

# テンポ・音量・音色の実時間制御を目的とした 表情付けシステム

馬場 隆<sup>1,a)</sup> Sergio Canazza<sup>2</sup> Antonio Roda<sup>2</sup> 片寄 晴弘<sup>1</sup> Giovanni De Poli<sup>2</sup>

**概要:** インタラクティブな自動表情付けシステムにおいて、テンポ (Agogik)・強弱 (Dynamik)・音色 (Klangfarbe) の3つを同時操作できるものは、ほぼ例をみない。本研究では、音色変化を実現した表情付けシステム CaRo と、テンポや音量の制御が制御可能な指揮システム VirtualPhilharmony を統合し、上記3つの音楽表情を制御できるインタラクティブ自動表情付けシステムを提案する。その初期的な実施例について報告する。

**キーワード:** インタラクション, モーションキャプチャ, 自動表情付け, 指揮システム

## A Realtime Performance Rendering System for Controlling Tempo, Intensity and Timber

BABA TAKASHI<sup>1,a)</sup> CANAZZA SERGIO<sup>2</sup> ANTONIO RODA<sup>2</sup> KATAYOSE HARUHIRO<sup>1</sup> GIOVANNI DE POLI<sup>2</sup>

**Keywords:** Interaction, Motion Capture, Performance Rendering, Conducting System

### 1. はじめに

音楽の演奏には、読譜・演奏計画・発音に係る楽器操作等、高次の認知的作業と運動能力が求められる。その中で、音楽解釈と呼ばれる楽譜の解釈においては、量子化された音価、ピッチの情報を担う音符に対して、タイミングや音量を与え、さらには、アーティキュレーションや音色を付与するといった計画・立案作業が必要となる。これまでに、このプロセスに焦点を当てたさまざまな研究 [1] が実施され、演奏の表情付け研究 (Performance Rendering) と呼ばれる研究領域が定着するに至っている。また、その国際的なコンテストも実施されている [2]。

演奏の表情付けシステムの多くは、楽譜 (通常、MIDI や MusicXML フォーマットで記載される) 入力に対して、表情のついた演奏を生成するプロセスを、オフラインの処理として実施してきた。これらは人間のプレイヤーとのインタラクションのモデルは持たない。その一方で、指揮システム [3] や自動伴奏システム [4] に代表されるように、演奏におけるテンポ制御やキューイングを目的としたインタラクティブシステムの開発も進められてきた [5]。インタラクティブシステムはプレイヤーとのインタラクションの方策が実装されているものの、システム自体に音楽解釈機構のモデルを持たそうとしたものはほとんどない。

筆者らはこれまでに、指揮システム VirtualPhilharmony の開発を進めてきた [6]。VirtualPhilharmony は表情のある楽曲テンプレートを利用しており、そのテンプレートはプレイ状況に適應して音楽性を保ちつつ変化する。VirtualPhilharmony は広義での音楽解釈モデルとインタラクションモデルを有する演奏システムであると言える。その

<sup>1</sup> 関西学院大学大学院理工学研究科  
Graduate School of Science and Technology,  
Kwansei Gakuin University

<sup>2</sup> Department of Information Engineering,  
Padova University

<sup>a)</sup> takashi-baba@kwansei.ac.jp

一方で、音色の制御は処理の対象としてこなかった。この問題に解決するものとして、2012年度に、Padova 大学で開発された、2次元の演奏印象マップを用いた実時間演奏表情制御システム CaRo とのコラボレーションプロジェクトを試行した。本稿では、その活動状況について紹介する。

以下、VirtualPhilharmony, CaRo の概要について紹介し、それらを利用したテンポ・音量・音色を実時間制御するためのコンバインドシステムと、そのシステムを利用した演奏の実施状況について述べる。

## 2. VirtualPhilharmony

指揮システム“VirtualPhilharmony”は、実際にオーケストラを指揮する感覚に焦点を当てた MIDI ベースの指揮システムである。開発者の指揮経験を活かして指揮に関する様々なヒューリスティクスをシステムに盛り込むことにより、VirtualPhilharmony は、従来の指揮システムにはなかったリアルな指揮感覚をユーザに提供することが可能となっている。指揮感覚とは即ち、指揮者とオーケストラとのインタラクションの感覚である。指揮者の振りに追従するだけの既存システムは言わば楽器演奏システムであり、意思を持った人間のオーケストラと指揮者とのインタラクションにより演奏を生成する、という相互依存関係が指揮システムには求められる。VirtualPhilharmony は、オーケストラの演奏モデルのヒューリスティックな構築によってこのインタラクションを再現している。

システム概要を図1に示す。ユーザの指揮動作を読み取るためのセンサには電子楽器テルミン(静電容量センサ)を用い、指揮動作における鉛直成分の極小点・極大点および加速度を検出、これらのデータから拍点と音量を抽出する。楽曲は、CDの録音演奏を聴取コピー作業によりMIDI化したものをテンプレートとして用いる。ユーザ入力の拍点と音量を基に、テンプレートの各音符をスケジューリングする部分は“コンサートマスター機能”と名付けられており、ヒューリスティックなオーケストラ演奏モデルを基に設計されている。オーケストラの全声部がそれぞれ個別のタイムラインに基いて動いており、旋律声部のみ独立して動く、ある箇所ではすべての声部のタイミングが揃う、といった細かな音楽表現が可能となっている。

コンサートマスター機能は、次拍のテンポ予測、テンプレート補正、ユーザとオーケストラとの駆け引き設定等8つの機能より成り、その多くは実録音の解析を基にした各種パラメータによりヒューリスティックに設計されている。楽曲のジャンルや拍子や局所的なアゴーク等を考慮に入れた詳細な次拍テンポ予測、ユーザのテンポに基づくテンプレートの実時間補正(拍内のリズム演奏法の補正等)、ユーザとオーケストラとのインタラクションの度合いのユーザによる調節等々が行われる。また、これら8つの機能は声部ごとにそれぞれ個別に適用される。

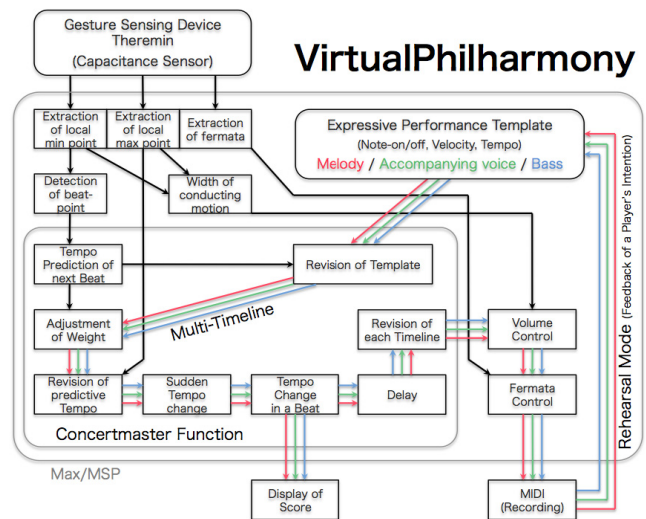


図1 VirtualPhilharmony の概略



図2 VirtualPhilharmony による演奏例

## 3. CaRo

Padova 大学は、音楽演奏の表情付けに関する研究を実施しているヨーロッパ有数の研究機関の一つである [7]。Padova 大学では、情緒ある演奏の生成に向け、ルールベースのシステムの構築を進められてきたが、生成演奏の表情 (Tempo, Legato, Intensity 等) を細かく制御すればするほど、ルール数、および、パラメータの設定 (図3参照) の設定が煩雑になるという問題に直面していた。この問題に対処するものとして、各ルール、パラメータの抽象化レベルでのメタコントローラの提供を企図して提案された表情付けシステムが CaRo である [8]。CaRo では、音楽演奏の印象の得点付けを行い、その因子分析によって得られた Perceptual Parametric Space (PPS) と呼ばれる 2次元演奏印象マップ (図3) が利用される。このマップでの2軸は、それぞれ、運動性とエネルギーに関連付けられるとされている。これら2軸により、音色やアーティキュレーションの制御を行うといった応用も可能である。

CaRo は、演奏の表情の生成に関連した運動、認知、音楽スキルの修得に向け、どちらかといえば、上級ユーザを想定して提案されたシステムであったが、その一方で、PPS 上にマッピングされた happy や sad などの情動情報や、soft や dark などの音色形容詞情報をメタファーとして用いることで、音楽スキルに乏しい音楽初心者の利用に

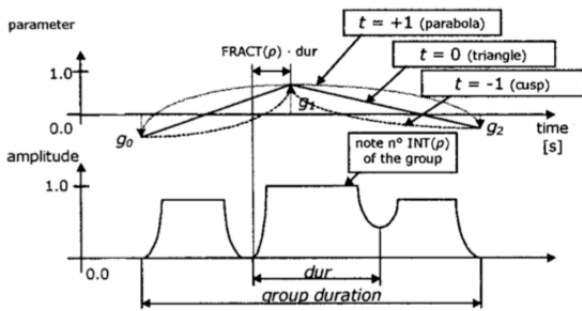


図 3 Padova 大学での表情付けシステムの演奏制御のルールとパラメータ (抜粋)



図 4 CaRo の PPS(左) と演奏インタフェース (全体). 右は PPS 上で与えられた座標位置によって変化する演奏パラメータ (一部)

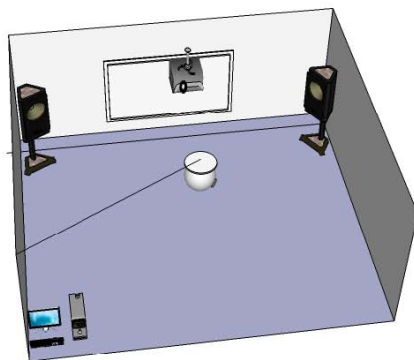


図 5 Stanza Logo-Motoria の全体図

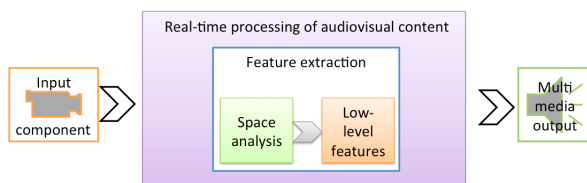


図 6 Stanza Logo-Motoria のシステム構成図

も有効であることが明らかになってきた。現在は、図 4 に示すような、PPS を用いたリアルタイムの演奏表情のコントローラが提供されている。

## 4. システム構成

本プロジェクトの狙いは、A) 広義での音楽解釈モデルとインタラクションモデルを有し、B) 音色の制御も行う拍打型演奏インタフェースの実装とそのシステムを使った演奏試行である。A) および B) での拍打型インタフェースについては、VirtualPhilharmony によって既に実現されている。CaRo との連携により、音色の制御を行うことがここでの取り組みの狙いとなる。CaRo はピアノ曲を対象としているため、コンバインドシステムでもピアノ曲のみを対象とした。これは、特にピアノという楽器が減衰系の楽器であり、かつ他のオーケストラの楽器に比べて音色のバリエーションに乏しく、音色変化の要因が音量やアーティキュレーション等に限定されているためである。CaRo は C++ にて実装されているが、VirtualPhilharmony との連携のため Max/MSP にて必要部分の再実装を行った。VirtualPhilharmony, CaRo とも、センシングの選択により、さまざまな形式での実装・使用状況が想定されるが、今回は、空間内で動き回って演奏を実施するという形式での実装を試みた。

システムのコンセプトを図 7 に示す。センサには、プレイヤが空間内を動き回る必要があることから、テルミンでなく iPhone 内蔵の 3 軸加速度センサを利用する。加速度データより指揮動作の極小点・極大点および加速度を求めて、これらから拍点と音量を抽出し、VirtualPhilharmony のコンサートマスター機能によって楽曲のすべての音符をスケジューリングする。アーティキュレーションと音色制御のための PPS に関する情報は、プレイヤの空間内の座標として与えられ、空間上部に設置されたカメラ、もしくは、Kinect によって取得され、CaRo により音色パラメータへと変換されて VirtualPhilharmony へと送られる。以下、カメラおよび kinect によるユーザの動きのキャプチャ法について述べる。

### 4.1 モーションキャプチャに Stanza Logo-Motoria を用いた場合

Stanza Logo-Motoria は、Padova 大にて開発された学習とコミュニケーションのためのインタラクティブ環境である [9]。空間上部中央に設置されたカメラにより人間の動きを二次元キャプチャする。出力として四方に設置されたサラウンドスピーカを備えている。visual-audio のマッピングを行うためのグラフィカルなプログラミング環境が整備されている。Stanza Logo-Motoria の全体図を図 5、システム構成図を図 6 に示す。

コンバインドシステムでは、Stanza Logo-Motoria の 2 次元平面座標を CaRo における PPS に見立てることにより、プレイヤの Stanza 空間内での動きを CaRo における

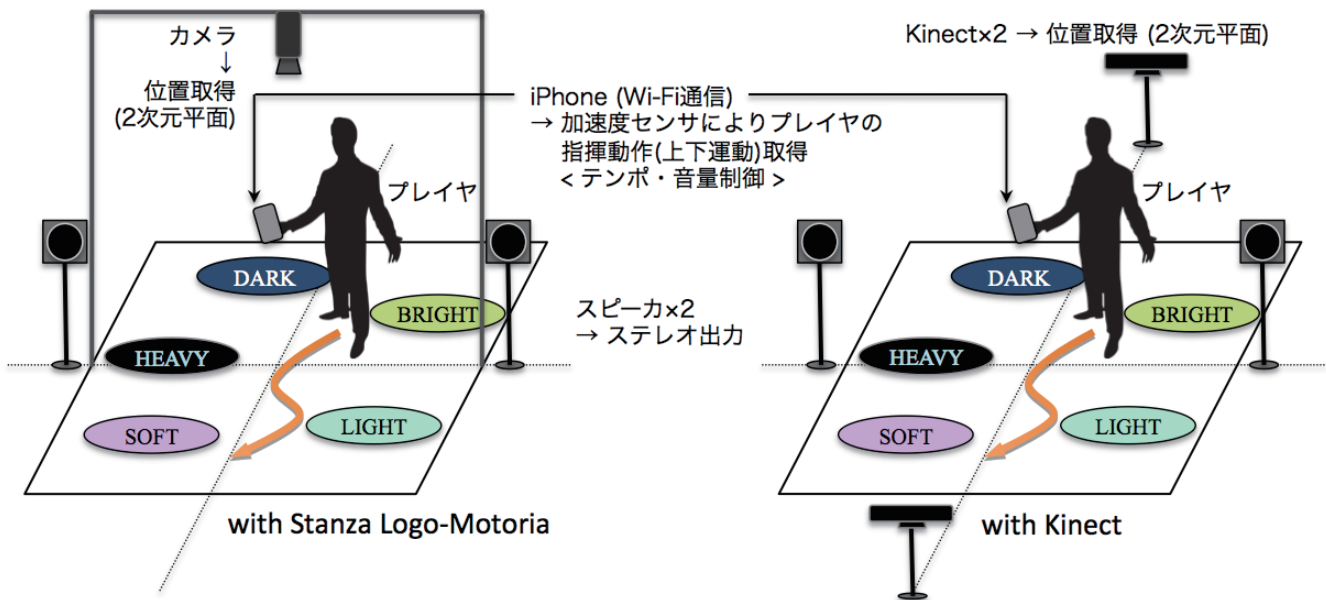


図7 コンバインドシステムのコンセプト. 左) Stanza Logo-Motoria を用いたシステム構成, 右) Kinect を用いたシステム構成

カーソルの動き (マウス操作) に置換している。プレイヤーは, Stanza Logo-Motoria の空間内を自由に動き回りながら, iPhone を右手に持って指揮を行う。即ち, プレイヤーの平面上の動きに合わせて楽曲のアーティキュレーションや音色が, 指揮動作によって楽曲のテンポや音量が, それぞれリアルタイムに変化する。VirtualPhilharmony の部分については, iPhone の加速度センサデータを基に得られた拍点及び音量の情報を基に, コンサートマスター機能を用いて各音符のスケジューリングを行った。音の出力に関しては, サラウンドスピーカは使用せず, 2ch のステレオスピーカとして使用した。コンセプトは図 7-左を参照。

#### 4.2 モーションキャプチャに Kinect を用いた場合

Kinect は, RGB カメラ・深度センサ等が内蔵された, Microsoft 製のゲームデバイスである [10]。RGB カメラにより水平方向の, 深度センサにより奥行き方向の, それぞれプレイヤーの動きをキャプチャする。これを 2 台用いて, Stanza Logo-Motoria と同様にプレイヤーの 2 次元平面座標上の位置をリアルタイムに取得する。その他の部分に関しては, Stanza Logo-Motoria を用いた場合と同様である。コンセプトを図 7-右に, 演奏例を図 8 示す。

### 5. ケーススタディ

#### 5.1 Stanza Logo-Motoria を用いた実験

2012 年 7 月, Padova 大にて Stanza Logo-Motoria を用いたコンバインドシステムの演奏実験を行った。演奏楽曲は, F. Kuhlau のピアノ・ソナタ作品 88 の第 3 楽章である。演奏の結果, テンポ・音量・音色の 3 つをそれぞれ自由に制御できることが確認された。ただし, この時点で



図8 コンバインドシステムの演奏例 (Kinect を用いた場合)

は CaRo の Max/MSP への再実装が一部に留まっていたため, PPS を完全に再現できていなかった。

#### 5.2 Kinect を用いた実験

2012 年 12 月, Kinect を用いたコンバインドシステムを, 岡山大学オーケストラの楽団員数名とその友人 (音楽経験少) に体験してもらった。併せて, VirtualPhilharmony 単独での演奏も体験させたところ, VirtualPhilharmony 単独に比べ音色の変化を耳で確認できたとのコメントを得た。さらに, 動き回る指揮者という新たな表現方法に対する好意的コメントが多かった。一方で, 動きの必然性が薄い, あくまで指揮の動作によって音色を変更させたい等の意見も寄せられた。今回は CaRo の PPS をそのまま空間に持ってきたが, 指揮者が動きながら音色を変える時に必要な PPS と CaRo の PPS とでは, 必ずしも一致しないということが示唆された。また, サラウンド環境にした上で, オーケストラ内 (=空間内) を指揮者が自由に動き回り, 近

傍にいる楽器の音を大きくしたりするのはどうか、との提案も受けた。指揮台に立つだけでなく、オケ内の様々な位置から指揮演奏を楽しめる、という新たなエンターテインメントの可能性を秘めており、実装を検討中である。

## 6. おわりに

本稿では、VirtualPhilharmony, CaRo を組み合わせ、iPhone とプレイヤ位置のセンシングによって、広義での音楽解釈モデルとインタラクションモデルを有し、音色の制御を行うパフォーマンスシステムの概略とその演奏試行について報告した。

VirtualPhilharmony を用いて楽曲のテンポと音量を制御するとともに、CaRo を用いて楽曲の音色を制御した。後者については、CaRo の PPS を実空間に置き換えることで、2次元平面上の人間(プレイヤ)の動きを基に音色制御を実現した。モーションのキャプチャには Stanza Logo-Motoria を用いたものと Kinect を用いたものの2種類を用意した。

Kinect を用いたものについては使用評価を実施し、様々なコメントを得た。今後は、指揮者の動きと音色の変化の関係性への考慮、評価者の対象の拡大、新たなコンバインドシステムの模索等を行なっていく予定である。

## 参考文献

- [1] Kirke, A., and E. R. Miranda: An Overview of Computer Systems for Expressive Music Performance. In A. Kirke, and E. R. Miranda, (editors) *Guide to Computing for Expressive Music Performance*. Springer London, pp.1-47 (2013).
- [2] Haruhiro Katayose, Mitsuyo Hashida, Giovanni De Poli and Keiji Hirata: On Evaluating Systems for Generating Expressive Music Performance: the Rencon Experience, *Journal of New Music Research*, Vol. 41, no. 4, pp. 299-310 (2012)
- [3] Mathews, M.: The Conductor Program and Mechanical Baton, *proc. Intl. Computer Music Conf. (ICMC)*, pp.58-70 (1989)
- [4] R. B. Dannenberg: An On-line Algorithm for Real-Time Accompaniment, *proc. Intl. Computer Music Conf. (ICMC)*, pp.193-198 (1984)
- [5] Fabiani, M., A. Friberg, and Bresin, R.: Systems for Interactive Control of Computer Generated Music Performance. In A. Kirke, and E. R. Miranda, (editors) *Guide to Computing for Expressive Music Performance*. Springer London, pp.49-73 (2013).
- [6] Takashi Baba, Mitsuyo Hashida, Haruhiro Katayose: VirtualPhilharmony: a conducting system focused on a sensation of conducting a real orchestra, *Proc. ACE 2010* (2010)
- [7] Canazza S, De Poli G, Roda A, Vidolin A: Analysis and synthesis of expressive intention in a clarinet performance, *Proc. Intl. Computer Music Conf. (ICMC)*, pp.113-120 (1997)
- [8] Canazza, S., G. De Poli, A. Rodà, and A. Vidolin: An abstract control space for communication of sensory expressive intentions in music performance. *Journal of New Music Research*, Vol.32, No.3, pp.281-294 (2003)

- [9] Camurri, A. and Volpe, G. and Canazza, S. and Canepa, C. and Roda, A. and Zanolla, S. and Foresti, G.L.: The Stanza Logo-Motoria: an interactive environment for learning and communication, *Proc. SMC 2010*, pp.353-360 (2010)
- [10] <http://www.microsoft.com/en-us/kinectforwindows/>