

古い録音メディアからの演奏情報の抽出

下条 敬洋[†] 片寄 晴弘^{††} 井口 征士[†]

[†]大阪大学基礎工学研究科

^{††}和歌山大学システム工学部

過去の貴重な音楽資料の多くは、低品質な音として記録・保存されており、快適な鑑賞が困難であったり再生自体が貴重な資料に損傷を与える場合がある等、利用が困難である。そこで、それらをデジタルアーカイブ化することにより、貴重な音楽資料の再利用が可能となる。デジタルアーカイブ化を行うには、採譜処理を用いた演奏の再現やノイズリダクション等が必要となるが、従来はそれらを個々に行っており、デジタルアーカイブ化を行っている研究例はない。本稿では、楽譜情報を用いた演奏情報の抽出とノイズリダクションによる演奏の復元を行うことによりデジタルアーカイブ化する手法を提案し、実際にその手法によりシステムを構築した。

Extraction of musical expression on music performance from audio signals including a harsh noise

Takahiro Gejo[†] Haruhiro Katayose^{††} Seiji Inokuchi[†]

[†]Faculty of Engineering Science, Osaka University

^{††}Faculty of Systems Engineering, Wakayama University

There are many rare recorded performances by famous artists. It is difficult to listen to those performances comfortably, because the quality of their sound is often very low. It will be possible that we listen to those performances comfortably by reviving those performances as Digital Archive. Digital Archive consists of Synthesis performance, Regenerate performance, and Original performance. First, it is effective in synthesizing to extract musical expression (for example pitch, velocity, and so on) from those performances. Second, it will be necessary in regeneration to reduce noise which those music signals include. In this paper, we propose the method which generate Digital Archive and we construct system that generate Digital Archive from those music signals.

1. はじめに

昨今の音楽メディアは、CDやDATのようにデジタル化された音響信号を扱うものが主流となっている。しかしながら今までに録音され残された貴重な演奏の中には、ろう管レコードやSPレコードのように、ノイズが多く鑑賞が困難であったり、再生自体が貴重な資産に損傷を与えかねないものが数多く存在している。そこで、低品質な音源として記録・保存されている音楽資産をデジタルアーカイブ化することにより、古い音楽メディアの再利用が可能となり、貴重な音楽資料を多くの人が共有できるようになると考えられる。

本研究で目標とするデジタルアーカイブの概念図を図1に示す。この図に示されているように、デジタルアーカイブ化を行うには、採譜処理による演奏情報の抽出やノイズリダクション等による演奏の復元が必要となる。採譜処理については、高見らの楽譜を用いたDPマッチング^{[1],[2]}によるものや、片寄らの知的採譜システム^[3]などを始めとして、数多くの研究が成されている^[4]。また、実際にろう管から演奏を復活させる試み^[5]も数多く成されている。

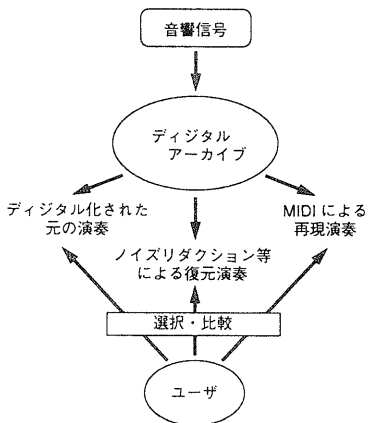


図1 デジタルアーカイブの概念

本稿では、対象をクラシックのピアノ独奏とし、音響信号から楽譜を用いて演奏情報を抽出する手法について述べる。これは、次に演奏される音の位置を推測しながら採譜を行う手法であり、人間が楽譜を見ながら音楽を聴く際に、現在の演奏のテンポと楽譜の音符の長さから、

次の演奏音の位置を予測する過程を実現させたものである。また実際に、この手法による演奏情報の抽出と、簡単なノイズリダクションによる演奏の復元を行うことで、デジタルアーカイブ化を行うためのシステムを構築した。

2. システム概要

本システムの概要を図2に示す。

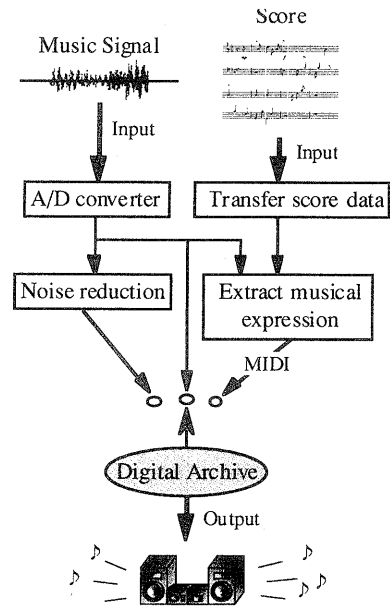


図2 システム概要

まず、本システムの入力部は、低品質音源のメディアの音をアナログフィルタを通してデジタル化し、コンピュータへの入力を行う。この音響信号のデータは、演奏情報の抽出部とデジタル信号処理による演奏の復元部へと送られる。同時に、必要な楽譜情報をデータ化し、コンピュータへ取り込む。楽譜データは、演奏情報の抽出部へと送られる。

演奏情報の抽出部では採譜処理を行う。ここでは、まず、得られた音響信号のデータに対しFFTを行い、周波数ごとのパワーの時間変化を得る（以下、ここで得られた時間と周波数のマップをTFマップと称する）。このTFマップと楽譜情報を用いてトップダウン的な採譜処理を行う。一般的に、ピアノ演奏における演奏情報としては以下のものをあげることができる。

- ・音量（ベロシティ）
- ・音の立ち上がり時間と終了時間
- ・音色（ペダル情報等）

上記の中でも、一般に誰もが容易にその違いを認識できるのは、音の立ち上がり時間と音量である。そこで、本稿ではこの2つの情報に重点を置いて情報抽出を行う。一方音色については、演奏者のタッチやペダルの踏み込み具合等によって微妙に変化し、さらには演奏するピアノによっても異なる⁶⁾。そのため、音色の情報を抽出するのは非常に困難であり、再現するのも容易なことではない。本稿では、ペダル記号を楽譜より取得し、ペダルを踏んでいると考えられる区間の音の終了時間を調整することで、音色の再現の代用としている。このようにして得られた演奏情報を用いることにより、MIDI機器で演奏を再現する。

演奏の復元部では、デジタル信号処理により演奏の復元を行う。本稿では、簡単なノイズリダクションにより演奏の復元を行っている。

上述の方法により、MIDI機器で再現した演奏とノイズリダクションにより復元した演奏、更に元の演奏の3つの演奏を得る。この3つの演奏によりデジタルアーカイブが形成される。

3. 楽譜を利用した情報抽出法

3.1 楽譜情報

楽譜には、楽曲に関する情報が数多く記されている。実際の演奏では、楽譜情報以外に演奏者固有の表現が付加されるが、演奏される音の音高や順番等の音楽を形作る基本要素は、基本的にどの演奏者でも同じである。本研究では、こういった情報に加えて、演奏表現に関する楽譜情報も利用することで、情報の抽出を行っている。ここで、本システムにおいて用いている楽譜情報を以下に示す。

- ・音高（ピッチ）
- ・音符の長さ（音価）
- ・同時に演奏される音
- ・楽譜に現れる（演奏される）音の順番
- ・演奏表現記号（アクセント・スタッカート・フェルマータ・ペダル記号）

これらの楽譜情報を用いた具体的なトップダウン処理としては、以下のものがあげられる。

- ・楽譜から得られる、音の出現順の採譜処理
- ・音符の長さからの次音位置推測
- ・採譜された音列のフィードバックによる音の位置補正
- ・演奏表現記号による位置の優先順位の決定

3.2 音の立ち上がりと終了時間の抽出

音の立ち上がり時間の抽出に用いるTFマップは、FFTの結果に対し、楽譜からのピッチ情報を利用して特定の周波数帯でのピークピッキングを行うことで得られる。以下に、楽譜とTFマップの一例を示す。

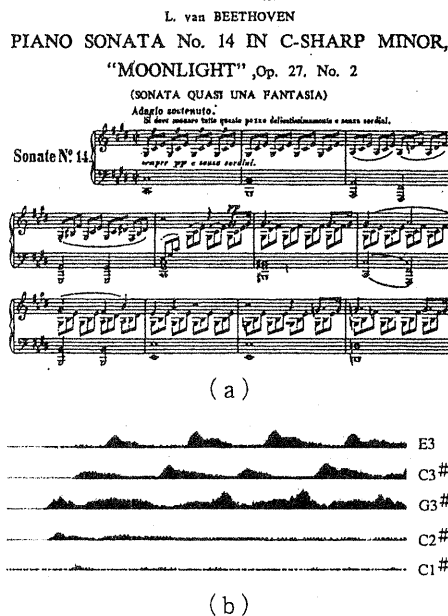


図3 周波数処理
(a) 楽譜 (b) TFマップ

本システムは、このTFマップを用いて、まず楽譜に現れる音の順番に音を探していく。ここでは、楽譜上で同じ拍位置にある音をまとめて1つのユニットとし、現在探しているユニット内の音が全て得られると次のユニット内の音を探す。この時、次に演奏される音（ユニット）の位置を推測しながら楽音の探索を行っていく。これは、楽譜に記された音符の長さを利用して図4に示す計算により行われる。但し、ユニッ

トの長さはそのユニット内に含まれる最も短い音符の長さである。

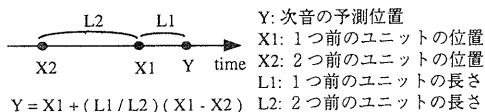


図4 次音の位置予測

また、これと同様の方法により、音符の長さを利用してすでに得た音の位置の補正を行う。これは、1フレーズごとに行われ、図5の計算により補正する音の候補位置を求める。

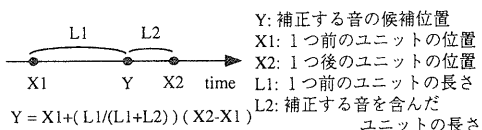


図5 音の位置補正

この候補位置の周辺で、すでに得た音の位置よりも候補位置に近い位置に、立ち上がりと考えられる点がある場合には、その位置を正しい音の位置であるとする。但し、アクセント記号、スタッカート記号がある音については立ち上がり時間が正確に得られることが多いため、音の補正において基準となる点と考え、補正はその他の音に対して優先的に行われる。

ここで、上記の2つの手法は音符の長さを利用しているが、音符の長さに関わらずフェルマータ記号がある場合には、演奏者はそれぞれ異なった間を置く。そのため、楽譜からフェルマータ記号を情報として取得しておき、これが音の立ち上がりは、予測位置より後にあると考え、また補正は行わずに情報抽出を行う。

以上の処理を利用して、最終的にはパワーの微分値を用いることにより詳細な立ち上がり時間を得る。

音の終了時間については、音の立ち上がり後のパワーの最大値からの微分値を調べ、その値が、ある threshold を下回る位置を音の終了時間とみなす。

3.3 音の強度情報の抽出

音の強度情報の抽出では、まず音の立ち上がり後のパワーの最大値（実験により約0.3[s]後で最大となることが分かっている）を取得する。

この後、本来ならば、パワーの最大値は同じ強さで打鍵しても音高により変化し、さらに個々のピアノの特性、録音の際の録音機器の設定によっても変化を生じるため、それに応じた補正を行う必要がある。しかしながら、入力として用いる演奏が成されたピアノの特性を調べることは困難であり、音響信号からその特性を得ることもほとんど不可能である。そこで、本システムでは演奏を再現する MIDI 楽器の音に対する FFT のパワーの特性を用いて補正を行う。これは、そのMIDI 楽器で、鍵盤の中央の音を C3としてC1からC5まで演奏させた音のパワー値を直線で近似することにより行う。

ピアノによってパワーの特性が異なることは上に述べた通りであるが、それに加えベロシティカーブもピアノによって異なる。そこで、上記の手法で大まかな補正を行った後、より適切な値を得るために本システムでは、音響信号のゲインを利用してベロシティの補正を行う。この手法の概要を図6に示す。

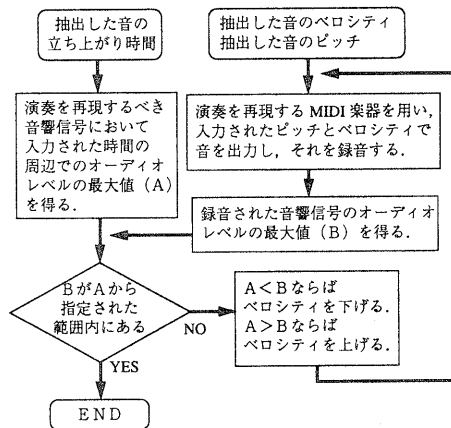


図6 ベロシティの補正

4. 雑音除去による演奏の復元

本来、演奏の復元を行う際には、アナログフィルタ、デジタルフィルタ（演算処理による雑音除去）、周波数領域での処理等を組み合わせることによって演奏を復元する必要があるが、本稿では、デジタルフィルタを用いたノイズリダクションのみを行っている。

まず、多項式適合による移動平均法を用いる

ことで全体的なノイズを大幅に押さえる。以下にその式を示す。

$$y(i) = \frac{1}{W} \sum_{j=-m}^m x(i+j)w(j)$$

$$W = \sum_{j=-m}^m w(j)$$

上式において、 $x(i)$ は音響信号のサンプル値、 $w(i)$ は重み関数の係数、 $y(i)$ が平滑値となる。重み関数は2・3次多項式適合のものを用いる。

次に、適応化平滑化法を適用することで、ベース部分の雑音を押さえる。以下に、その式を示す。

$$\hat{s}(i) = \frac{\partial_x^2(i) - \partial_n^2}{\partial_x^2(i)} \{x(i) - \bar{x}(i)\} + \bar{x}(i)$$

$$\bar{x}(i) = E\{x(i)\}$$

$$\partial_x^2(i) = E\{(x(i) - \bar{x}(i))^2\}$$

$x(i)$ は上記のものと同様であり、 $\hat{s}(i)$ は信号の推定値、 ∂_n^2 は雑音成分の分散である。雑音成分の分散は、音響信号の音の鳴っていない部分から取得する。

上記の2つのノイズリダクション処理を行うことで演奏の復元を行う。

5. 実験と考察

5.1 システム

図7に構築したシステムのユーザインタフェー

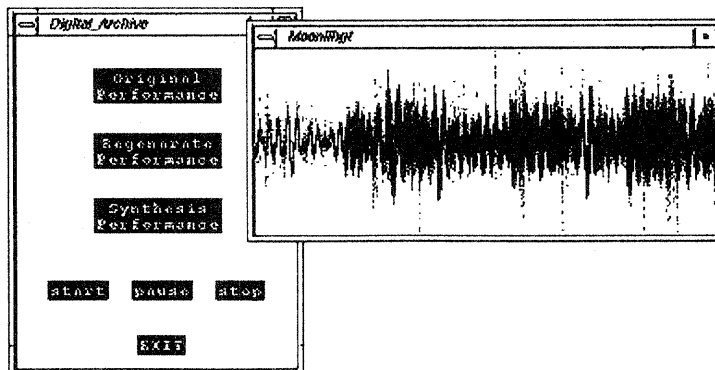


図7 デジタルアーカイブ化システム

スを示す。図のようなダイアログにより元の演奏、復元演奏、再現演奏を選択することが可能であり、選択した演奏を波形を見ながら鑑賞することが出来る。これにより、3つの演奏を比較することが可能であり、快適な鑑賞も可能となる。また、3つの演奏はすべてデジタル化された音響信号、あるいはMIDI信号であるため、ユーザは容易に波形編集や演奏の編集を行うことが可能である。

5.2 実験

演奏の再現と演奏の復元に対する評価を行うために、LPレコードに録音された、ベートーベン作曲によるピアノソナタ14番「月光」を用いてデジタルアーカイブ化を行った。図8にコンピュータに取り込んだ音響信号の波形を示す。

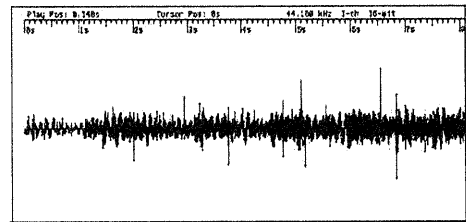


図8 「月光」の原波形

これは、44.1kHz、16bit、モノラルでコンピュータに取り込まれたものである。取り込みと本システムの処理は、Silicon Graphics Octaneで行う。また、演奏の再現に用いるMIDI楽器はYAMAHA Clavinova CLP-560である。

5.3 実験結果

再現演奏に対する結果として、元の演奏と再現演奏のトップノートの音の立ち上がり時間のずれを図9に示す。元の演奏におけるトップノートの立ち上がり時間は、波形と音を参考にして取得しており、再現演奏の音の立ち上がり時間は、FFTの窓幅と音高を考慮して計算している。

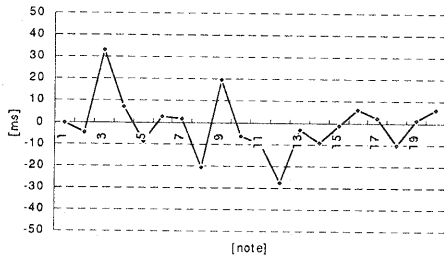


図9 音の立ち上がり時間のずれ

上図は、最初に演奏される音を同期させて計算しているが、この結果から、ほとんどの音のずれが+30[ms] ~ -30[ms]の間に入っていることが分かる。人間が、音のずれを認識できるのは、一般的に約10[ms] ~ 30[ms]のずれを生じた場合といわれている。そのため、この結果から、再現演奏と元の演奏の間で人間が感じ取ることが出来る程のずれは生じていないと言える。ただし、元の演奏の立ち上がり時間は、先述したように波形と音のみを参考に手動で行っているため、この結果には多少の誤差が含まれていると考えられる。また、再現演奏を聴く場合には、元の演奏との時間のずれが負から正、あるいは正から負へと急激に変化すると、相対的なずれにより人間は違和感を感じる可能性があるが、音量や音色の違いにより感じ方が異なる。そこで、定量的ではないが再現演奏を数人の音楽経験者に聴いてもらったところ、ほとんど違和感を感じられない、という結果が得られた。

復元演奏についても、再現演奏と同様に数人の音楽経験者に聴いてもらったところ、ノイズが軽減されて鑑賞し易くなった、という結果が得られた。但し、平滑化によって音がこもっているように感じられるという意見もあった。

5.4 考察

本システムは、デジタルアーカイブを行

うシステムとしては、まだプロトタイプの段階にすぎない。しかし、本システムで用いた演奏情報の抽出手法は、ノイズを含んだ音響信号からの演奏情報の抽出に有効であると考えられる。また、DPマッチングを用いた演奏情報の抽出手法^{[1],[2]}に対し、本手法では局所的な演奏速度を考慮した予測が可能であるため、演奏速度が大きく変化するような楽曲に対しても適切な演奏情報の抽出が可能であると考えられる。これについては今後実験を行っていく予定である。

6. おわりに

本稿では、古い演奏メディアをデジタルアーカイブ化するための手法と、実際に構築したデジタルアーカイブ化システムについて述べた。貴重な音楽資料を利用可能な形で後世に残すためには、デジタルアーカイブ化は必要不可欠であり、多くの人が望んでいることである。本システムは、まだ不完全ではあるが、それを実現するための指針を示したと言えるであろう。

今後、今よりもクオリティの高いデジタルアーカイブ化を行うために、より詳細な演奏表現の抽出を行い演奏を再現する手法や演奏の復元法について考えていきたい。また、抽出した演奏表現は、演奏家モデルの構築や音楽理論での利用が可能であると考えられるため、これについても検討していきたいと考えている。

参考文献

- [1] 高見啓史, 片寄晴弘, 井口征士: ピアノ演奏における演奏情報の抽出, 電子情報通信学会論文誌, Vol.J72-D2, No.6, pp.917-926 (1989).
- [2] 高見啓史, 片寄晴弘, 井口征士: DPマッチングを用いた演奏情報の抽出, 昭63音響学秋期論文集, 1-4-9.
- [3] 片寄晴弘, 井口征士: 知的採譜システム, 人工知能学会誌, Vol.5, No.1, pp.59-66 (1990).
- [4] 片寄晴弘: 自動採譜 (概論), コンピュータと音楽の世界, bit別冊, pp.74-88, 共立出版, (1998).
- [5] 魚住純, 牛坂健, 伊福部達: 古ろう管からの名演の復活, 日本音響学会誌, Vol.54, No.6, pp.443-449 (1998).
- [6] 中村勲, 難波精一郎: ピアノの音色はタッチによって変化するか, 日本音響学会誌, Vol.52, No.5, pp.349-355 (1996).