

Virtual Performer : 適応型カラオケシステム

6H-6

竹内庸博* 片寄晴弘** 井口征士*

*大阪大学基礎工学部 ** (財) イメージ情報科学研究所

1. はじめに

我々は、ソロ演奏者に対して視覚・聴覚的效果によりあたかもセッションをおこなっているかのような環境を提供するシステム「Virtual Performer」を提案する[1]。

本稿ではVirtual Performer上のアプリケーションとして、自動伴奏システムで歌唱を対象とした場合の応用例について示す。自動伴奏システムとはソリストによる旋律音の入力に対応する伴奏音をリアルタイムで出力するシステムであり、あらかじめロードされた楽譜データと入力を照合しながらソリストのテンポ変化や音量変化にも追従した伴奏を行うシステムである[2][3][4]。通常のカラオケはCDなどに記録された伴奏を一方的に出力するだけで、人間はそれに合わせなければならないが、本システムでは伴奏の方が人間に合わせてくれるという適応型のカラオケを実現している[5]。

ここでは、自動伴奏システムで特に歌唱を対象とした場合に発生する問題点とそれに対するアルゴリズム上の工夫について述べる。

2. システムの概要

本システムのハードウェア構成と概念を図1、図2に示す。

イベント検出モジュールは入力された音響信号から旋律音のアタック時刻、パワー、周波数を抽出する(これらの情報はシステム中ではMIDIデータに変換して用いられる)。これらの情報はマッチャーへ送られる。マッチングの結果、正しい音高入力であった場合にはパラメータ修正エージェントがテンポ、音量を修正する。誤った入力がなされた場合には、その頻度に応じて歌唱位置探索モジュールが起動される。テンポ、音量、演奏位置が決定されると、スケジューラがタイムテーブルを作成する。そのタイムテーブルに基づきシーケンサが順次伴奏を出力する。

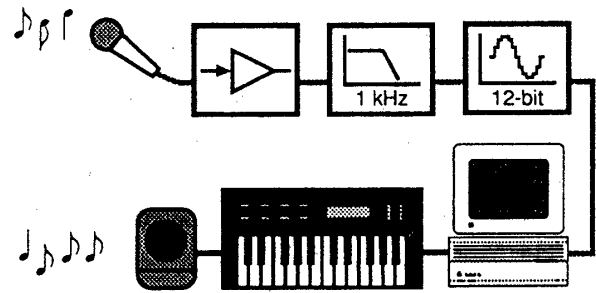


図1 ハードウェア構成

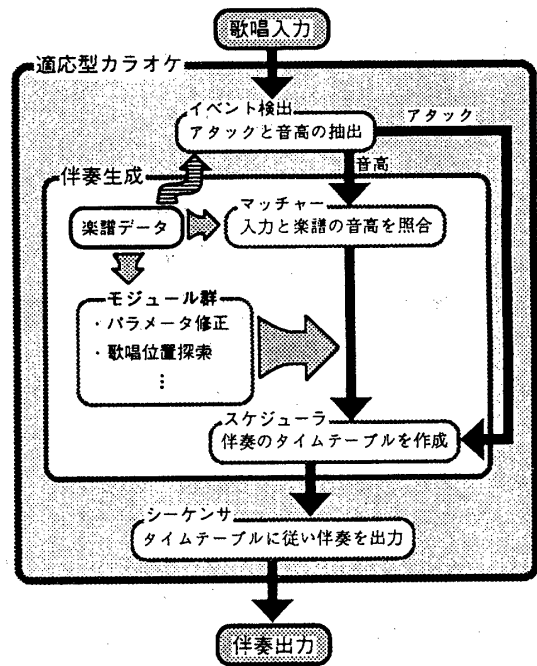


図2 システムの概念

3. イベントの検出

歌唱を対象とした場合、通常アタック時刻から50~60msec経たないと周波数が安定しないため、アタックの直後に周波数を計算すると誤認識が多いため、本システムではアタックが検出された時点から周波数が安定するまでの平均をとり、これをアタック時刻における周波数として用いている。

また、歌唱ではタンギングの明確な部分以外では周波数のみが増減し、パワーはあまり変化していない場合が多い。また、声のふらつきによるパワーのゆれも生じる。よって単純なパワーの変化率や閾値

処理などでアタックを検出するのは非常に困難である。そこで本システムでは、楽譜データから予測される次のイベント発生時刻が一番大きく、そこから遠いほど小さくなるような重み関数をパワーに掛け合わせ、その値によってアタックかどうかを判定している(図3)。これにより、上記のような方法よりも正確かつ確実にアタックを検出できるようになった。

また、周波数の計算には簡単な自己相関法を用いているが、全周波数に対して計算を行うのは時間の無駄である。そこで本システムでは、必要な周波数(全声域中で平均率音階にあてはまる周波数)とその近辺に対してのみ計算を行い、その際の窓幅も計算する周波数に応じて変化させることにより処理の高速化を行い、リアルタイム性の向上を図っている。

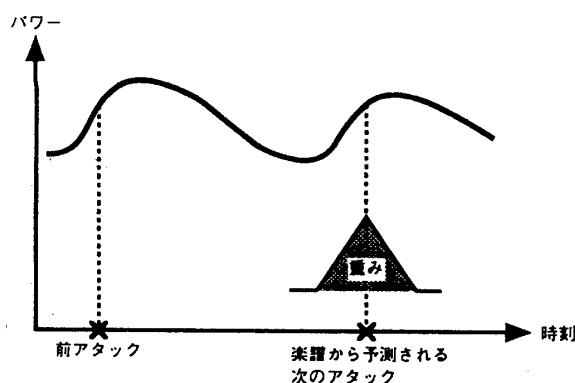


図3 アタックの検出

4. マッチングとシーケンシング

上記のようなイベント検出方法を用いているため、アタック検出時刻から音高抽出時刻まで約60msecの遅れが生じる。このため、音高が抽出されてからマッチングを行いスケジューリングしては適切な伴奏を行うことができない。そこで、本システムでは明確なアタックが検出された場合、即座に次の伴奏のスケジューリングを行い、マッチングは後で行うという方法を取っている。そしてマッチングの結果、正しい入力であった場合にはそのまま伴奏は続行されるが、誤った入力であると判断された場合には、次にマッチングに成功した時点でタイムテーブルを修正する。

また、次の入力予測される時刻までに明確なアタックが検出されなかった場合にも、予測どおりに入力があったものとしてスケジューリングを行う。これにより、歌詞で母音が続くときなど、アタックの検出が困難な場合にも対応できる。これは、通常カラオケにおいて極端なテンポ変化はしないという

ことを前提としている。しかし、フレーズの切れ目などでは楽譜上に記述されない「間」が挿入されることが多いため、上記のような「先読み」はせず、フレーズの頭のアタックで伴奏を合わせるようにしている。フレーズ情報については楽譜データにフラグとして与えている。

5. 適応性のためのモジュール

本システムは適応性を向上させるための様々なモジュールを備えている。

パラメータ修正モジュールは、楽譜データと入力を照合し、各旋律音の単位音価あたりの時間を計算することによりテンポを修正する。また、入力のパワーの変化に伴い伴奏の音量も修正する。

歌唱位置探索モジュールは、誤った入力が続く場合には歌手が曲の途中を飛ばし、楽譜上の他の位置を歌っているものと判断し、入力シーケンスとマッチする部分を探索する。

キー調整モジュールは、歌手が常に自分の最も歌いやすいキーで歌えるように曲全体を移調する。

この他にも、歌い出しで曲を検索するモジュール、サビの部分で入力に対してハーモニーを付加するモジュールなどがある。

6. まとめ

通常のカラオケにおける「歌手は伴奏に合わせてなければならない」という問題点を解決し「伴奏が歌手に合わせてくれる」カラオケシステムを実現した。

7. 今後の課題

イベント検出のアルゴリズムを再検討し、リアルタイム性および音高抽出の信頼性の向上を図りたい。

また、リハーサルを行うことにより、歌手特有の癖などにも対応できるシステムを実現したい。

【参考文献】

- [1] 片寄 他: "Virtual Performer の概要", 本大会 6H-03
- [2] 竹内 他: "リアルタイム性を考慮した自動伴奏システム", 音楽情報科学研究会会報, Vol.38, pp.24-25, 1991
- [3] Roger B. Dannenberg: "An On-Line Algorithm for Real-Time Accompaniment", ICMC Proceedings, pp.193-198, 1984
- [4] Barry Vercoe: "The Synthetic Performer in the Context of Live Performance", ICMC Proceedings, pp.199-200, 1984
- [5] 井上 他: "歌声自動伴奏システム", 音楽情報科学研究会夏のシンポジウム'92 Proceedings, pp.79-82, 1992