

Virtual Performer :

6H-4

Virtual Musicianの音響・モーションセンサ

汪 増福* 金森 務* 片寄晴弘* 佐藤宏介** 井口征士**

* (財) イメージ情報科学研究所 **大阪大学

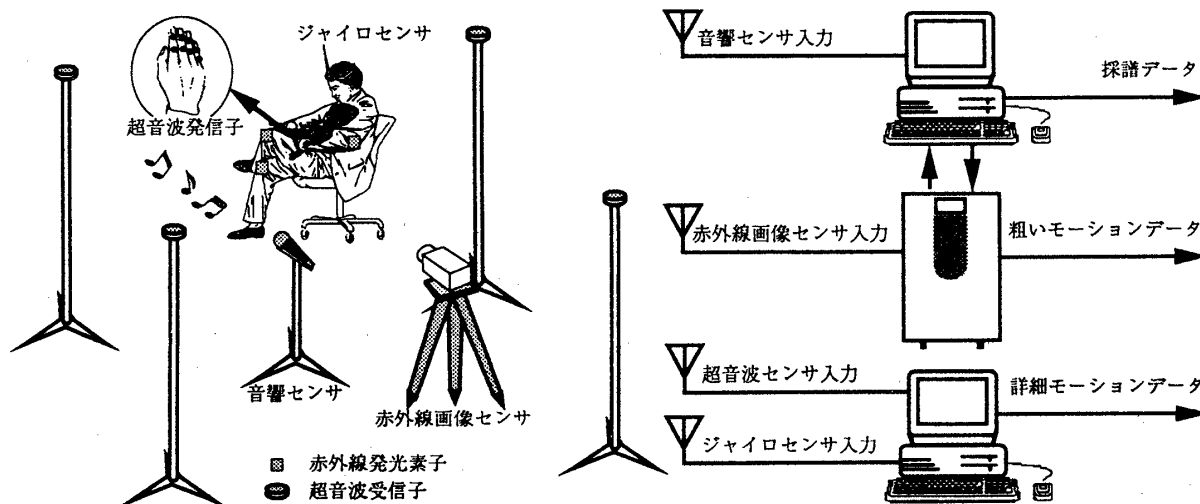


図1 マルチセンサの概要

1. はじめに

我々は、Virtual Performer, すなわち、人間と共演できるコンピュータ生成のパフォーマー（音楽家、舞踊家など）の実現を目指している。Virtual Musicianはその典型の例の一つである⁽¹⁾。Virtual Musicianを実現するためには色々な要素技術が必要であるが、なかでも演奏者の音楽ジェスチャー（演奏音、演奏動作など）をリアルタイムで計測する技術が重要な役割を果たすと思われる。本稿では、音響センサ、赤外線画像センサ、超音波センサとジャイロ角速度センサを含んだマルチセンサによりリアルタイムで演奏に関する音符情報とモーション情報を抽出する手法を提案する。

2. システム概要

Virtual Musicianシステムにおける演奏者の各種ジェスチャーを計測するマルチセンサの概要を図1に示す。

音響センサは生の音響信号を拾い、その中から演奏者の音響ジェスチャーを表す楽音の音符情報を検出する。また、赤外線画像センサ、超音波センサとジャイロ角速度センサからなる複合モーションセンサは各センサの特徴を活かして、演奏者のモーションジェスチャーを表す位置、速度、角速度などのモーション情報を検出する。

3. マルチセンサとその実現

3.1 音響センサ

音符情報を抽出するためには、デジタル周波数解析の方法を用いた。ここでは、音名に対応する特定の周波数成分のみを抽出すればよいので、FFTで全ての周波数成分を求めるのではなく、DFTで特定の周波数成分のみを直接計算する⁽²⁾。そこで、コンピュータにより、並列にDFTを計算し、周波数解析を高速に行っている。

音符情報：

- ・音の高さを示すピッチ情報
- ・音の強さを示すパワー情報
- ・ノート・オンとノート・オフ情報

抽出処理は下記のように行なう：

(1) 音響データから時間周波数マップを作成する。
ここで、サンプリング周波数は50KHzとしている。また、最大窓幅は50msである。12個のコンピュータはパイプライン的に接続され、それぞれ割り当てられた周波数成分のDFTを並列に計算する。窓を滑らせながら計算を行うので、ある時刻でのDFTを計算するのに、窓内の全データを計算する必要はなく、前の時刻のDFTを基にデータを更新するのみでよい。ただし、割り当てられた周波数成分に応じて、最大窓幅の範囲内で最適な窓幅を使用する。こうすることによって、時間周波数マップの時間分解能は1m

Virtual Performer: Acoustic and Motion Sensors for Virtual Musician

Z. Wang*, T. Kanamori*, H. Katayose*, K. Sato* and S. Inokuchi**

* Laboratories of Image Information Science and Technology,

** Osaka University

s 程度を実現している。時間周波数マップと滑り窓の関係を図2に示す。

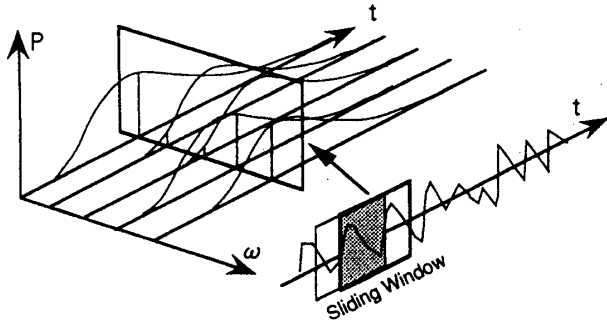


図2 時間周波数マップと滑り窓

(2) 各音ごとに対応するスペクトル成分の時間マップ(エンベロープ)からノート・オンとノート・オフのタイミングを検出する。

検出ルール:

- ・立ち上がりはエンベロープの最大値の位置に対応
- ・立ち下がりにはエンベロープの値がノート・オンの状態からあるしきい値まで下がった位置に対応

図3には検出ルールを適用した例を示す。

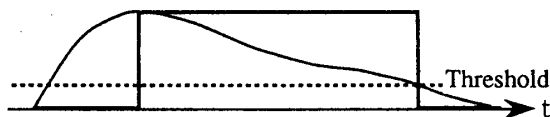


図3 検出ルールの適用例

(3) 検出された音に対して、お互いに倍音関係を持つ音の間に倍音解析を行ない、実際に弾かれた基本音のピッチとパワーを同定する。

上に述べた手法は単音だけでなく、和音を含めた複数の音を同時に検出できる。また、計算処理の時間遅延は最大窓幅の半分超の30ms位に押さえている。

3.2 モーションセンサ

モーションセンサは赤外線画像センサ、超音波センサとジャイロ角速度センサから構成される。赤外線画像センサは粗い位置データを、超音波センサは詳細な位置データを得るのに用いられる。また、ジャイロ角速度センサは首の動きを表す首の角速度を得るのに用いられる。

3.2.1 赤外線画像センサ

演奏者の四肢のモーションを検出するために演奏者の主要関節部に赤外線発光素子を付けて、その発光素子の画像中の位置を検出する。

市販の普通のCCDカメラに可視カットフィルタを付けて用いている。こうすることによって、画面中から赤外線発光素子部分(輝点)だけを検出できている。

輝点検出処理もコンピュータにより並列処理で実現している。画像のサイズは256×256である。処理の流れは以下の通りである:

- (1) 入力画像を2値化し、輝点だけを検出する。
- (2) ラベリング処理を施し、輝点の重心位置などの物理量を計測する。
- (3) 輝点の追跡を行う。

3.2.2 超音波センサ

演奏の詳細なモーションジェスチャー、例えば、演奏する際の指の使い、を検出するために超音波センサを用いている。8つの超音波発信子を時分割的に使って、8点のモーションを計測する。例えば、発信子を演奏者の指の関節部に付けて計測を行う。発信子からの超音波を受けるために受信子を設けるが、単一の受信子では奥行き情報しか得られないので、多方向に受信子を設けて計測を行うことにした。2方向からの奥行き情報から発信子の位置を決定できる。ただし、固定した2方向からは演奏者自身による隠蔽や解のあいまい性などの問題がしばしば起きるので、質の高いデータを得るために多方向に受信子を設け、隠れのない位置データを得るように工夫をした。

超音波センサの計測周期は100msであり、また、空間分解能は5mmである。

3.2.3 ジャイロ角速度センサ

演奏楽器の種類によって、演奏者の首の動きは重要である。例えば、尺八の演奏では、首振り動作は重要な意味を持つ。そこで、演奏者の首の動きを計測するために、演奏者の頭部に(ヘアバンドタイプの)ジャイロ角速度センサを付けて首の3軸の角速度の変化を計測する。現在は有線でセンサを結合しているが、将来は無線方式に変更する予定である。

ジャイロ角速度センサの計測周期は20msであり、また、角度分解能は360°/10ビット程度である。

4. おわりに

複数種の視覚、聴覚センサによる演奏者のジェスチャーを計測する方法を提案した。これにより、音楽における人間と機械間のコミュニケーションが円滑に行えるようになりつつある。今後、提案したマルチセンサを積極的に応用していきたい。

参考文献

- (1) 片寄 他: "Virtual Performerの概要", 情報処理学会第46回全国大会一般講演6H-03.
- (2) Kaset S. 他: "構造可変型並列処理システム及びその応用", 1990年電子情報通信学会春季全国大会D-398.