

歌声ピッチの自動補正システム

7Q-5

井上 渉 橋本 周司 大照 完

早稲田大学 理工学部

1 はじめに

筆者らはすでにメロディの楽譜を参照しながら、歌声から自動的に歌唱テンポを検出し、伴奏をつけるシステム、つまり「歌に合わせて伴奏が流れる」適応型カラオケの試みについて発表している[1]。このシステムでは、任意の調で、また曲の途中から歌っても、歌唱の進行している小節を判定し、自動的に歌声に合わせた伴奏をつけることができる。

今回報告するシステムは、カラオケにおいて、歌の音程がはずれているときに、正しい音程とのズレを検出し、自動的に歌声の音程を補正することにより正しい音(歌声)を実時間で出力させるものである。

2 システム構成

本システムのハードウェア構成は、システム制御用のパーソナルコンピュータ、ピッチ検出用のDSPボード、ピッチ補正用のエフェクター、および伴奏出力用のMIDI音源からなっており、曲に関する知識としては主旋律(歌)と伴奏の楽譜を持っ

ている。

図1にシステムの基本構成を示す。まず、マイクからの入力信号をDSPで処理し、歌声のピッチ検出をリアルタイムで行う。検出したピッチと主旋律の楽譜から、正しい音程とのズレを求め、この情報をMIDI信号でエフェクターに送信する。エフェクターは受信した情報により、ピッチチェンジを行い、効果音(音程を補正した歌声)だけをスピーカーより出力する。

3 歌声のピッチ検出

まず、マイクからの入力信号をローパスフィルタに通して高調波を除去し、サンプリング周波数20kHz、12ビットでA/D変換する。次に、DSPで歌声の基本周波数およびこれに対応する音名の決定を行う。

対象としている歌声は80~1000Hzまでで、音名の決定は半音間隔で行い、DSPからホストとなるパソコンには、決定された音名のMIDIノートナンバーが送られる。

4 ピッチ補正

4.1 楽譜情報

楽譜情報は、音名をMIDIデータと同様に7ビットで表し、その音長を最小音長音符の整数倍で表す。最小音長音符はシステムで扱う最も短い音符であり、本システムでは32分音符を最小音長音符としている。そして伴奏出力時の音量、和音等を示すフラグ、MIDIチャンネルと合わせたものを一つのイベントデータとし、5バイトで表している。

また歌詞データとして、歌詞と、その歌詞(1文字)に対応する音符の音長を楽譜データと同様に最小音長音符の整数倍で表したものをコンピュータに持たせている。これにより、パソコンのモニタに歌詞を一覧表示し、曲の進行に伴って、歌うところの歌詞の表示色を変えるようにした。

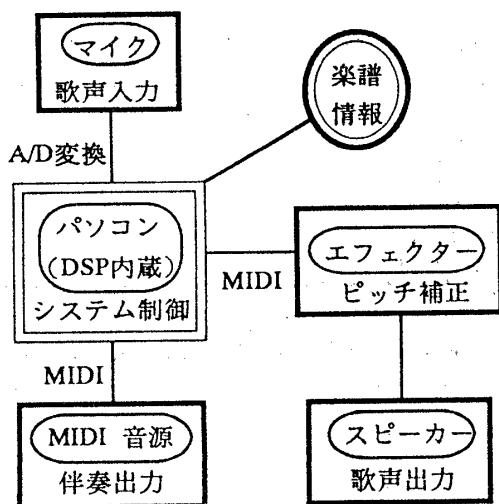


図1 システム構成

4.2 補正

DSPで検出した歌声ピッチと主旋律の楽譜から、正しい音程とのズレを求め、その情報を以下のようにMIDI信号でエフェクターに送信している。

ピッチ補正は、エフェクター（ヤマハ Digital Sound Processor SPX90II）のピッチチェンジプログラムを使用している[2]。

ピッチ変化の量は、“BASE KEY”パラメータで指定した音名と、MIDI信号で入力された音名との音階差で決められる。例えば、“BASE KEY = C3（MIDIノートナンバは48、以下同様）”とすると、MIDI信号でC4(60)が入力されると1オクターブ上のピッチに信号が変化し、F3(53)が入力されると5半音上に変化する。

ピッチは、±1オクターブ（±12半音）の範囲で可変で、BASE KEYとMIDI信号の音名との音階差が12半音以上のときは、12半音以内の範囲に制限される。

図2にDSPで決定された音名、正しい音名（楽譜）、およびエフェクターに送信するMIDI信号の音名を示す。例えば、時間[a]では楽譜の音名はC3(48)、DSPの音名はH2(47)であるので、ズレは+1となる。“BASE KEY = C3(48)”に設定しているので、エフェクターに送信するMIDI信号の音名はC#3(49)となる。

4.3 補正による音色の変化

もとの歌声（原音）とエフェクターによるピッチチェンジ後の歌声（効果音）とでは、声質（音色）の変化が感じられる。図3、4はそれぞれ原音および効果音の波形とその周波数スペクトルの一例である。図でわかるように、効果音のスペクトルは倍音以外の成分にも見られる。そこで、フィル

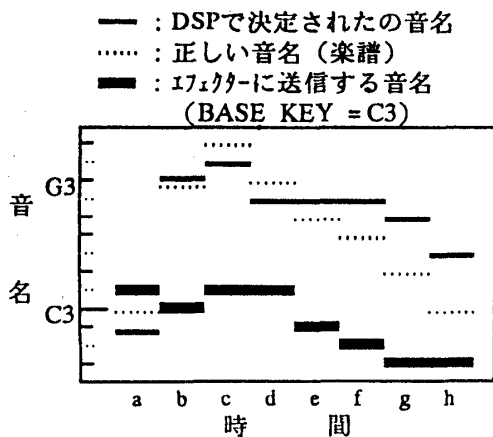


図2 各音名

ターを用いて倍音以外の成分を減少させ、十分ではないが簡単な補正を行っている。

5 おわりに

実時間で歌声のピッチと正しい音程とのズレを検出し、エフェクターを使うことによって歌声ピッチの補正を試みた。もとの歌声とピッチチェンジ後の歌声とでは、声質（音色）が変化してしまうが、音程は正しい音程に補正されてスピーカーから出力することができた。

今後の課題として、声質に変調が生じないようにピッチ補正方法、また前述の適応型カラオケシステム[1]との結合などを考えていきたい。

参考文献

[1]大照、橋本他：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(1992)

[2]ヤマハ DIGITAL SOUND PROCESSOR SPX90II 取扱説明書

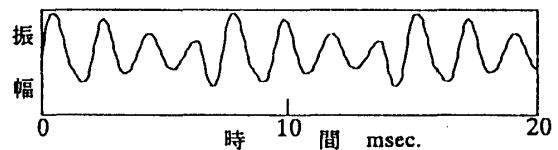


図3 (a) 原波形

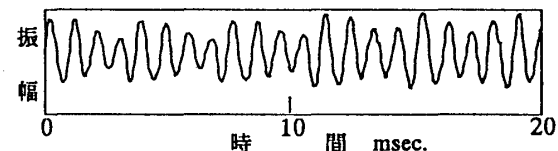


図3 (b) 1オクターブ上に変化させたときの波形

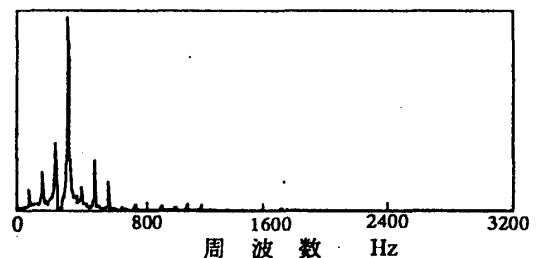


図4 (a) 原音の周波数スペクトル（基本周波数130Hz）

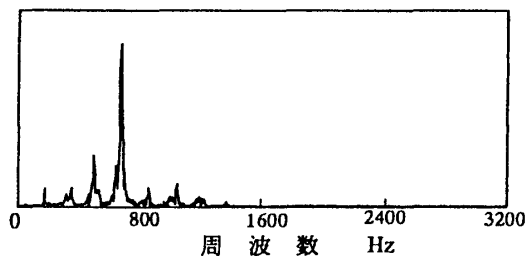


図4 (b) 効果音の周波数スペクトル（1オクターブ上に変化させたとき）