

スラーに基づく演奏表現：パラメータ推定と生成手法の考察

橋田 光代^{†,††} 片寄 晴弘^{†††,†}

本稿では、スラーに関する演奏表現モデルを提案し、その効用について考察する。ここでは、1) 演奏ルール抽出アルゴリズムを利用したスラー演奏モデルの効用調査、2) のべ95曲に対するスラー演奏パラメータの抽出と分類、3) 同じ楽曲で解釈の異なる版（楽譜）に対するスラーパラメータの適用の3種類の実験を行った。「別れの曲」では、再構成した演奏と実演奏との相関係数が、Velocity 0.979、テンポ 0.877 という高い値を得た。また、スラーに対応すると考えられるフレーズが必ずしも「山型」の表現にはならないこと、グルーピングのまとまり感が、音量とテンポの相補的な相互作用によって表現されている可能性があることを示す興味深いデータが得られた。

A Study on Expression of Slurring: Performance Model and Extraction of Parameters

MITSUYO HASHIDA^{†,††} and HARUHIRO KATAYOSE^{†††,†}

This paper proposes an expression model of slurring and considers the effectiveness of the model. This paper describes 1) an investigation of the model contribution to performance rendering, 2) an analysis of slur parameters and classification of 95 performances, and 3) applying the extracted slur parameters from a piece to the other interpretation. The coefficients between the source and the reconstruction were 0.987 for velocity and 0.877 for tempo. We obtained the interesting data, which show the general expression of phrases doesn't always become a *mountain-type*, used to be taken for granted, and which illustrate a complementary interaction between tempo and dynamics gives the taste of small phrases in terms of slurs.

1. はじめに

情緒あふれる演奏の自動生成は、自動作曲・聴取モデルとならば、音楽と人工知能に関連した主要研究テーマである。最近の報告では、提案されたシステムによる演奏生成例の中には、人間の演奏に劣らないものもあるという評価もされている¹⁾。その一方で、フレージングに関しては十分な解明が進んでいない。自動化に関する研究は限定されており、根本的な部分はなお人間の介入を前提としている^{2),3)}。

筆者らは、自らの演奏体験、打ち込みによる（情緒ある）演奏データ作成体験を通じて、1) 楽譜に明記されたスラーの表現には、従来から知られているアーティキュレーション表意に加え、フレーズの一般表現のような演奏情報も与える機能がある、2) フレー

ズの頂点と、スラーの幾何学的な頂点が、より近い場所で一致する、という仮説を持つに至った。演奏家にとっては暗黙のうちに理解・実施されている可能性はあるが、今まで定量的に検討されたことはなかった。本稿の主題の一つはこのことを定量的に確認することである。

スラー表現にかかわる演奏ルールを精緻化することの第一のメリットに、より自然な演奏を自動生成するということがあげられる。機械によるフレージングが難しい理由の一つとして、フレーズ（ここでは、比較的短い範囲のものを想定する。サブフレーズとしての捉え方も可能）の始点、終点、頂点の同定問題があるが、もし、楽譜上のスラーがその情報を担うとすれば、演奏ルール（特に条件節）をより明確に生成することが期待できる。加えて、本稿では、フレーズの頂点がスラーの幾何学的な頂点に当たるという仮説に基づいて、演奏パラメータを解析し、さらに、そこで得られたパラメータを用いて演奏を生成し、評価を行うこととした。

フレーズ中で最も重要になる音。

† 科学技術振興事業団さきがけ研究 21「協調と制御」領域 PRESTO/JST

†† 和歌山大学システム工学研究科
System Engineering of Wakayama University

††† 関西学院大学理工学部情報科学科
Information Technology of Kwansai Gakuin University
音を滑らかにつなぐという意味。legatoでも同様。

以下、第2章では、フレーズとスラーの演奏表現に関する基本的な考え方を整理し、スラーに基づく演奏生成モデルと実験概要について述べる。第3章で、実験を実施するためのシステム環境、およびモデルの記述手法について説明を行った上で、第4章で、3種類の実験と考察について述べる。

2. スラーに基づく演奏生成モデルと実験概要

2.1 フレーズ表現とスラー

音楽におけるフレーズは、音楽構造の一部を表す短い単位として一般にも広く理解されている⁴⁾。その一般的な表現は、音量と演奏速度が、頂点に向かって増え、逆に、頂点を過ぎれば小さくなるとされている。

比較的大きな構造におけるフレーズは、和声分析などによって特定が比較的容易で、演奏は一般表現にほぼ等しいと考えられる。一方で、より小さいレベルのフレーズ(以下、下位フレーズ)ほど、その演奏表現は不規則であり、また頂点の位置も一定ではない。そのため、下位フレーズの位置と頂点を同定することは、演奏生成において最も重要な課題となっている。

スラーは、通常、音価(音長)に対する演奏指示として知られている(図1)。しかし、それだけでは、ひとつのグループとしてスラーのまとまり感を表現するのは難しい。演奏者は、効果的な聴取のために、音量変化(デューナミック)とテンポ変化(アゴーギク)を組み合わせさせて与えている。

我々は、自らの音楽体験と楽譜の調査から、ふたつの仮説を持つに至った。1)多くの楽曲において、スラーの与えられた部分と下位フレーズとが一致する。2)各楽曲において、下位フレーズの頂点とされる場所が、多くの場合、スラーの幾何学的頂点の位置に近い場所で一致する。1)については Jackendoff からも指摘しており、フレーズを特定する根拠のひとつに、スラーの存在を取り上げている⁵⁾。裏返せば、楽譜上のスラーに対し、フレーズとしての演奏表現を与えることで、より豊かな表情のある演奏を生成することが期



図1 一般的なスラー表現
(ダンパーペダルを使用しないピアノ演奏の場合)

概ね 8-16 小節を想定するが、楽曲の構造や基本テンポ等により長さは一定ではない。

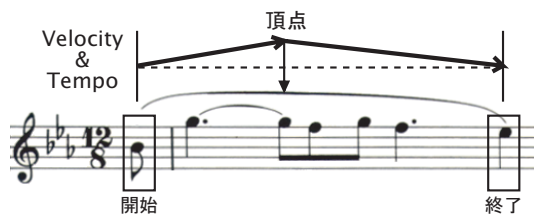


図2 スラーに基づく演奏モデル

待できる。

2.2 スラーに基づく演奏生成モデル

前節の観察に基づき、我々は、以下のようなスラー表現モデルを設定した。

- (1) スラーは下位フレーズを規定している。
- (2) スラーの幾何学的な頂点に近い音符がフレーズの頂点となる⁷⁾。
- (3) 音量とテンポに対し、スラーの頂点位置に表現が変化する屈曲点がある。

表現に関する基本的なモデルを図2に示す。この図は、1)音量は、スラー開始から頂点までクレッシェンド、頂点からスラー終点までデクレッシェンドする。2)テンポは、スラー開始から頂点までアツェレランド、頂点からスラー終点までリタルダンドする。3)スラー開始および終了時に、その前後の音符グループとの間には何らかの音量変化またはテンポ変化が現れることを模式的に表したものである。この図自体は「山型」で説明しているが、実際のモデルは頂点で屈曲点を持つという汎化したものを用意している(3.2節)。

2.3 実験の概要

本研究の主題は、提案したスラーモデルを検証することである。その方法は、分析的な手法と生成的な手法とに大別される。前者は、実際の演奏で本モデルと一致するような表現が行われているかを、後者は、本モデルで生成した演奏を、聴取実験等を通じてその正当性を探るアプローチである。

本稿では、分析的な手法として、1)演奏ルール抽出アルゴリズムを利用したスラーモデル効用の調査、2)さまざまな楽曲におけるスラーパラメータの抽出と分類を実施することにした。生成的な手法については RENCON アクティビティがまさに相当する⁶⁾。ここでは、分析的な手法と生成的な手法を組み合わせた3)同じ楽曲の異なる版(解釈)に対するスラーパラメータの適用を取り上げて、検討を行うことにする。

FIT 連根⁶⁾での聞き比べにおいて、我々はスラー表現を重視した Analysis by Synthesis によるルールベースの演奏生成を行った。音楽的な表現の観点で上位となり、その際、フレージングが良いという審査員のコメントを得ている。

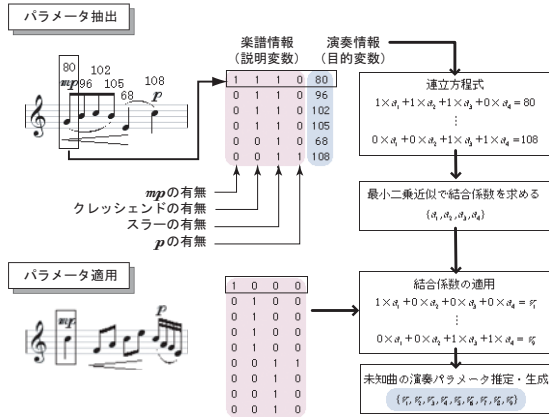


図 3 重回帰分析の楽曲への適用

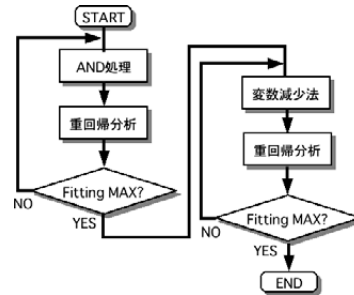


図 4 イタレーションアルゴリズム



3. 実験環境

3.1 実験環境

音楽の表現モデルを検証するには、1) 楽曲からのモデルパラメータの抽出、2) 抽出したパラメータに基づく演奏生成に関する環境が必要となる。そのため、我々は、演奏生成エンジンとパラメータ抽出エンジンからなる統合的な実験環境を用意した。

演奏生成エンジンでは、表情のない演奏に対し、楽譜上の表意記号あるいは楽譜を解析して得られる理論情報をもとに、演奏ルールを適用し、その結果を足し合わせて演奏表情を与える。ルール処理は OPS、その他の処理は C 言語で実装した。演奏ルールには、音量 (Velocity)、テンポ (Beat-time)、発音時刻、拍内表情 (一拍以下の微細表情) に働くモデルを用意している。演奏ルールのパラメータは、人間が直接入力できる他、パラメータ抽出エンジンの結果を用いることもできる。

パラメータ抽出エンジンでは、演奏ルール抽出のために開発された重回帰分析のイタレーションアルゴリズム²⁾を用いる。重回帰分析では、与えられた説明変数に結合係数を乗じた値の線形和によって目的変数が与えられるという仮定のもとに、結合係数を最小二乗法で求める (図 3)。本手法では、処理効率を上げるために、抽出したパラメータによる演奏の再現で誤差の最も大きい部分を検出し、一種の山登り法で、AND 結合した説明変数を付加し、さらに変数減少法によって説明変数選択を行う繰り返し処理を行う (図 4)。

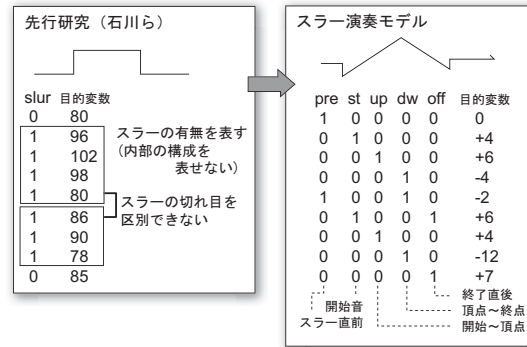


図 5 スラーの説明変数表現

3.2 スラーモデルの記述

スラーモデルのパラメータ抽出では、スラーの位置に合わせて、計 5 つの説明変数を用意した (図 5)。“up” と “dw” は、頂点で表現の変化が屈曲するという仮定に基づき、スラーの頂点を境にした前半部と後半部を示す。“pre”、“st” および “off” は、連続したスラーの区別や、スラーの演奏表現の可能性 (2 節) を考慮するものである。

本研究で用いるスラーモデルは、図 2, 5 のように、基本的には、一次関数の組み合わせによって記述される。最小二乗法で単純に結合係数を求める場合であれば、説明変数の値を直線に乗るような形で与えれば、モデルのパラメータを直接に得ることができる。しかし、本研究でのパラメータ抽出エンジンでは、説明変数を AND 結合して用いるため、量的な値を与えることはできない。ここでは、演奏データ、すなわち目的変数を差分系で与えるとともに、スラーのパラメータも差分系で与えることにより、この問題に対応した。

1 拍当たりの時間をミリ秒の単位で表したものとす。通常用いられる 1 分間当たりの拍数 (BPM) の逆数にあたる。

4. 実験と検討

4.1 演奏ルール抽出アルゴリズムを利用したスラーモデル効用の調査

解析の対象となった演奏が、抽出されたパラメータによってどの程度再現できるかをみることで、ルールの有効性を検証することができる。ここでは、スラーの表現モデルの有効性を確認するために、クレッシェンドをはじめとする7個の基本的な演奏ルールに、

- A スラーに関するルールなし
- B スラーの存在を示すルール²⁾
- C スラー演奏モデル

を、オプションとして加える形で、生成された3つの演奏を比較することにした。

この実験では、ショパン「エチュード Op.25-3 (別れの曲)」パデレフスキ版の楽譜を基準とした。演奏データには、音楽を専門に学び教師の経験もあるピアニストの演奏を用いた。図6にそれぞれの生成結果、表1に、実演奏と生成演奏との相関係数(以下、再現率とよぶ)を示す。再現率は、演奏データと重回帰分析時に与えた差分系データの双方について求めた。

全体として、スラー情報のない演奏 A よりも B の方が良く、B よりも本研究で提案する手法を用いた C の方が再現率が向上した。差分系データ、演奏データ

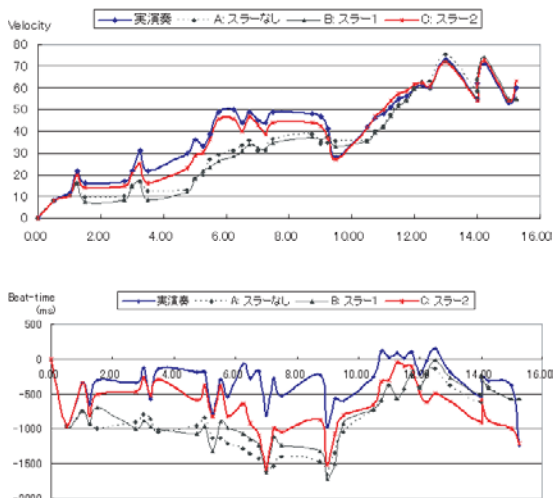


図6 「別れの曲」生成結果

表1 実演奏と生成演奏の相関係数
(差分系データ/演奏データ)

生成パターン	Velocity	Beat-time
A	0.784 / 0.936	0.599 / 0.470
B	0.828 / 0.931	0.718 / 0.539
C	0.979 / 0.989	0.877 / 0.747

いずれも、スラーに与えるルールを精緻にするほど再現率が向上した。特に Velocity では、スラールールを含まない A でも 0.936 という高い再現率が得られていたが、C はそこからさらに 0.053 向上させることができた。しかし B では、差分系データの再現率は向上したにもかかわらず、生成された演奏ではほとんど変化がみられず、むしろ、ごくわずかではあるが値が低下 (-0.005) する結果となった。この結果から、単にスラーの有無を与えただけでは、必ずしも生成する演奏に適切な表現を与えられるとは限らないことがいえる。

Beat-time においても、C は A から約 27%、B から 20% 近く再現率を向上させることができた。図6(下)では、中間部で実演奏と大きく隔たりがあるが、演奏データの形状としては、A、B と比べて C は実演奏にかなり近い。スラーモデルは、演奏表現の変化の仕方を重視するものである。したがって、実演奏に平行するようなデータとなった C は、このモデルが有効に機能していることを示している。

4.2 さまざまな楽曲における

スラーパラメータの抽出と分類

より多くの楽曲にスラー演奏モデルを与え、そこからパラメータを抽出することで、スラーモデルが一般性を持っているかを確認できる。ここでは、特に、スラー開始から頂点(“up”)と頂点から終点(“dw”)の変化を重視し、95曲の実演奏データに対するパラメータ抽出実験を行った。55曲は著者自身による演奏、2曲を前節のピアニストの演奏とし、残り38曲は、YAMAHAのサイレントピアノ用演奏データを用いた。楽曲は、楽譜全体にスラーの記述が多いことを条件に、ショパンやベートーヴェン、ドビュッシーなどクラシックのほか、ポピュラー曲のピアノアレンジも対象とした。演奏を記録する際は、できるだけ楽譜の表記に素直に従うようにした。

パラメータ抽出にあたり、スラーの種類を大きく3つに分類することにした。本研究で生成される演奏データはパラメータの線形和であるため、スラーが長くなるほど、演奏は極端に大きく変化することになる。実際の演奏では、ある一定の範囲に収まる表現が行われているため、スラーの長さに合わせたパラメータを取得する必要があると考えられる。そこで、各曲の基本となるテンポをもとに個々のスラーが平均的に必要とする演奏時間を求めた。その分布と音楽的な構成とを照らし合わせたところ、1) 1秒以下、2) 1秒以上6秒以下、3) 6秒以上の3つに分けるのがよいであろうと判断した。

表 2 スラーの演奏パラメータ抽出結果

上段: Velocity	曲数	説明変数	
下段: Beat-time	のべ 158	up	dw
(~1 秒)	41	0.964	-0.774
(1~6 秒)	88	2.112	-0.729
(6 秒~)	29	-1.176	-0.648
平均		1.210	-0.725
(~1 秒)	41	56.837	-22.587
(1~6 秒)	88	4.291	-14.960
(6 秒~)	29	-14.196	-99.982
平均		12.740	-32.787

表 3 スラーパラメータ分類

Vel.	Beat-time				合計
	++	+-	-+	--	
++	14	22	4	6	46
+-	20	28	8	13	69
-+	5	6	6	4	20
--	5	7	6	5	23
合計	44	63	24	27	158

表 2 に、抽出したスラーパラメータを分類別に平均した一覧を示す。値は小さいものの、Velocity, Beat-time, スラーの長さに関係なく、両者の値には、はっきりと差が現れた。このことから、頂点を境にスラーが屈曲点を持つことが言える。しかし、両者の値の傾向は、必ずしもフレーズの一般表現として予想されたものにはならなかった。一般的とされるフレーズの表現は、Velocity が + 方向から - 方向の「山型」、Beat-time が - 方向から + 方向の「谷型」(BPM では「山型」となる)という組であるが、表 2 では、Beat-time が「山型」になった。また、6 秒以上の長いスラーでは、屈曲点はあるものの、Velocity も Beat-time も右下がりの方向、つまり、ディミヌエンドしつつ演奏が速くなる値となった。

表 3 に“up”と“dw”の方向の組み合わせの分布を示す。Velocity, Beat-time とともに、“up”では + 方向へ、“dw”は、ばらつきがあるものの、どちらかといえば - 方向、つまりどちらも「山型」となる方が多かった。この結果からも、一般表現として思われていた事例は必ずしも多くないことが言える。

4.3 同じ楽曲の異なった版(解釈)に対するスラーパラメータの適用

意味のある演奏パラメータが抽出されているかどうかを確認する手法として、そのパラメータを別の音楽構造(解釈)に適用してみるという方法がある。もし、スラーに関する演奏モデルが正しいとすれば、抽出したスラーパラメータを別の音楽構造(解釈)に適用しても、聴取者は正しいスラーの表現を認知できるはずである。ここでは、Velocity に対して、スラー表記の異なる複数種類の楽譜が出版されているベートーヴェ

ンのピアノソナタ「悲愴」第 2 楽章(図 7, 図 8)を対象にパラメータを適用し、その聴取実験を行った。

図 9 は図 7 による演奏データ、図 10 は、図 9 で得たパラメータを図 8 に適用した演奏データを表す。図 7, 8 の楽譜の範囲は、図 9, 10 のカギ括弧で示した部分にあたる。図 10 の実演奏は、生成演奏に対する参考データとして別途記録した。適用する演奏ルールには、スラーのほかに、臨時記号、スタッカートを含めた。数値的な比較では、図 9 の再現率が 0.904 であったのに対し、図 10 では 0.736 に下がった。しかし、演奏データの形状としては比較的良好である。

次に、楽器経験のある 3 人の被験者に対し、生成した演奏データをそれぞれ聴取してもらった。演奏データはいずれも、Velocity は抽出パラメータ、テンポは一定(BPM:四分音符=60)とした。

聴取の結果、全員が、2 つの演奏において音楽の構造に違いがあることを判別した。そのうち 2 人は、図 10 が図 8 に基づく演奏であると回答した。しかし、同時



図 7 「悲愴」ヘンレ版(下段のスラーは省略)



図 8 「悲愴」YAMAHA による出版

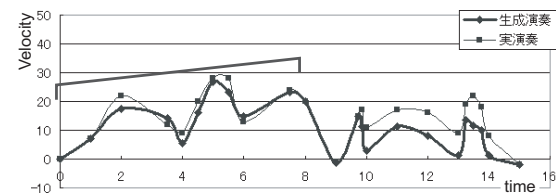


図 9 図 7 による生成演奏(再現率 0.904)

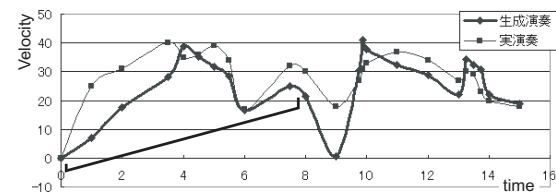


図 10 図 8 による生成演奏(再現率 0.736)

に、3小節目冒頭のC音に対しては、図8の演奏として強すぎると指摘した。一方、他の2人よりは音楽経験が浅い残りの1人は、そのC音を新しいグループの始まりであると聴取した。

続いて、Velocityは変えずに、テンポを人間の演奏から抽出したもので合成して、再度聴取実験を行った。この結果は、演奏としても非常に自然なものとなり、被験者は3人も正しいスラー付与を判定できた。このことから、Velocity値が同じでも、前後の音との時間的な距離によって、演奏の音楽的なバランスが変わる可能性があることが言える。

4.4 検 討

第1の実験からは、スラーモデルを導入することで、演奏の再現性が高まることが確認できた。特に、Beat-time、すなわちテンポ表現については2割から3割の再現率の向上が見られる。演奏生成のためのルールとして、スラーが非常に重要な役割を持っていることが確認できる。

第2の実験からは、屈曲点を持つという基本モデルの中で、いろいろなタイプのスラー表現が存在することが確認された。楽譜の調査から、かなりの頻度でスラーと下位レベルのフレーズが一致することもわかっている。フレーズ表現の基本的なアーキタイプとしては「山型」モデルが知られていたが、必ずしもそのようなものが多くないことが確認された。特にBeat-timeが、頂点に向かって長くなる、すなわち、テンポが遅くなるものは少なくない。このような状況に対する説明は今まででなされてこなかったが、例えば、音量での制御が飽和した際に、表現上のエネルギーが時間に転換されるという説明も可能であろう。

第3の実験は、抽出したパラメータの汎用性の確認を目的に実施したものである。スラー表記の異なる別の楽譜に対して、抽出したスラーパラメータからVelocity情報を与えただけでは、正しいグルーピングを満足して聴取することができなかった。これにテンポを合わせることで、楽譜表記に相当するグループ境界を認知できた。Velocityとテンポ間の相互作用に関する研究の必要性を強く要請するデータの一つである。

また、第3実験で用いた「悲愴」では、Velocityは下降系、テンポは上昇系といった特有の演奏表現がなされていた。このような特徴は、第2の実験から、メンデルスゾーン「ロンド・カプリチオーソ」冒頭部、ドビュッシー「月の光」などにも現れた。これらの楽曲は、基本的に、1拍にかかる時間が1.5秒から2秒前後と比較的ゆったりしている点で共通している。ひとつの可能性として、このようなタイプの楽曲に多

く見られる比較的長いスラーで、音符が前半と比べて後半に集中する場合にこのような特徴が現れているのではないだろうか。さらに細かな検証が必要であるが、このように、スラーのかかり方によって、特徴的に相互作用が起こる可能性が多いものと予想される。今後の興味深い研究テーマである。

5. おわりに

本稿では、スラーに関する演奏表現モデルの提案を行い、その検証を実施した。フレーズの場所や頂点を同定するのは容易ではないが、楽譜に記されたスラーの利用によって、比較的小さい単位のフレーズに対し効果的なフレージングを実現できる可能性を示した。

3種類の実験の結果から、1)スラー演奏モデルが演奏生成に対して効果的であること、2)フレーズ表現として一般的に理解されていた表現が必ずしも「山型」になるわけではないことが確認された。また、3)グルーピングのまとまり感が、音量とテンポの相補的な相互作用によって表現されている可能性を示す興味深いデータが得られた。

現段階では、複数の声部で別々にスラーが与えられた場合の演奏モデルについての検討が不足している。今後は、その対応と、音長に対するベダリングや運指のモデルについて考慮していきたい。

謝辞 本研究は、科学技術振興事業団さきがけ研究21「協調と制御」領域研究として実施されました。

参 考 文 献

- 1) 平賀瑠美, 大島千佳, 西本一志. Rencon を外と内から眺めたら … . 情報処理学会研究報告音楽情報科学 2002-MUS-50, pp. 19–24, May 2003.
- 2) 石川修, 片寄晴弘, 井口征士. 重回帰分析のイタレーションによる演奏ルールの抽出と解析. 情報処理学会論文誌, Vol. 43, No. 2, pp. 68–76, 2002.
- 3) G. Widmer and A. Tobudic. Playing mozart by analogy: Learning phrase-level timing and dynamics strategies. In *International Conference on Auditory Display (ICAD) Proc.*, pp. 28–35. ICMA, 2002.
- 4) 講談社. ニューグローブ世界音楽大事典, 1993.
- 5) Lerdahl and Jackendoff. *A Generative Theory of Tonal Music*. MIT Press, 1983.
- 6) 橋田, 野池, 平賀, 平田, 片寄. Fit 2002 rencon workshop —報告と課題—. 情報処理学会研究報告音楽情報科学 2002-MUS-48, pp. 35–99, Dec. 2002.
- 7) 橋田光代, 片寄晴弘. スラーに基づく演奏表現の一検討. 情報処理学会研究報告音楽情報科学 2003-MUS-49, pp. 19–23, Feb. 2003.