

音楽音響と映像信号を対象としたジェスチャによる指揮システム

鈴木 健嗣[†] 金尚 泰^{††} 小尾 正和^{†††} 橋本 周司^{†††}

A gestural conducting of music using audio and video signal rendering

KENJI SUZUKI,[†] KIM SANGTAE,^{††} MASAKAZU OBI^{†††}
and SHUJI HASHIMOTO^{†††}

1. はじめに

我々は、ジェスチャによる演奏情報により音楽を制御することで、誰もが気軽に音楽を指揮することができるシステムの構築を試みている。これまで、計算機による指揮システムは、MIDIによる音響生成、及びCGによる演奏の視覚的提示が主な手法であり、CDなどの音響信号を対象とし、実時間で人間の指揮に追従した音響生成を実現したシステムは少ない^{1),2)}。そこで本稿では、ユーザの指揮動作による音響信号の制御だけでなく、オーケストラ等の演奏映像を用いて映像の再生制御と音響制御を同時に行なうことで、より臨場感の高い演奏制御を実現するシステムを提案する。また、ジェスチャ取得のために手首部に装着する無線デバイスを開発した。これより、指揮棒による制約を越えて、自由に音量・テンポといった演奏情報を制御できる音楽指揮システムについて報告する。

2. システム概要

本システムの構成を図1に示す。提案する指揮システムでは、ユーザは、加速度センサと通信機器が内蔵されたセンサグローブを右手首部に装着し、予め定められた指揮動作により、音響・映像制御を行なうことができる。音響・映像制御、及びセンサデータ解析には、Max/MSPとEyesWebを用いている。音楽音響とその演奏映像データからなるMPEGファイルは、それぞれ音響部・映像部に分離し、演奏制御のためにシステムバッファに取り込んでおく。さらに、楽音・映像データに加え、予め楽音データから作成した拍点の位置を表す拍点データを利用することも可能である。装着するセンサグローブは、三軸加速度センサと

Bluetoothによる無線通信機、及びバッテリーからなる。加速度センサはグローブに固定されており、体軀から右腕に対し前後方向をX軸、左右方向をY軸、上下方向がZ軸となる。センサによって得られる加速度データは、10bitでA/D変換され、計測時刻とともにBluetooth通信を介して、約60Hzの間隔で制御用PCへ送信される。制御用PCでは、受信したデータを解析しテンポを抽出することで、音楽の音量や時間伸張、及び映像の再生速度を同期して制御する。処理された音響信号はスピーカを通じ、また映像信号はディスプレイを通じてユーザに提示される。

3. 演奏情報の取得と音響・映像制御

図2に、ジェスチャ認識による演奏情報の取得と、音楽・映像制御の概要を示す。図3は、開発した無線デバイスの概要である。ここでは、オーケストラ及びカルテットの演奏映像を対象とし、その演奏制御とは、センサによって得られた加速度の情報により、演奏の主音量とテンポ、及び映像再生の時間制御を指す。

3.1 主音量の制御

主音量の制御には、加速度の強度を利用する。三軸方向の加速度の絶対値により、演奏の主音量を調整することとした。これにより、ユーザのジェスチャの大小が、直接演奏の音量の大小に反映される。

3.2 テンポ抽出

テンポ制御には、Z軸方向の加速度の変化から抽出した拍点の位置から求めたテンポを用いる。指揮法において、演奏者にテンポの指示をするのに用いられる拍点は、基本的には鉛直方向の加速度の変曲点によっ

[†] 筑波大学 大学院システム情報工学研究科
Graduate School of Systems and Information Engineering,
University of Tsukuba

^{††} 筑波大学 大学院人間総合科学研究科(芸術学系)
Institute of Art and Design, University of Tsukuba

^{†††} 早稲田大学 理工学術院/理工学研究科
School of Science and Engineering, Waseda University



図1 システム構成：音響映像制御システム
Fig. 1 System Overview: A music conducting system

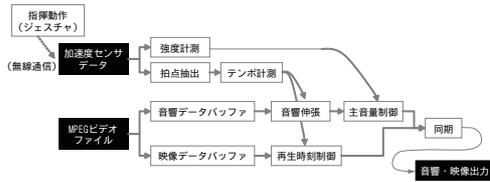


図 2 ジェスチャ認識と音響・映像制御のフロー
Fig. 2 Gesture recognition and audio/video rendering



図 3 AccGlove : センサグローブの構成
Fig. 3 AccGlove : Overview of sensor glove

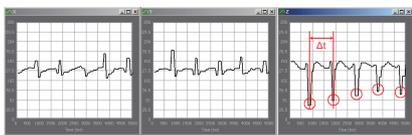


図 4 加速度データと拍点抽出 (左から各 xyz 軸に相当)
Fig. 4 Acceleration data and beat extraction (x,y,z axis from the left figure)

て示される．そこで今回はセンサを上から下へ振り下ろした瞬間（最下点）を拍点と定めた．指揮動作の最下点においては，Z 軸方向の加速度値が極大値となるため，加速度変化に伴い局所的に閾値を定めるように工夫することでピーク検出を行い，拍点を検出することができる．図 4 に，拍点抽出の例を示す．予備実験では，例外動作を除いた全ての点で拍点抽出に成功しており，被験者は若干の練習により操作が可能であった．

3.3 再生速度制御

再生速度は，拍点の時間間隔の逆数から求められるテンポに応じて定められる．抽出されたテンポに基づき，時間伸張された音楽音響データと演奏映像データを同期させて再生する．しかしながら，MIDI データによる演奏とは異なり，音響信号の制御では，再生速度に応じてピッチ変調が生じる．本システムでは，音響信号の時間伸張のために，固定時間長に分割された音響データの一部を削除・重ね合わせを行なう代表的な手法である CUT&SPICE 法に基づく手法を応用し，さらに時間伸張を行なった波形を 1/4 周期ずつずらして重ね合わせることでノイズ低減を行うことで，実時間にて良好な品質の音響再生制御を実現した．

3.4 拍点追従制御

演奏の“間”を実現するため，拍点データを利用した再生制御を可能とした．拍点データは，予め演奏前にユーザがセンサグローブを用い，時間伸張のない

音響・映像データを見ながら拍点の位置を入力することで作成する．つまり，曲の先頭から各拍点までの時間が 1 ミリ秒単位で記録されたデータである．演奏時に拍点データを用いる場合，ユーザの指示する拍点の位置と実際の演奏の拍点の位置を常に合わせながら音楽・映像を再生する．再生時間の伸張により，記録された拍位置と演奏の拍位置にズレがある場合は，同期制御時に拍点合わせの処理を行なう¹⁾．

3.5 指揮動作の開始・中断・再開

本システムでは，演奏を始める際，センサを上下に振る動作より開始テンポを指定し，ユーザの指揮動作（ここでは上下運動）から抽出されたテンポが楽曲本来のテンポの 70 % ~ 130 % の間になった時にシステムは再生を始めることとした．一方，ユーザが途中で指揮動作を中断し，演奏に対して二拍以上の遅れが生じた場合は，最後に得られたテンポに基づいて引き続き演奏を行なうこととした．また，途中でユーザが再びジェスチャ動作を始め，拍点が認識されると，最も近い拍点の位置から音楽とジェスチャを同期させ，再びテンポ制御を行なうことを可能とした．

4. まとめと今後の課題

本稿では，開発した無線デバイスによるジェスチャ取得に基づき，身近にある CD などの音楽音響の再生制御だけでなく，音楽演奏映像の再生制御も同時に行なうことで，臨場感の高い演奏制御を実現するシステムについて報告した．これより，オーケストラや弦楽四重奏の演奏映像を用いて良好な演奏制御を実現した．既開発システム¹⁾を用いて行なった被験者実験より，初めて使った人でも比較的簡単に楽しめることが示されているが，本システムは入力デバイスの開発と映像情報を含めた拡張と言え，より多くの人々が手軽に音楽表現を楽しむことのできるシステムを目指した．今後，再生音響の更なる改善とともに，ユーザの多様な動作をパート別の演奏制御へ関連付けるなど，シリアスゲームとして演奏指揮の学習への応用も検討したい．

謝辞 本研究は，JST CREST の研究課題「人を引き込む身体性メディア場の生成・制御技術」，及び中山隼雄科学技術文化財団の支援によるものである．

参考文献

- 1) 小尾 正和, 鈴木 健嗣, 橋本 周司, “音楽音響信号を対象としたジェスチャによる音楽指揮システム”, 情報処理学会第 68 回全国大会, 工学院大学, 2L-6, 2006 .
- 2) E. Lee, et al., Toward a Framework for Interactive Systems to Conduct Digital Audio and Video Streams, *Comput. Music J.*, 30(1), pp.21-36, 2006.