音声解析を利用した芸術表現

市　川　　衛
助教授
デザイン学科
平成13年度

研究の背景

私のインタラクティブアート作品では、各種センサーや音声など外部からの物理的情報をトリガーとして、その情報によって音楽や映像をコンピュータによって生成する技術を用いているものが多い。外部入力として音声を利用した作品では、従来ではギターシンセサイザーのような音量のオンオフ的な情報だけを扱うシステムはあったが、音の細やかなニュアンスや表情に反応するような汎用システムは無く、そのようなシステムを独自開発することが長年の課題であった。

しかし、近年のコンピュータ技術の急速な発展によって、音声をリアルタイムに解析するための技術的なハードルが次第に低くなってきた。2001年にマッキントッシュで動作するMIDIおよびデジタルサウンドのビジュアルプログラミング環境であるMAX／MSPがバージョン4にアップされて、音声解析に関する機能が大幅に強化された。また、DSPを内蔵した高性能マイコンや開発キットが低価格で提供されるようになり、マイコンレベルでの音声解析も個人の開発のレベルで手が届く範囲となってきた。そのような技術的環境が整ってきたため、課題であったリアルタイムの音声解析を芸術表現に応用できるシステムの研究を行った。

研究の課題と目的

音声情報の主な特徴は、音量と周波数成分の時間的な変化である。音声の細やかな特徴を抽出するには、音量だけでなく音声に含まれる周波数成分のリアルタイムの解析が不可欠であると考えた。そのために、外部音声情報をデジタイズして、リアルタイムで周波数分析を行う実用的なテクニックを確立することが本研究の主要な課題となった。

リアルタイムに周波数分析を行うために、MAX／MSP4を用いたプログラミングを行い、どのような方法による解析が有効であるか、実験を進めた。

また、芸術作品に組み込むことができるシステムを実用化していくことが最終目標なので、パソコンだけではなく、DSP内蔵マイコンボードによる音声解析の試作を行い、基礎的な動作検証をした。

MAX／MSP4による音声の周波数スペクトル解析

マッキントッシュ対応のMAX／MSP4で行った音声の周波数解析の方法の概要を以下に述べる。

(1) adcオブジェクトでデジタイズした音声データをfftオブジェクトに入力し、リアルタイムで高速フーリエ変換（FFT）の演算を行う。FFT演算はサンプル時間の整数倍の正確な周期的な波形に対しては正確な結果が出せるが、一般に他の周波数成分が混入してしまうので、サンプルデータに窓関数と呼ばれるデータを掛け算してから行った。窓関数には代表的なハニング窓を採用した。

(2) FFT解析結果の実数成分と虚数成分から、cartopolオブジェクトを用いて周波数スペクトルの振幅を計算し、これをscopeオブジェクトに入力してオシロスコープのように視覚的に観察できるようにした。

(3) 周波数スペクトルの中から最大の振幅をもつ周波数を数値的に検出するために、maxiオブジェクトというMSPオブジェクトを独自にC言語でプログラミングをして、一定のしきい値以上のスペクトル成分を数値的に抽出できるようにした。

周波数スペクトル解析の結果

MAX／MSP4による上記の方法を用いて、実際の音声の周波数スペクトル解析を行ったところ、音声の音程を決定するメインの周波数をリアルタイムに数値的に検出することに成功した。

しかし、音声の特徴や時間経過による変化に対して、実際の聴感上の変化を大きく上回る周波数成分の振幅の変化が生じているように思われた。この原因は、音声の周波数成分の微妙な変化に対して、メインとなる周波数は離散的周波数成分に正確に一致しないため、前後のスペクトルにまたがる成分が実際には生じてしまい、主成分のスペクトルの振幅が必ずしも音量とは比例せず、聴感上の印象とは矛盾するような振幅変化となるのが原因であろうと予想される。理想的なスペクトル解析が実現されれば、音声の主周波数の倍音構成による音質の特徴なども抽出できると期待していたが、現状の音声スペクトル解析の方法では、そのような精緻な周波数成分分析までは無理であろうと判断された。

また、予想されていたことではあったが、ドラムの音のように特定の音程をもたない音声に対しては、このような周波数分析は無効であった。これらの音声に対しては、それらがノイズ成分となって間違った判断を下さないようなプログラミング対策が将来必要であることが明確となった。

DSPマイコンボードによる周波数スペクトル解析

日立製のDSP機能を搭載したSH2マイコンを搭載する音声解析評価ボードを用いて、MAX／MSP4同様の方法によってマイコンによる周波数スペクトル解析を、アセンブラー言語によるプログラミングで行った。マイコンレベルでも音声のスペクトル解析が実用になることを、試作レベルで動作検証できた。

今後の展開

外部音声情報をもとに、その主たる周波数成分を数値としてリアルタイムに抽出することができたので、今後はパソコンレベルではMAX／MSP4でそれを視覚情報に表現して、それが聴覚上の感覚と一致するような表現の実現を目指したい。その表現手法が確立されれば、音声解析部分をマイコンシステムに移植してシステムとして独立させ、利用していきたい。