

加速度センサを用いた手振り解析と実時間サウンド制御

6S-5

大倉新也 澤田秀之 橋本周司

早稲田大学 理工学部

1 はじめに

我々の日常生活における情報伝達、中でも意志や感情の伝達においては、言語以上に非言語的手段（ノンバーバルコミュニケーション）によるところが大きいと言われる。何気なく用いているジェスチャーが機械によって理解できれば、伝達者の感情や情緒などの微妙な表現にも対応できる自然な入力デバイスが実現できる。ジェスチャー情報は障害者用のコミュニケーションエイドに用いられている一方で、音楽にも良く利用されている[1][2]。これは楽器が、動作から音響への感性的な変換器であるためである。筆者らもデータグローブや画像処理を応用して動作によって音楽を制御する試みを行ってきたが[3][4]、システムが複雑で高価になるという問題があった。

本研究ではジェスチャーの中でも特に、「手振り」の加速度センサによる認識を取り上げた。手振りにおいて、人間の意志や感情は、その空間的な位置よりも、手に加えられる力に、より顕著に現れると考えられる。力の測定法は種々あるが、運動中に働く力は加速度によって検出することができる。加速度センサには構造が簡単であるという利点もある。また、身体にセンサあるいはマーカーなどを取り付けることによって得られる軌跡から力を算出する従来の方法では、時間的遅れが問題となる。

ここでは、3次元加速度センサを用いた手振りの認識と、実時間サウンド制御システムの試作について報告する。

2 システム概要

2-1 加速度センサ

センサ内部には、x, y, z方向の加速度を独立に検出するために、各々の軸方向で金属お

もりが圧電素子に接続してある。測定範囲は、 $-25 \sim +25$ Gであり、加速度にともなう圧電素子の電圧の変化が出力される。この信号は増幅されて、A/D変換ボードを通じて12ビットのデジタル値として計算機に取り込まれる。

センサから得られた3次元の加速度成分のうち、ここでは垂直平面内2方向（y, z方向）の情報を用いた。これを加速度ベクトル、 $A(t) = (A_y(t), A_z(t))$ とし、これから特徴パラメータを算出して手振り認識を行う。取り込み周波数は25 Hzで、1回の動作についてのサンプリング数は30とした。

2-2 演奏制御システム

図1のようにまず、上で定義した加速度ベクトルを次章のアルゴリズムで処理し、あらかじめ学習してある片手の動作をコンピュータに認識させる。次にテンポ情報と音色情報を認識結果によって変化させ、MIDI音源を通して出力する。

2-3 音響制御システム

図2に示すように、加速度センサから得られるパラメータをニューラルネットワークを通してMIDI音源とエフェクタに出力する。ニューラルネットワークは、手振りにより得られる物理量と音源の音響生成パラメータを使用者の好みに応じて関係づける。

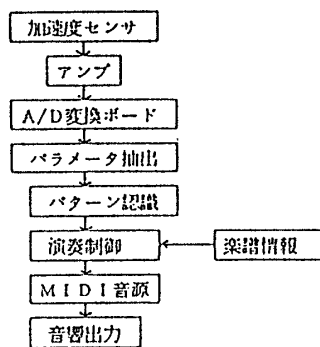


図1 演奏制御システムブロック図

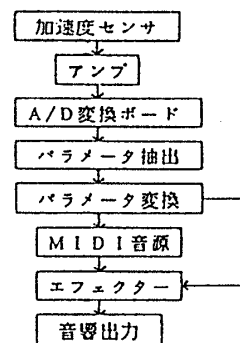


図2 音響制御システムブロック図

3 手振り認識アルゴリズム

3-1 運動特徴の抽出

得られた加速度ベクトル系列を図3のように y, z 平面上で捉え、次の12個のパラメータを運動の特徴量として算出する。

p_1 : ベクトルの時間差分

p_2 : 連続して得られたベクトルの外積

$p_3 \sim p_6$: $\arctan(Az / Ay)$ の各象限ごとの平均

$p_7 \sim p_{10}$: y, z 成分の最大、最小点へのベクトルの p_3 から p_6 方向への射影

p_{11} : 外接長方形の縦横の比率

p_{12} : 外接長方形の対角線の大きさ

3-2 手振り認識

認識は基本的には標準パターンとの比較により行っている。まず、学習モードで認識させたい動作のサンプルを5回づつ入力し、前記の各パラメータについての平均および標準偏差を求め、標準パターンデータとする。

認識モードでは未知動作を入力するたびに各パラメータについて標準パターンの平均値との二乗誤差を計算し、それぞれ標準偏差で割ったものを重み付き誤差とする。この重み付き誤差を各動作ごとに p_{12} 以外のパラメータについて足し合わせたものを「類似度」とする。類似度が最小で、かつ、あるしきい値よりも小さいものを標準パターンから選び認識結果とする。

p_{12} については、加速度の大きさ、つまり手に加えられた力の大きさを表すため、認識した動作の“強さ”の評価値とする。

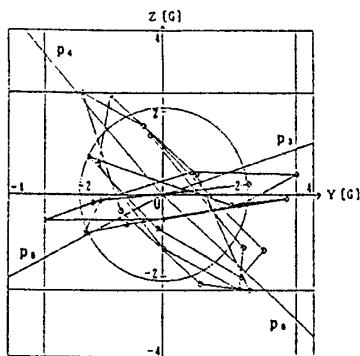


図3 y, z 平面における加速度ベクトル

4 サウンド制御

4-1 演奏制御

演奏制御では、現在次の12種の手振りを認識している。

[縦振り、横振り、斜め振り、時計回し、反

時計回し、強いひと振り、上に引き抜く、下に引き抜く、星を描く、十字を描く、×字を描く、静止]

これらの認識結果を、演奏のスタート、ストップ、テンポ変化、楽器選択、音量調節などに用いて実時間制御を行い、簡単な操作で自由度の高い演奏制御が可能になっている。

4-2 音響制御

音響パラメータとしては非常に多くの種類が考えられるが、ここでは、白色雑音に近い音源を用いて、その振幅包絡線の形とポストフィルタの遮断特性を身振りにより制御する。現在制御している音響パラメータは以下の6個である。

1) サスティン 2) ディケイ 3) リリース 4) アタック 5) ポストフィルタ高域特性 6) ポストフィルタ低域特性

ニューラルネットワークの学習は既発表[5]の手法を改良し連想記憶型としたため、対応付けの適応度が高まり扱いやすいものとなっている。

5 あとがき

現在は、 y, z 軸方向の加速度のみを用いているが、今後は更に x 軸方向の加速度も用いることにより、より複雑で細かな手振りの解析も可能になると考えられる。また、サンプリング時間および取り込み周波数についても検討中である。3D加速度センサを用いることにより、データグローブや画像を用いる場合に比べて、簡単で自由度の高いマン-マシンインタフェースが実現できると思われる。

本研究の一部は、文部省科学研究費(重点(2)「感性情報」課題番号06212225)によっている。

参考文献

- [1] 音楽情報処理の技術的基盤 平成4年度文部省科学研究費総合研究(B)報告書 第4章
- [2] Katayose et al. "Virtual Performer", Proc. of ICMC, 1993
- [3] Morita et al. "Computer Music System that Follows a Human Conductor", IEEE Computer, Vol. 24. no. 7, 1991
- [4] Sato et al. "Singing and Playing in Musical Virtual Space", Proc. ICMC, 1991
- [5] Hartono et al. "Adaptive Timbre Control Using Gesture", Proc. ICMC, 1994