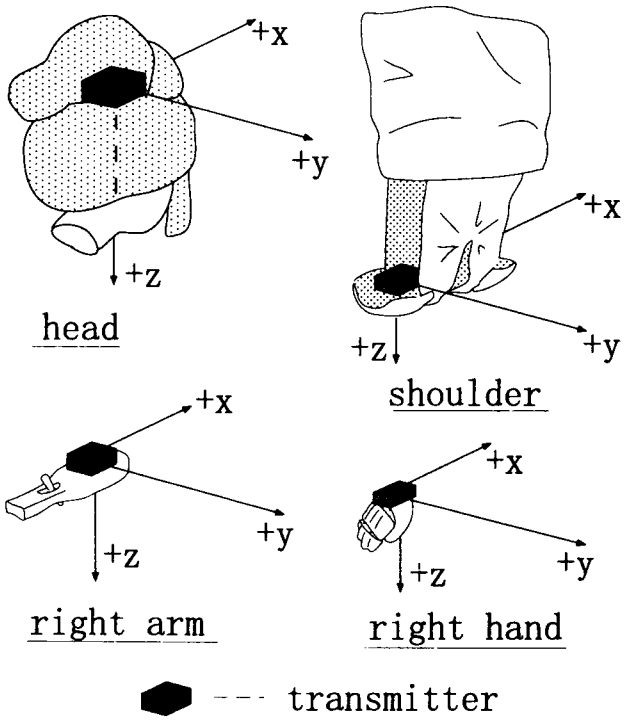


Figure 6.8: Initial setting of the puppet objects.

6.3.3 グラフオブジェクト

一つのグラフオブジェクトの構成を Figure 6.9に示す。上方にそれぞれの位置関係を、下方にオブジェクトの階層構造をを示す。

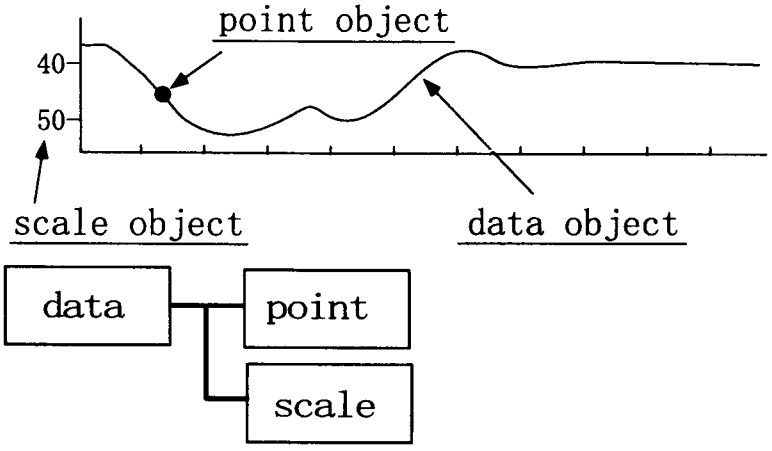


Figure 6.9: Structure of the graph object.

一つのデータオブジェクトを共通の親として、一つの点オブジェクト、複数の目盛りオブジェクトにより構成されている。

データオブジェクトは、時刻を横軸とし、自由度の値を縦軸としたグラフである。横軸には 2.0 秒毎に目盛が打ってあり、最右端の時刻は、$0.2 \times 106 = 21.2$ 秒 となる。点オブジェクトはある時刻における自由度の値を示す。

目盛オブジェクトはグラフの縦軸の目盛を１０刻みで示す。

グラフオブジェクトは実際には、Figure6.10のように表示される。

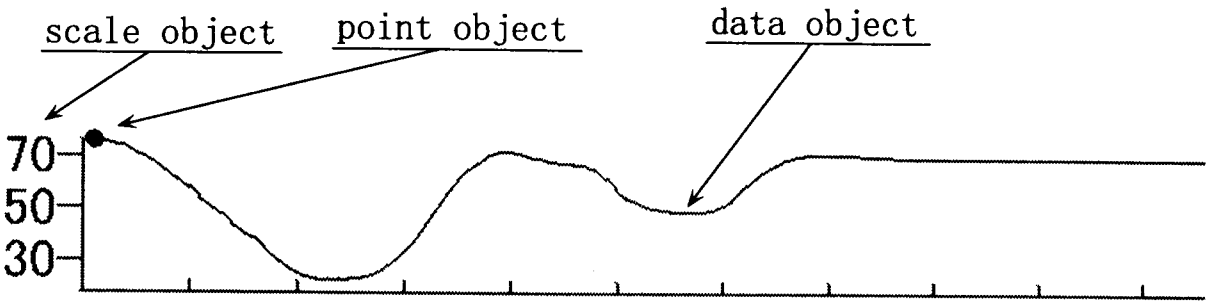


Figure 6.10: Figure of the graph object.

6.3.4 動作軸オブジェクト

一つの動作軸オブジェクトの構成を Figure6.11に示す。上方にそれぞれの位置関係を、下方にオブジェクトの階層構造を示す。

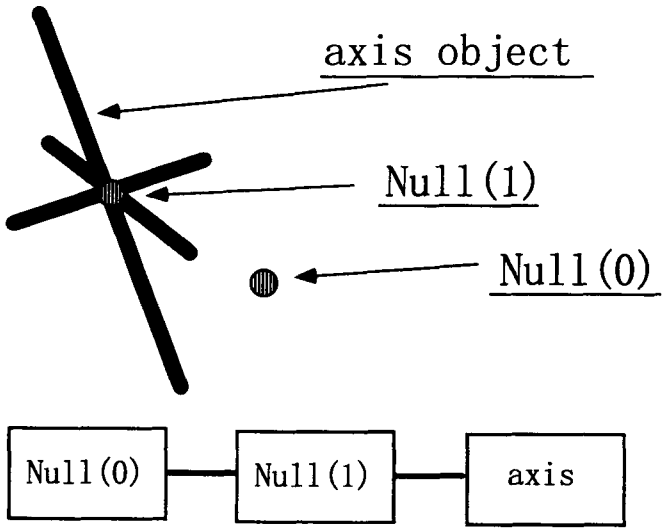


Figure 6.11: Structure of the motion axes object

軸オブジェクトの親として、一つの Null オブジェクト (Null(1)) があり、その親オブジェクトとしてもう一つの Null オブジェクト (Null(0)) がある。Null(1) の位置は、その子である軸オブジェクトの中心点と一致している。これにより、Null(1) を回転させることで、軸オブジェクトの位置を変えることなく座標軸を回転させることができる。また、Null(0) の位置や姿勢によって、動作軸オブジェクト全体としての位置や姿勢が決定される。

第 5 章で述べたように、動作軸とは、文楽人形の動作を代表するものである。動作解析においてこの動作軸に注目することにより、扱う自由度が 6 自由度に減少し、解析が容易になるというものである。このオブジェクトは直交する 3 つの軸を表す棒により、動作軸を表している。

動作軸オブジェクトは実際には Figure6.12のように表示される。

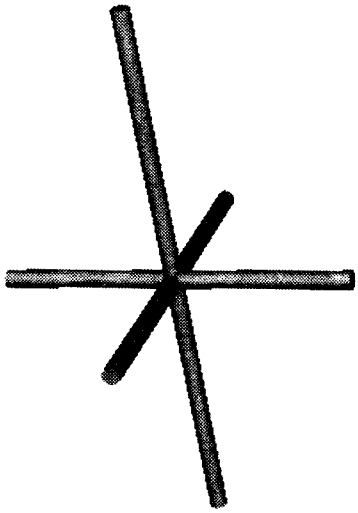


Figure 6.12: Figure of the motion axis object.

6.3.5 人形部位の動作軌跡オブジェクト

人形部位の動作軸オブジェクトの構成を Figure6.13に示す。上方にそれぞれの位置関係を、下方にオブジェクトの階層構造をを示す。

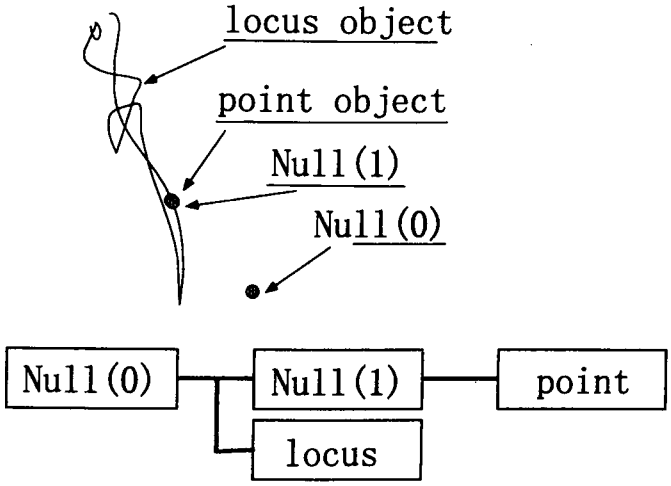


Figure 6.13: The object of the orbit of each part.

Null(0) を共通の親として Null(1) と軌跡オブジェクトが構成され、Null(1) を親として点オブジェクトが構成されている。Null(1) オブジェクトの位置は、その子であ

る点オブジェクトの中心と一致している。これにより、Null(1) を回転させることによって、点オブジェクトの位置を変えることなく座標軸を回転させることができる。

軌跡オブジェクトは、ある人形部位の始めから終りまで動作したとき、レシーバーが描く軌跡である。ポイントオブジェクトはある時刻における、人形部位に取り付けられたレシーバーの位置を示している。

人形部位の動作軌跡オブジェクトは実際は Figure6.14のように表示される。

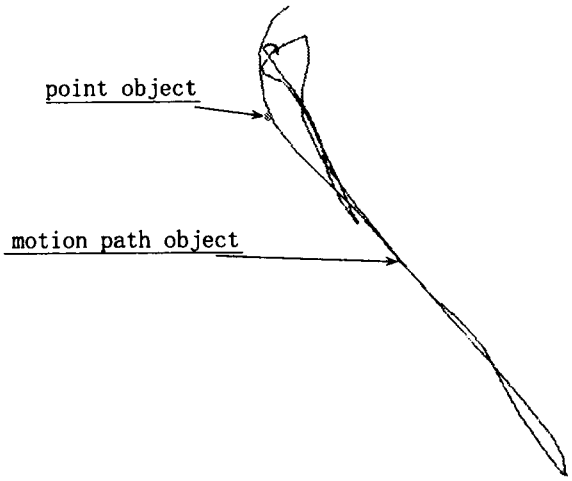


Figure 6.14: Figure of the orbit of each part.

6.4 シーンの構成

シーンは各オブジェクト、カメラ、ライトによって構成されている。各オブジェクトにデータを適用し、ライトによってオブジェクトを照らし、カメラの視点から見た画像がアニメーションになるのである。

本節では、これらの設定について述べる。これらの設定は、シーン生成プログラムにオプションパラメータを渡す時、その値によって決定される。また、生成されたシーンを LightWave3D に読み込んだ後、LightWave3D の機能を使って設定を変更することも可能である。

6.4.1 オブジェクトの設定

望みのシーンを得るためには、各オブジェクトを適切に設定する必要がある。6.3 節と同じようにオブジェクトを分類し、それぞれの設定について詳しく述べる。

人形オブジェクト

Figure 6.15に示すように、LightWave3D で扱われる座標系は左手座標系であり、回転は y 軸回転、x 軸回転、z 軸回転の順番である。しかし FASTRACK により計測されたデータは右手座標系で回転は、z 軸回転、y 軸回転、x 軸回転の順番である。

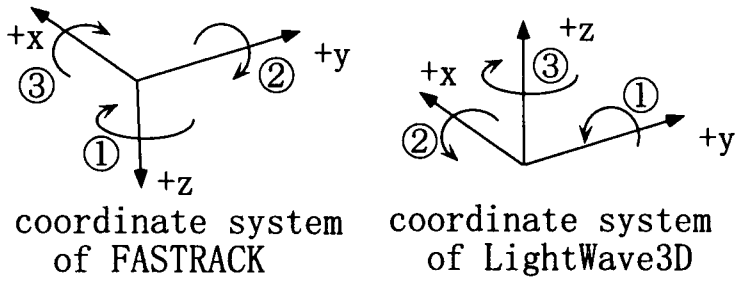


Figure 6.15: Coordinate systems.

このことから、FASTRACK で計測されたデータを LightWave3D で扱えるようにするには、右手座標系から左手座標系への変換、及び回転順序の変換を行なう必要がある。

まず、座標系の変換について説明する。Figure6.15を見ると分かる通り、z 軸を反転させ、x 軸についての回転、及び y 軸についての回転を反転させればよい。z 軸を反転させるために、z 軸データに-1 を乗じたものを用いる。また、x 軸についての回転、及び y 軸についての回転を反転させるために、pitch データ、yaw データそれぞれに-1 を乗じたものを用いる。

次に、回転順序の変換について説明する。各人形部位オブジェクトの親である Null オブジェクトを y 軸について 90[*deg*]、z 軸について 90[*deg*] 回転させる。これにより Figure6.16のように、人形部位の x 軸、y 軸、z 軸が、それぞれ絶対座標の y 軸、z 軸、x 軸に対応するようになり、絶対座標に対しての回転順序は z 軸回転、y 軸回転、x 軸回転の順番になる。

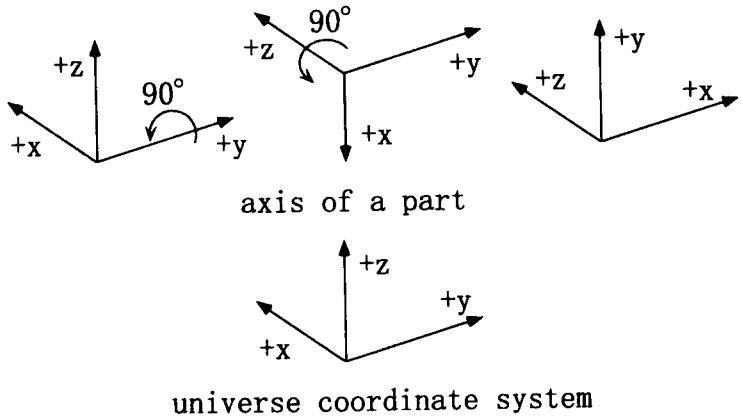


Figure 6.16: Coordinate systems.

このように、LightWave3D 上では人形部位オブジェクトの座標軸が回転し、Figure6.16 のようになるので、Figure6.8に示した初期状態の座標軸はそれぞれ回転後の、対応する座標軸に置き換えたものになる。Figure6.6に示した座標系と LightWave3D の絶対座標系を z 軸だけ反転させたものは同じであるから、LightWave3D 上での人形部位の初期状態は Figure6.17のようになる。

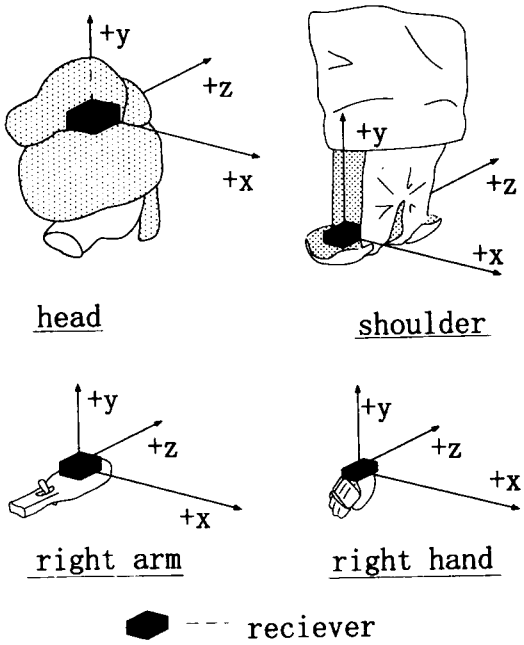


Figure 6.17: Initial setting of puppet objects in LightWave3D.

人形部位オブジェクトには 6 自由度のデータの内、姿勢データ $(roll, pitch, yaw)$

のみ適用する。位置データ (x, y, z) については、座標軸を回転させているので適用できない。そこで、親である Null オブジェクトに位置データを適用する。Null オブジェクトの座標軸は LightWave3D の絶対座標軸と一致しており、また人形部位オブジェクトのレシーバ取付け位置と一致しているので、人形部位オブジェクトを位置データ通りに動かすことができる。この Null オブジェクトが移動すれば、各人形部位オブジェクトも同じように移動する。人形部位オブジェクトが回転しても、親の Null オブジェクトは影響を受けない。これにより、人形部位オブジェクトは、位置、姿勢データの両方を適用したような動作をする。このように、人形部位オブジェクトと Null オブジェクトにそれぞれ姿勢データと位置データを別々に適用することで、シーンにおける人形オブジェクトの動作を生成している。

一つのシーンにおいて人形オブジェクトを 2 つ同時に表示する場合の設定を、Figure6.18 に示す。右側にそれぞれの位置関係を、左側にオブジェクトの階層構造をを示す。一つの Null オブジェクトを共通の親として、人形オブジェクトを 2 つ用いる。個々の人形オブジェクトについての設定は、一つだけ表示する場合と同じである。

カメラの見る角度によって人形が重なってしまわないように、二つの人形オブジェクトの位置は、カメラがどの角度に設定されても常に左右にならんで表示されるようにように設定される。

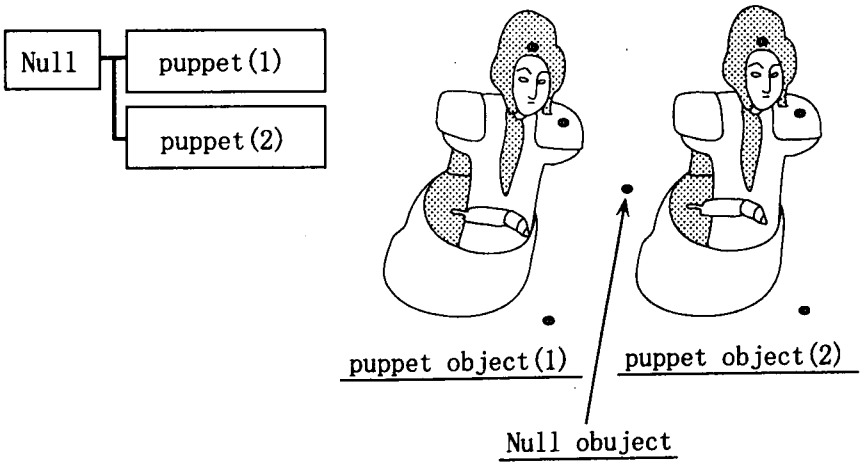


Figure 6.18: Construction of 2 puppets in a scene.

グラフオブジェクト

シーンにおいて、グラフオブジェクトの親として、一つの Null オブジェクト (Null(1)) が設定され、その親としてさらに一つの Null オブジェクト (Null(0)) が設定さ

れる。グラフオブジェクトを複数表示する場合には、Null(1) を親としてグラフオブジェクトが追加されるように設定される。シーンにおけるグラフオブジェクトの設定を Figure6.19に示す。

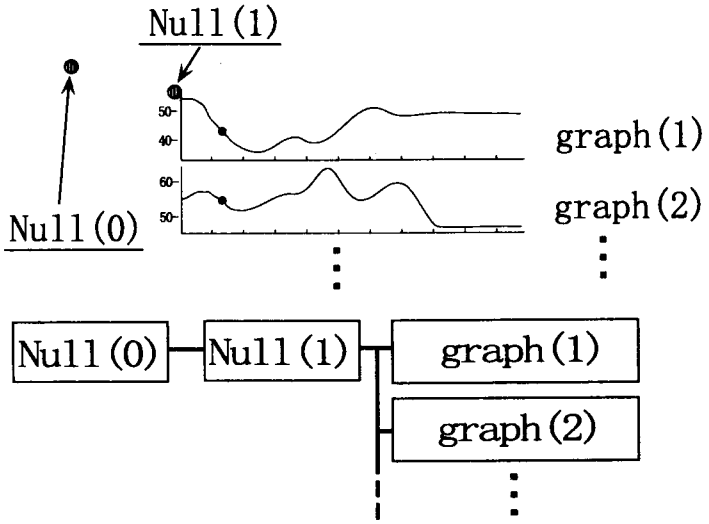


Figure 6.19: Setting of the graph objects in a scene.

Null(1) の位置及び姿勢はカメラのそれと同じである。これにより、常にカメラの正面にグラフを表示することができる。

使用するデータが一つのシーンにおいて、グラフの表示パターンは、

1. 全ての人形部位について 6 自由度すべて表示する (２４グラフ)

2. 一つの人形部位について 6 自由度全て表示する (6 グラフ)

3. 全ての人形部位について 1 自由度だけ表示する (4 グラフ)

4. 一つの人形部位について 1 自由度だけ表示する (1 グラフ)

の 4 パターンある。

1 は、１２個ずつ縦に並べて、画面全体にグラフオブジェクトが並べられる。2,3 は、画面の右半分にグラフオブジェクトが並べられる。4 は、画面の下方にグラフが配置される。

使用するデータが二つのシーンにおいて、グラフの表示パターンは、

1. 全ての人形部位について 6 自由度すべて表示する (４８グラフ)

2. 一つの人形部位について6自由度全て表示する(12グラフ)

3. 全ての人形部位について1自由度だけ表示する(8グラフ)

4. 一つの人形部位について1自由度だけ表示する(2グラフ)

の4パターンある。

1は、12個ずつ縦に並べて、画面全体いっぱいにグラフオブジェクトが並べられる。2は、6個ずつ縦に並べて、画面の下方にグラフオブジェクトが並べられる。3は、4個ずつ縦に並べて、画面の下方にグラフオブジェクトが並べられる。4は、横に並べて、画面の下方にグラフオブジェクトが配置される。

動作軸オブジェクト

動作軸オブジェクトの設定は、人形オブジェクトの一部位の設定と同様である。ただし、適用するデータが動作軸用のデータであるという点が異なる。

Figure6.20のように、6.3.2で述べた人形オブジェクトの最上階層に位置するNullオブジェクトと、6.3.4で述べた動作軸オブジェクトの最上階層に位置するNullオブジェクトが一致するように設定される。

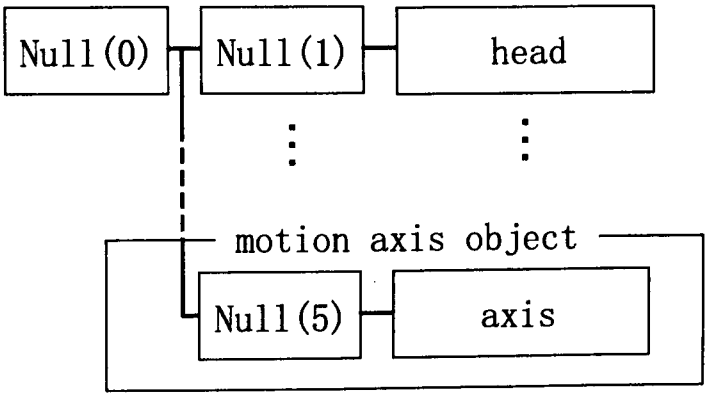


Figure 6.20: Setting of the motion axes objects in a scnene.

軸オブジェクトの親であるNullオブジェクトをy軸について90[*deg*]、z軸について90[*deg*]回転させる。これにより、6.4.1で述べたように、軸オブジェクトの座標軸が回転し、絶対座標に対しての回転順序が変わる。そして動作軸用のデータの内、位置についてはNullオブジェクトに、姿勢については軸オブジェクトに適用する。

このように親子関係にある2つのオブジェクトを用いて、シーンにおける動作軸オブジェクトの動作が設定されている。

人形部位の動作軌跡オブジェクト

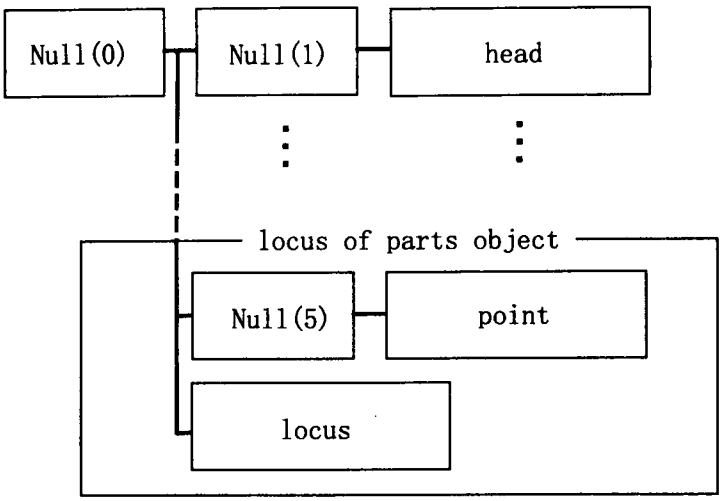


Figure 6.21: Setting of the objects of the orbits in a scnene.

人形部位の動作軌跡オブジェクトは、Figure6.21のように、6.3.5で述べた動作軸オブジェクトの最上階層に位置する Null オブジェクトと、6.3.2で述べた人形オブジェクトの最上階層に位置する Null オブジェクトが一致するように設定する。これにより、点オブジェクトは各人形部位オブジェクトと同じ動作をすることになり、軌跡オブジェクトはその動作の軌跡と一致する。すなわち、人形オブジェクトの各部位の動作軌跡と一致する。

ただし、点オブジェクトは人形オブジェクトが表示されない時にのみ表示される。

6.4.2 カメラとライトの設定

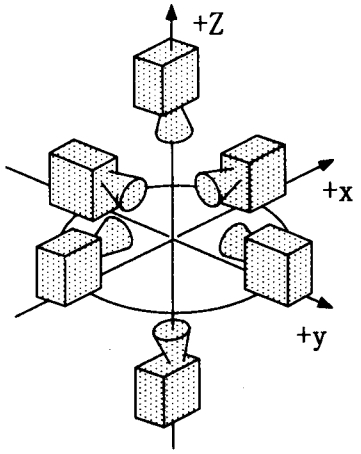


Figure 6.22: Setting of the camera in a scnene.

Figure6.22のように、カメラはxy平面上を、原点を中心とした円周上、及びz軸上の2点に設定できる。また、カメラはいつでも原点を向いている。

人形を真上から見下ろした画像を得るには、カメラがz軸上にあり、しかもz座標が正の値をとるように設定する。人形を真下から見上げた画像を得るには、カメラがz軸上にあり、しかもz座標が負の値をとるように設定する。また、人形を正面、横、背後などから見た画像を得たいときには、xy平面にある円周上の位置を、カメラがy軸上の正の値をとる位置にある時を0[*deg*]として、回転角で設定する。

シーンでは、カメラの親オブジェクトとして、一つのNullオブジェクトを設定する。このNullオブジェクトを回転させることによって、カメラの位置を任意の円周上に設定することができる。このNullオブジェクトは、6.4.1で述べたグラフオブジェクトの設定において最上階層に位置するNullオブジェクトの親としても設定される。

ライトは、一つのNullオブジェクトを親として設定する。このNullオブジェクトは、6.4.1で述べたグラフオブジェクトの設定において最上階層に位置するNullオブジェクトと一致するように設定される。これにより、ライトはカメラと同じ位置に設定され、常に視線方向の光線が各オブジェクトに当たっていることになる。

シーンにおけるグラフオブジェクト、カメラ、及びライトの階層構造の設定をFigure6.23に示す。

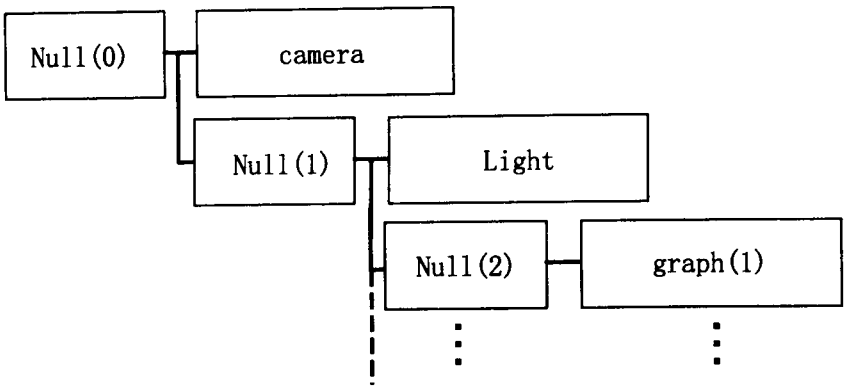


Figure 6.23: Setting of the graph object , the camera and the light in a scnene.

6.5 マルチモーダル表示システム

本システムでは、人形オブジェクト、グラフオブジェクト、動作軸オブジェクト、人形部位の動作軌跡オブジェクトうちのいくつか、または全部を同時に表示し、お互いを関連付けてとらえられることができる。

それぞれのオブジェクトを一つだけ選んで表示した場合に得られる情報は以下の通りである。

- 人形オブジェクト

  適用したデータによって生成された人形の動作を見ることができる。

- グラフオブジェクト

  適用したデータのグラフの波形、及びある時刻においてどの値をとっているかを見ることができる。

- 動作軸オブジェクト

  適用したデータによって生成された動作軸の動作を見ることができる。

- 人形部位の動作軌跡オブジェクト

  適用したデータによって生成された人形の各部位の動作軌跡を見ることができる。

本章では、表示するオブジェクトの組合せによって、どのように関連づけてとらえることができるかを示す。

6.5.1 一つの動作時系列データに基づいて表示する場合

- 人形とグラフ

  人形オブジェクトとグラフオブジェクトを同時に表示することにより、人形の動作の変化とグラフの波形の変化を、対応させてとらえることができる。

- 人形と動作軸

  人形オブジェクトと動作軸オブジェクトを同時に表示することにより、動作軸に対する各人形部位の相対的な動きをとらえることができる。

- 人形と人形部位の動作軌跡

  人形オブジェクトと人形部位の動作軌跡オブジェクトを同時に表示することにより、ある時刻の人形の各部位の位置が、動作全体の流れ中でどのような位置にあるのかをとらえることができる。

- グラフと動作軸

  グラフオブジェクトと動作軸オブジェクトを同時に表示することにより、動作軸の動きの変化とグラフの波形の変化の対応をとらえることができる。

- グラフと人形部位の動作軌跡

  グラフオブジェクトと人形部位の動作軌跡オブジェクトを同時に表示することにより、グラフの波形の変化と、軌跡の形状の変化や、グラフの値の変化と、軌跡上の位置の変化を対応づけてとらえることができる。

- 動作軸と人形部位の動作軌跡

  動作軸オブジェクトと人形部位の動作軌跡オブジェクトを同時に表示することにより、動作軸の動きの変化と、各部位の位置の変化を対応づけてとらえることができる。

- 人形とグラフと動作軸

  人形オブジェクトとグラフオブジェクトと動作軸オブジェクトを同時に表示することにより、動作軸に対する各人形部位の相対的な動作と、それに対応するグラフの波形や値の変化、また人形全体としての動作の変化を同時にとらえることができる。

- 人形とグラフと人形部位の動作軌跡

  人形オブジェクトとグラフオブジェクトと人形部位の動作軌跡オブジェクトを同時に表示することにより、各人形部位の動作全体の中での位置と、それに対応するグラフ上での位置をとらえることができる。また、人形の動作の変化と、グラフの波形の変化、及び動作軌跡の形状の変化を同時にとらえることができる。

- グラフと動作軸と人形部位の動作軌跡

  グラフオブジェクトと動作軸オブジェクトと人形部位の動作軌跡オブジェクトを同時に表示することにより、グラフ上での値の変化とそれに対応する人形部

位の軌跡上での位置の変化、また、動作軸の動きと、それに対応する人形部位の軌跡の動き、及びグラフ上での値の変化を同時にとらえることができる。

- 人形とグラフと動作軸と人形部位の動作軌跡

  人形オブジェクトとグラフオブジェクトと動作軸オブジェクトと人形部位の動作軌跡オブジェクトを同時に表示することにより、以上に述べた各オブジェクトについての関係をすべて同時にとらえることができる。

6.5.2 二つの動作時系列データに基づいて表示する場合

異なる二つの動作時系列データを適用された二つの人形オブジェクト、グラフオブジェクト、動作軸オブジェクト、人形部位の動作軌跡オブジェクトをそれぞれ同時に表示することができる。各オブジェクトの表示組合せは、前節で述べたものと同じである。それぞれの組合せによってとらえることのできる関係は、前節でそれぞれ述べたものに加え、異なるデータを適用された各オブジェクトとの比較ができる。

機能や情緒表現の違いによる各オブジェクトの違いを、それぞれ同時に比較することができる。また、計測したデータをある手法を用いて加工したものと、計測したデータそのままのものについて、各オブジェクトの動きや形状を同時に表示し比較することにより、その手法の有効性について多くの情報を元に、しかもその情報同士の関連も考慮しながら評価できる。

6.6 動作表示の例

「感情をこめないでお茶を渡す」という動作についてのデータを用いて、人形オブジェクト、グラフオブジェクト、動作軸オブジェクト、人形部位の動作軌跡オブジェクトをそれぞれ組み合わせてできる全ての表示パターンについて、実際に表示した画像を示す。ただしグラフオブジェクトは、頭の六自由度についてのグラフを表示するように設定されている。

また、「感情をこめないでお茶を渡す」という動作についてのデータと、「敵意を持ってお茶を渡す」という動作についてのデータとの二つのデータを用いて、両方のデータについての人形オブジェクト、グラフ、動作軸、人形部位の動作軌跡の全てを同時に表示させた画像を次ページ以降に幾つか示す。

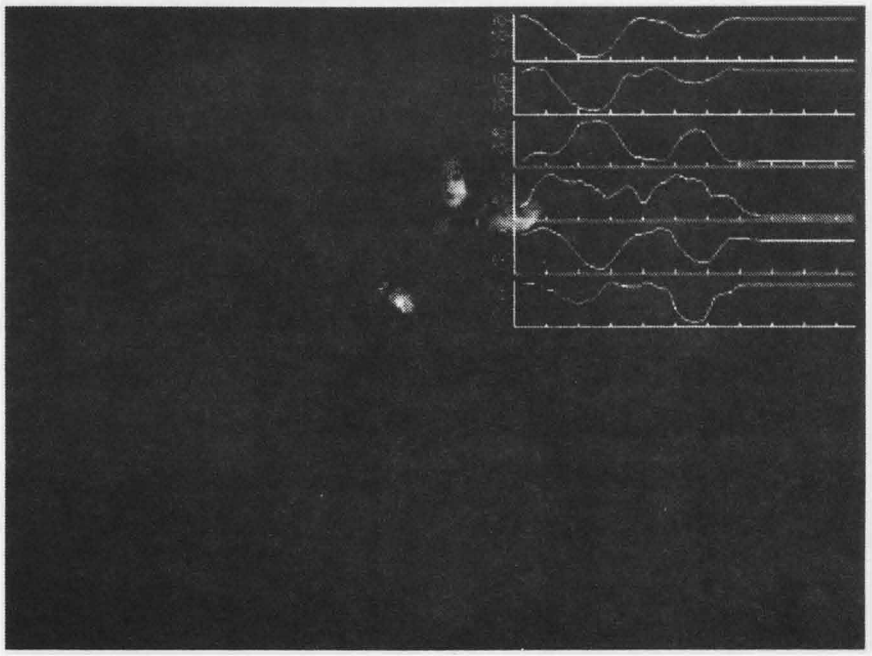


Figure 6.24: The puppet and the graphs.

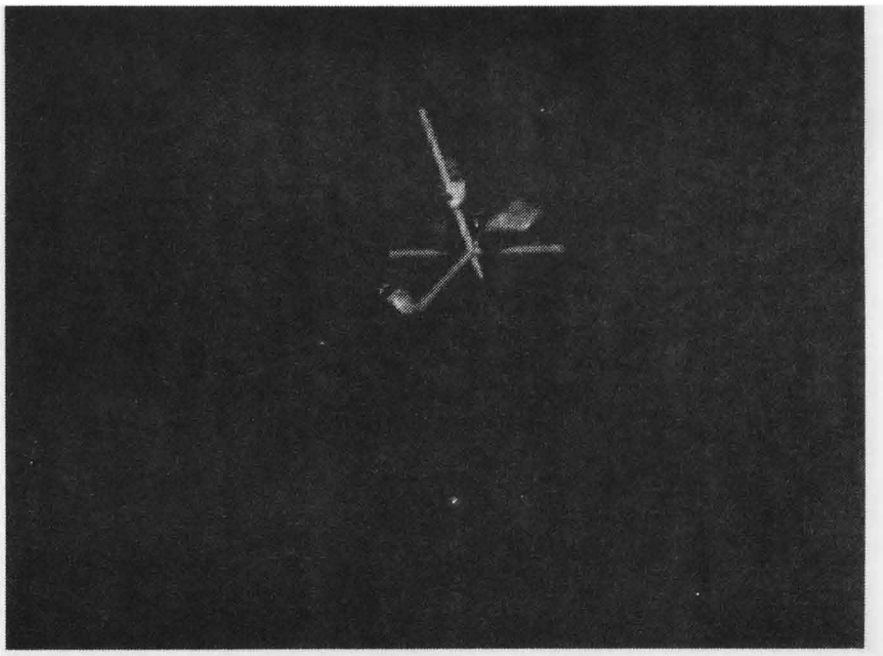


Figure 6.25: The puppet and the motion axes.

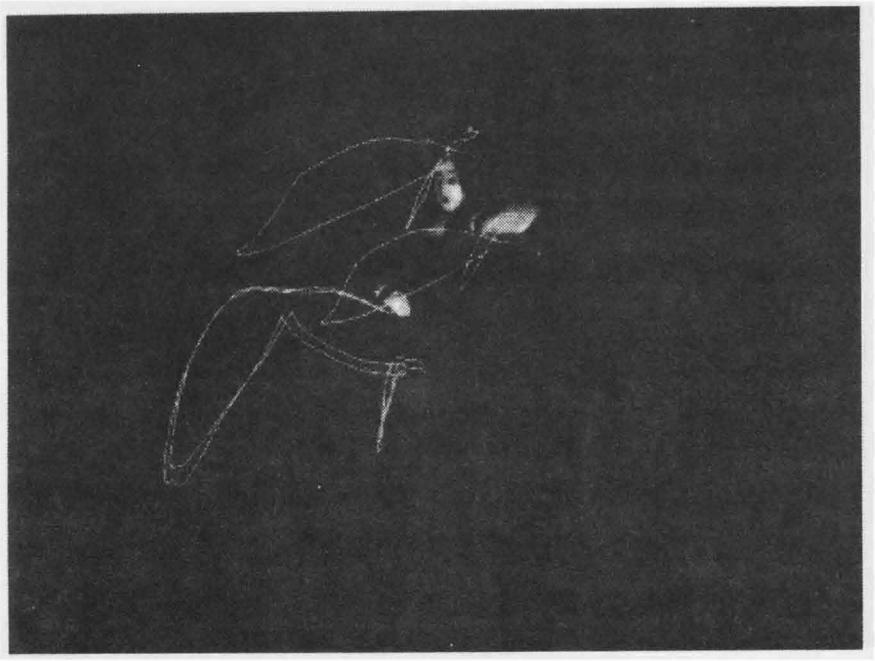


Figure 6.26: The puppet and the orbits.

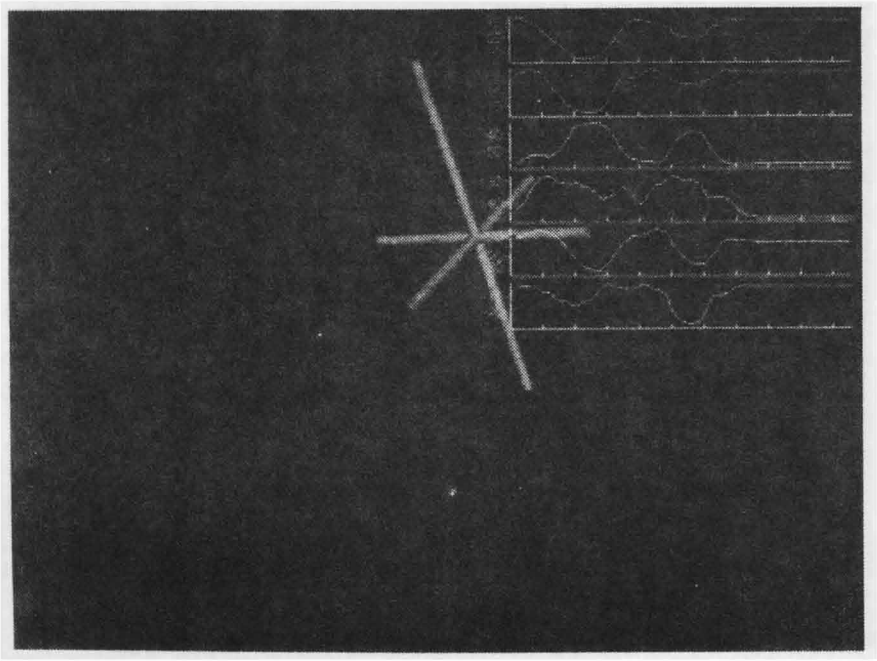


Figure 6.27: The graphs and the motion axes.

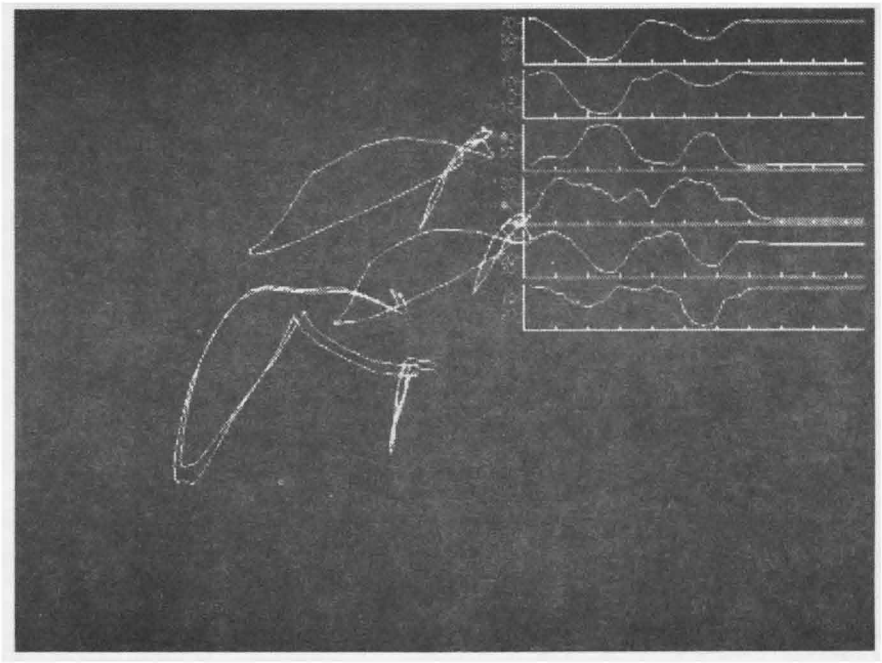


Figure 6.28: The graph and the orbits.

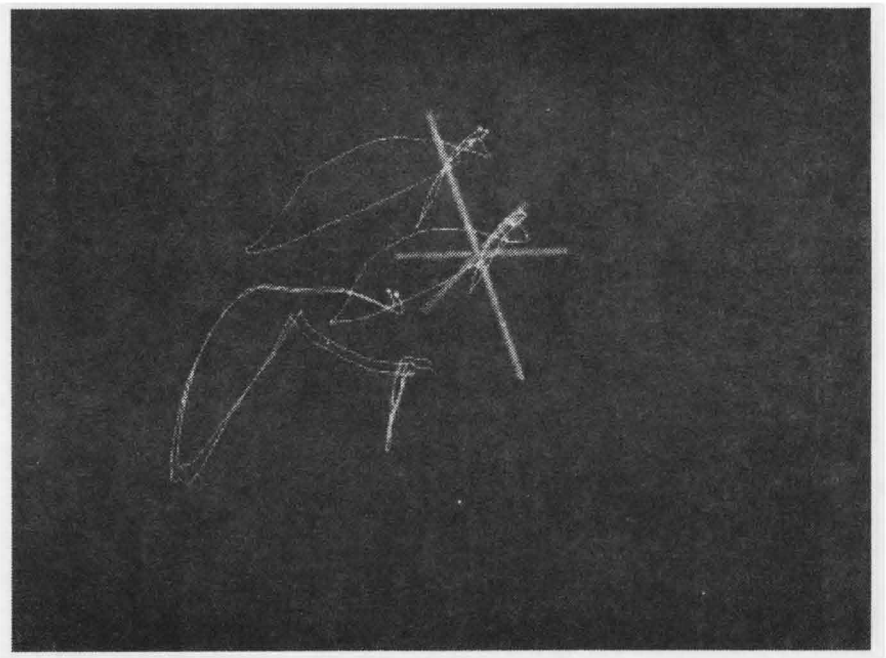


Figure 6.29: The motion axes and the orbits.

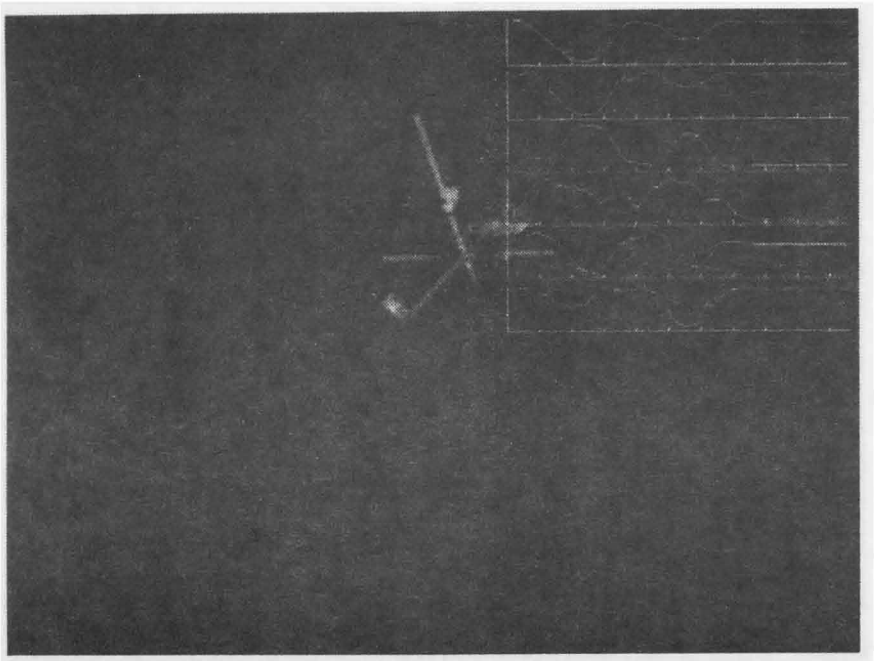


Figure 6.30: Puppet and graph and motion axis

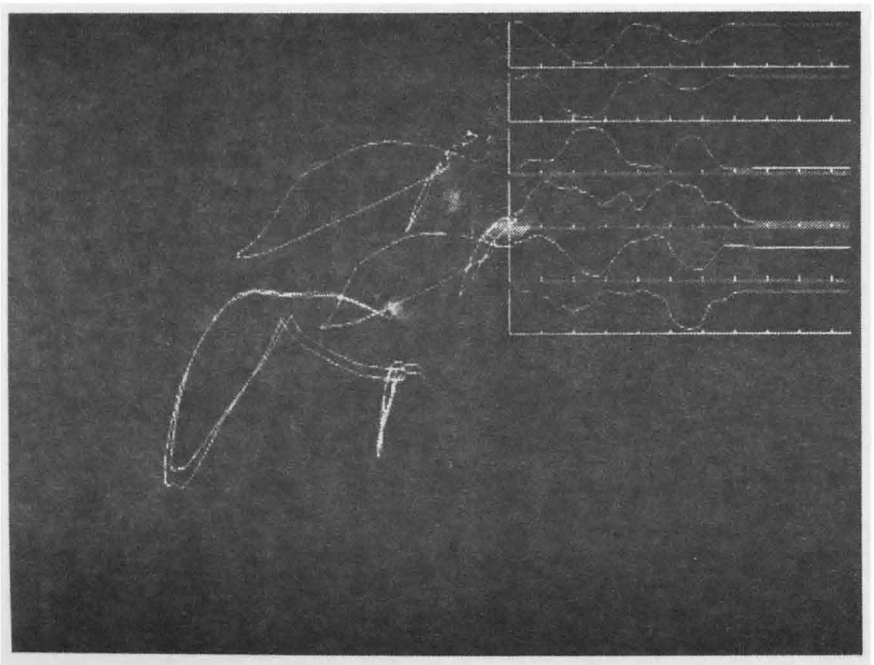


Figure 6.31: The puppet, the graphs and the orbits.

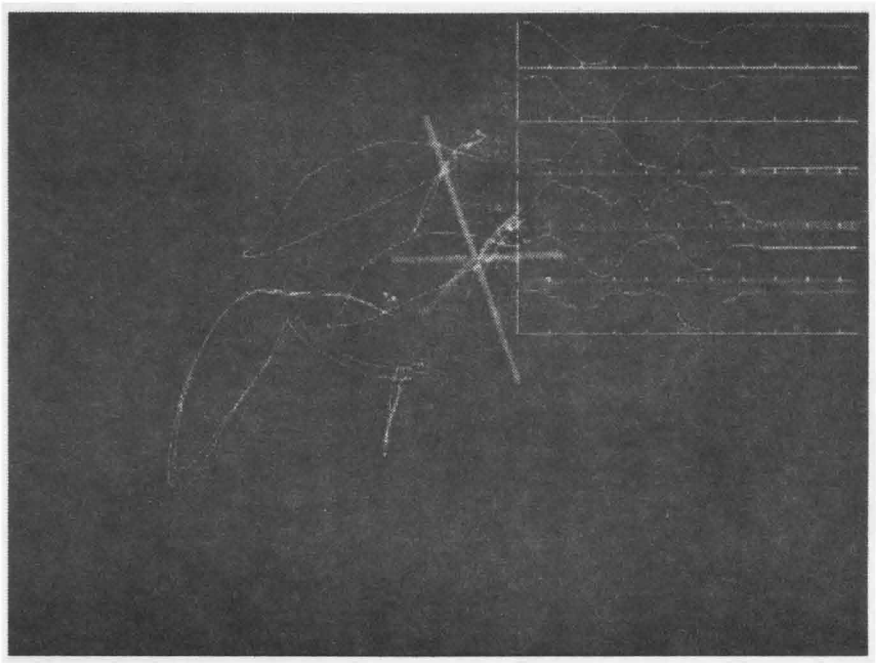


Figure 6.32: The graphs, the motion axes and the orbits.

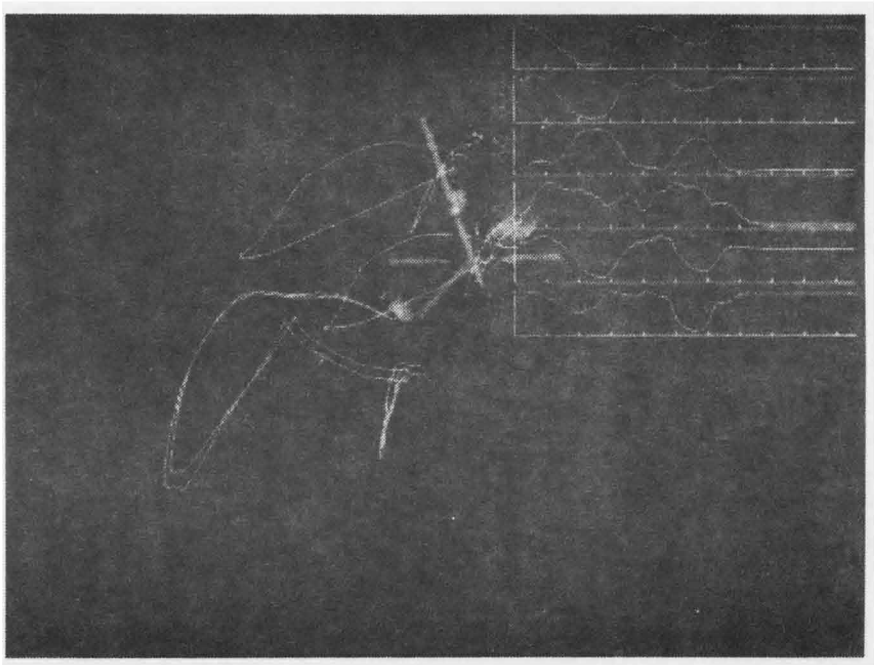


Figure 6.33: The puppet, the graphs, the motion axes and the orbits.

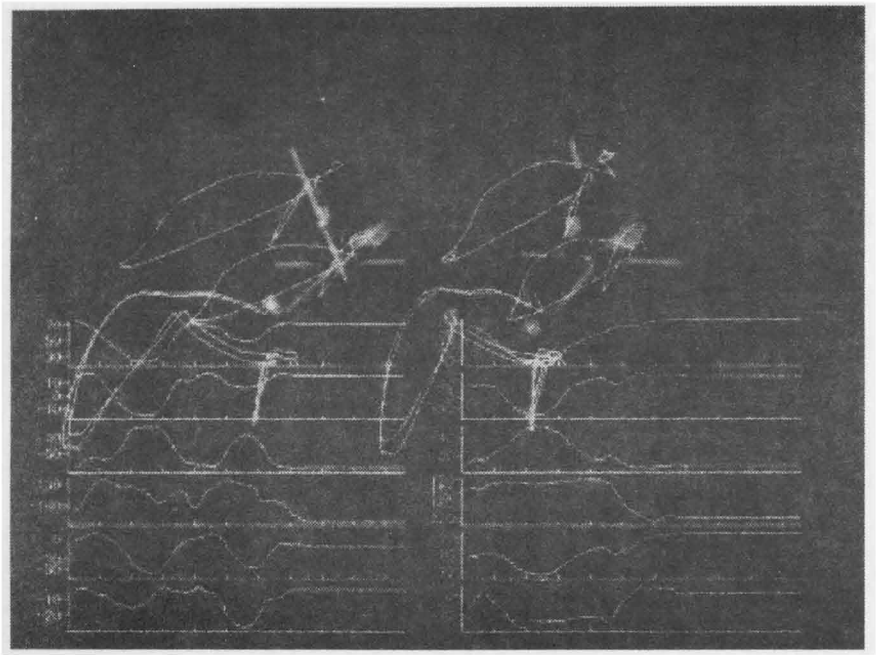


Figure 6.34: The 2 puppets, the graphs the motion axes and the orbits.

6.7 結言

本章では、文楽人形の演技動作の解析の支援及び解析結果のシミュレーションを、コンピュータグラフィクス・アニメーションを用いて行なうことを目的として開発したシステムを紹介した。

本システムでアニメーションを生成するまでの流れは次のようになる。

1. まず、人形オブジェクト、グラフオブジェクト、動作軸オブジェクト、人形部位の動作軌跡オブジェクトをそれぞれ用意する。

2. シーン生成プログラムによりそれぞれのオブジェクトに計測されたデータを適用し、カメラやライトの設定をすることによりシーンを生成する。

3. シーンを LightWave3D で読み込み、アニメーションを生成する。

用意したオブジェクトのいくつか、または全部を同時に表示させ、お互いの関連をとらえることができる。また、異なる二つのデータを適用したオブジェクトを同時に表示することで、機能、情緒の異なる人形の動作やグラフを比較することができる。これによりそれぞれを同時に考慮しながら効果的に解析を行なうことができる。また解析結果に基づいたシミュレーションにおいても、データや動作について比較や参照が行なえるため、結果のシミュレーションに対する評価をしやすくなる。

今後の課題としては、現在の着物を省いた人形オブジェクトを着物を着た人形オブジェクトにして、シミュレーションを行なうことが挙げられる。実際の文楽人形は着物を着た状態で演技動作を行なうため、着物の動きにも情緒表現が含まれていると考えられるからである。また、動作対象の表示や、動作譜の表示など、解析やシミュレーションに有効であると思われるものをさらに追加したシステムへ改良することも計画している。

第 7 章　結論

本論文では、日本の伝統的な情緒表現が典型的に表現されている文楽人形の演技動作を解析し、人間型の動作に特有なデータ処理方法・主要な情報の抽出方法などを考案し、それらに基づいて人間型ロボットの動作設計方法の幾つかの手法を提案した。各章の結果を要約すると下記のようになる。

第 3 章では、文楽人形の動作時系列の確率的な揺らぎを補正して標準化を行うために、同じ条件のもとで複数回観測された時系列たちから、同一の事象に対応する時刻たちを求めて対応付ける方法を提案した。鍵となる考え方は、時系列たちのグラフが局所的に似ている時刻を対応させるという点であり、グラフの局所的な形を数量的に表す物として、時系列の各時刻における種々のスケールのウェーブレット係数に着目した。つまりウェーブレット係数が近い時刻どうしを対応する時刻と考えて、時系列の標準化を行う手法を提案した。比較するウェーブレット係数のスケールを適切に選ぶことによって、高周波雑音の影響を受けない時系列の標準化が可能である。

これら時系列の標準化手法を、文楽人形の動作時系列に対して適用し、標準化された動作時系列のグラフを検討し、それらに基づいてコンピュータ・グラフィックス上に生成された動作を観察することにより、上記の手法の問題点について考察し、標準化手法を時間区間・時間発展・人形の部位どうしの関連を考慮するように改良した。

また上記の標準化手法を用いて、時刻と状態の双方が確率的に揺らいでいる複数の時系列から平均の時系列を求める方法を提案した。そのようにして平均化された時系列を参照信号として標準化を行い、良好な結果を得た。

提案された時系列標準化や平均時系列の計算手法は、文楽人形の演技動作の解析の今後の研究に有効であるだけでなく、異なる種類の複数の信号から対応する時刻を求める必要のあるセンサフュージョンや音声信号・ロボットへの教示信号などの

応用分野の信号処理においても有用である。

第４章では、機能・情緒・確率的揺らぎの線形構造に基づいて、文楽人形の演技動作を解析し、人間型のアミューズメント・ロボットの動作生成に使用する動作時系列の情緒を、強調させたり減衰させたりして加工するシステマティックな手法を提案した。様々な機能と様々な情緒の文楽人形の動作動作時系列から、機能の要素と情緒の要素を推定・抽出し、情緒の要素を強調させたり減衰させたりして、新しい動作を計算することができる。

第５章では、人間型アミューズメント・ロボットの動作を設計するにあたって、身体の各部位の動きを　ばらばらに設計するのではなく、それらの総体としての整合性のある動作を設計する技法を求めて、文楽人形の動作時系列を解析した。

測定された文楽人形の動作時系列を、（１）人形の慣性主軸の動きという主要な部分と（２）慣性主軸に対する相対動作という副次的な部分　の２つに分解する事を考えた。

そして、動作の主要な部分を表している慣性主軸の動きが、動作の情緒の違いによって変化する様子を、時間伸縮と振幅値の増幅･減衰によって説明できる事を示した。

さらに時間伸縮と振幅値の増幅・減衰を適切に設計する事で、情緒のない動作の軸の動きを加工して、情緒のある動作の軸の動きを生成する事を提案した。動作の主要部分を設計する一つの技法を示した事になる。位相要素の対応を示す簡単なグラフの作成と振幅を増幅・減衰する数個のゲインの設定だけで、動作軸の動きの設計を行える事が、本手法の特色である。

上記に研究された動作解析および設計の手法の有効性を検証するために、文楽人形や動作軸などの動きを CG で視覚的に表示しながら、同時に動作時系列や動作軸のグラフなども確認できる マルチモーダル動作表示システムを開発したので、その概要を第６章にて紹介した。

以上のように人間型の動作の設計方法を求めて、文楽人形の演技動作の解析方法を工夫して来た。人形の動作時系列（各時刻の各部分の位置と姿勢の値）という数

値データを出発点とし、最初は数値に依存した動作の解析・設計方法を模索していたが、やがて第5章で提案された動作軸のように、細かい具体的な数値から余分なものをそぎ落として、動作の本質的な表現を求めて行く方向で進んで来たし、今後もこの方針に従って研究を進めるべきであると考える。

人間型の動作を通じて、機能を実現したり情緒的な精神的な意図を伝達するにあたって、身体の各部位の各時刻の位置と姿勢という膨大な量の数値が本質なのではなく、それらからもっと簡略化された動作の本質だけを保存できる表現方法を、さらに深く究明すべきであると考える。その表現方法は、人間型の身体の動きを表現するにあたって汎用的なものであり、人形や人間型ロボットだけでなく人間やアニメーションの人物など人間型のものに共通なものであり、身体の形や大きさなどの個性に依存しないものが望ましいと考える。

このような人間型動作の記述・解析方法の研究は、人間型ロボットの動作設計にとって必要であるだけでなく、ビデオに記録された舞踊や文楽の身体の動きを3次元データに変換して、それらの動きの設計技法の解析や文化的な考察による分類などを客観的に行う比較舞踊学などの分野においても有用であると考える。

このような動きの科学を客観的な考察に基づいて進めるにあたって、本論文で報告された内容は、基礎的な方向付けを提案するものである。

謝辞

本研究を遂行するにあたり、小生の素養を絶えず考慮されながら適切な指導を頂きました高森年教授（神戸大学工学部）に心より感謝いたします。数値データだけでは捉えきれない文楽人形の動きの側面について、貴重で適切な指導を頂きました山田和人教授（同志社大学文学部）に深く御礼申し上げます。親切に相談に乗ってくださり、文楽人形の様々な動きを操作してくださいました吉田玉女氏 (文楽協会・人形遣い) に深く感謝いたします。本論文をまとめるにあたりまして、ヒューマノイドの動作認識という立場から貴重な御指導を頂きました上原邦昭教授（神戸大学工学部）賀谷信幸教授（神戸大学工学部）に深謝いたします。

また、信号処理や制御理論といった理論よりの研究から出発した小生に対して、工学の実用的な側面を御指導くださいました田所諭助教授（神戸大学工学部）
木村一郎教授（大阪電気通信大学）小林滋助教授（神戸高専）前田潔助教授（海技大学）に深謝いたします。研究室の研究基盤を整えるにあたり、様々な側面でお世話になりました山崎義治技官（神戸大学工学部）に深謝いたします。共同研究者である辻誠彦氏（関西電力）中坊保則氏（シャープ）北川智広氏（NTT 関西移動通信網）西澤俊君（神戸大学自然科学研究科）佐藤清昭君（神戸大学自然科学研究科）古田智君（神戸大学工学部）に感謝いたします。

信号処理や制御理論の研究を始めるにあたりまして、初心者の小生を導いてくださいました西尾真喜子先生（神戸大学理学部名誉教授）樋口保成教授 (神戸大学理学部）故砂原善文先生（京都工芸繊維大学名誉教授）大住晃教授（京都工芸繊維大学）亀島鉱二教授（大阪工業大学）に深謝いたします。

また、本研究を遂行するにあたりまして、比較舞踊学会や Labanotation 研究会における舞踊の先生方より貴重なアドバイスを頂いたことに感謝するとともに、このような先生方と議論できる素養を与えてくださいました中学時代以来のソルフェージュ・ピアノの先生方に感謝いたします。

最後に、工学者としての自覚を小生に促してくださいました、視覚障害を持たれる情報処理技術者に深謝するとともに、大阪府立盲学校情報処理専攻における研究会にて、これらの方々と、情報処理・情報工学の文献の電子化・音声化・点字化にも力を注ぐ所存です。

参考文献

[1] 井口信洋, "アミューズメント・ロボットの設計思想", 日本ロボット学会誌 8 巻 3 号, pp.315-319, 1990

[2] 池浦良淳, 猪岡光:"ダンシングロボット," 日本ロボット学会誌, Vol. 14, No. 2, pp. 200-203, 1996

[3] 松浦壮一, 藤城智子, 山本一道, 井口信洋:"ダンスロボットの開発とその主観的評価," 日本ロボット学会誌, Vol. 9, No. 2, pp. 177-183, 1991

[4] 藤原春夫:"阿波踊りロボットの制御," 計測と制御, Vol. 31, No. 12, pp. 1264-1270, 1992

[5] KOKORO NEWS, no.54, 1995

[6] わらび座デジタル・アート・ファクトリのホームページ: http://www.warabi.or.jp/daf/index.html

[7] 八村広三郎:"人文科学とイメージ処理- 情報科学の新しい展開 -", Ritsumeikan University symposium on image processing and analysis, Biwako-Kusatsu campus, Ritsumeikan University, pp.53-68, January 14, 1999

[8] 青木繁・山田和人 構成、園田学園女子大学 近松研究所 編集："吉永孝雄の私説 昭和の文楽"、和泉書院 1995

[9] 大西重孝："文楽人形の演出"、大西重孝著作刊行会 発行、中尾松泉発売 (1974)

[10] 竹本住大夫・豊澤富助："文楽カセット 文楽入門"、週間カセット出版、1986

[11] Tomoich Takahasi, "Time normalization and analysis method in robot programming from human demonstration data", Proc. of the 1996 IEEE on robotics and automation, pp.37-42, 1996

[12] 中川聖一, “確率モデルによる音声認識,” 電子情報通信学会,(1988) pp.92-94.

[13] 紙谷卓之, 山本方子, 丸谷洋二, “DP マッチングを用いた困り行動の認識,” Human Interface,Vol.12,pp.475-482,(1997).

[14] 岡部靖憲:“非線型予測理論と因果解析,” システム/制御/情報, Vol. 33, No. 9, pp. 478-487, 1989

[15] Y.Okabe:“Langevin equations and causal analysis,”Amer. Math. Soc. Transl. (2), Vol. 161, pp. 19-50, 1994

[16] Y.Okabe, Y.Nakano:“The theory of KM_2O-Langevin equations and its applications to data analysis (1) : Stationary analysis,”Hokkaido Mathematical Journal, Vol. 20, pp. 45-90, 1991

[17] 中溝高好:”信号解析とシステム同定”, コロナ社, 1988

[18] 川嶋弘尚, 酒井英明:“現代スペクトル解析,” 森北出版, 1992

[19] The Math Works,Inc., “ウェーブレットツールボックス ユーザーズガイド,” サイバネットシステム株式会社,(1996).

[20] チャールズ K. チュウイ:“ウェーブレット入門,” 東京電機大学出版局,1993

[21] 榊原進:“ウェーヴレットビギナーズガイド,” 東京電機大学出版局,1995