

コンピュータによる音楽データ処理

山本 順人
(国立民族学博物館)

1. はじめに

近年、音楽の分野において、従来のヨーロッパを中心とした音楽以外の音楽、民族音楽の研究が盛んになってきている。ヨーロッパ音楽は永い研究にうまげられた音楽理論を持ち、記譜法や様式分析などが確立してきていると言える。一方、非ヨーロッパ系の音楽においては、その研究の歴史も浅く、また、音楽自体も、固有の記譜法で書かれていたり、得られるデータとしては演奏の録音のみである場合が多い。比較音楽学の立場からは、対象とする部族、民族、地域などが、どのような音楽様式上の特徴を持つかに興味を持たれる。このような様式の分析をおこなう時、音楽情報に関するデータベースの必要が生じてくる。データベースの構築には、多くの原データが音響的記録として得られるため、よく訓練された聴覚により人手で採譜が行われてきている。しかし、これら民族音楽の特徴としての、微妙に変化する旋律やリズムの構造が、重要な情報を背負っていたりする場合もあり、客観的なデータの抽出に困難を生じている。

そこで、これらの困難を減ずるべく、コンピュータにより採譜を行い、データベースを構築し、分析処理に供するとともに、その結果を音楽的にわかりよい形式で出力する事を試みた。本システムは、音響的に音楽を入力する事により、自動採譜し、音符コードとして蓄積格納する。また、シンセサイザーを自動演奏させる事により、分析結果を音として出力し直接聴覚に働きかけ、レーザビームプロインタにより、楽譜形式で印刷出力を得る事ができる。

2. システム構成

本システムは、図1に示されるごとく、複数の計算機の結合した形として構成されており、それぞれの機能を分散処理している。

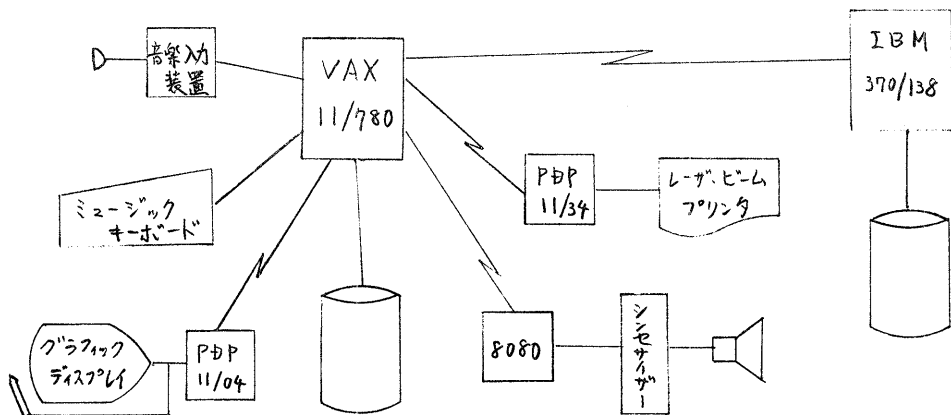


図1. システム構成

入力系としては、音楽入力装置、ミュージックキーボードを持っている。音楽入力装置は、マイク、録音機等からのオーディオ信号を受け入れ、各時点における音高、音量データを出力してゆく。ミュージックキーボードは、ピアノ等の鍵盤に類似した配置を持つ音高キー、音価キー、制御キーより成る端末である。

蓄積系としては、VAXの外部記憶装置を一次的な格納に、IBMの外部記憶装置を二次的な格納に用いる。

出力系としては、音響的な出力をシンセサイザーでおこない、プリントイメージとして楽譜をレーザビームプリンタにより印刷する。また、グラフィック端末では、インタラクティブな分析処理を行うためにグラフィカルな表示を得ている。

3. 音符コード

音楽を表記するためには、数字譜や五線譜、また、邦楽にもちいられている譜など、さまざまな形式が考え出され用いられて来ている。これらの中で、最も一般的に使用され、表現能力に豊んでいるのは五線譜である。本システムも、この五線譜を基本にした音楽の取り扱いを行っている。五線譜には、音高・音価による旋律、音の強弱、表情、ニュアンスなどが表現されているが、その中で旋律情報に主眼を置いた分析を行う事を目的とし、音高・音価によるコード化をはかった。システム内では、この音符コードにより、分析入出力処理等の統一した取り扱いをおこなっている。

図2は、12音音階における音高と、対応するコード(P)を表わしている。また各音は、A₄(=33)を440Hzとする絶対音高により表現される。

表1は、音価とそれに対応するコード(V)である。

このP, Vの組により、旋律を構成している各音を表現してゆく。(図3)

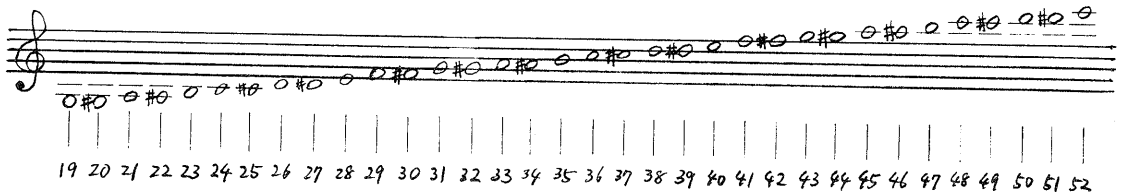


図2. 音高コード(P)

表1. 音価コード(V)

○	A	○.	H	○..	D
♩	B	♩.	I	♩..	P
♪	C	♪.	J	♪..	Q
♫	D	♫.	K	♫..	R
♩	E	♩.	L	♩..	S
♪	F	♪.	M	♪..	T
♫	G	♫.	N	♫..	U

(休符も同コード)



31C 28D 31D 36B 33C 36D 36D 31B

図3. 音符コードによるコーディング例

4. 採譜

音信号から旋律情報を取り出す方法には、フィルタ列による弁別や、フーリエ変換によるスペクトル分析などが考えられる。フィルタ列による方法では、各フィルタが狭帯域の通過特性を有せねばならず、工学的に実現するには困難がともなう。フーリエ変換により基本周波数をもとめる方法は、最も確実であると考えられる。しかし、短時間に生じる過渡的現象の場合と異なり、数分ないし数十分にわたり演奏される音楽に付しては、取り扱わねばならないデータ量が膨大であり、また、フーリエ演算にも長時間を要す。しかも、旋律やリズムの変化は、サンプリング周期に比して充分遅いものである。

そこで、本システムでは、すべてをデジタル方式で取り扱うのではなく、アナログ方式により音高・音価を抽出し、以後の処理をデジタル的におこなう、ハイブリッド方式としている。

・ピッチ抽出

オーディオ信号として入力された音楽は、まず、ローパスフィルタにより高調波成分を除去される。これにより得られた波形を整形し、方形波に変換する。この方形波の立ち上り t_u で、対象波形発生回路をスタートさせ、立ち下り t_d でストップ、ホールドさせる。(図4) 今、 t_e は、対象信号のその時点での周期となっているので、この時に得られる P の値がもとめる音高となる。このようにして、連続的に取り出す事によりピッチ波形を得ている。

・データ転送

上述の如く得られたピッチ波形および音量波形をA/D変換して、計算機に送出する。各時点での音高、音量データは、それぞれ12ビットで表現されており、4バイトに分割しバッファメモリに蓄積される。転送は9600ボーラインを用い、200ms毎に、128バイト単位でおこなわれる。この様にする事により、転送から転送までの間を利用し、ディスク装置へデータを格納するソフトウェアの動作時間を得ている。

・スムージング

得られた原波形データには、微細な変化が多く含まれているので、波形のスムージングを行う。スムージングには、時間的に平均してゆく方法と、クリッピングを行う方法と二方法を用いている。

時間平均法は、現時点の値を先立つ数時点の平均値として決定する方法である。

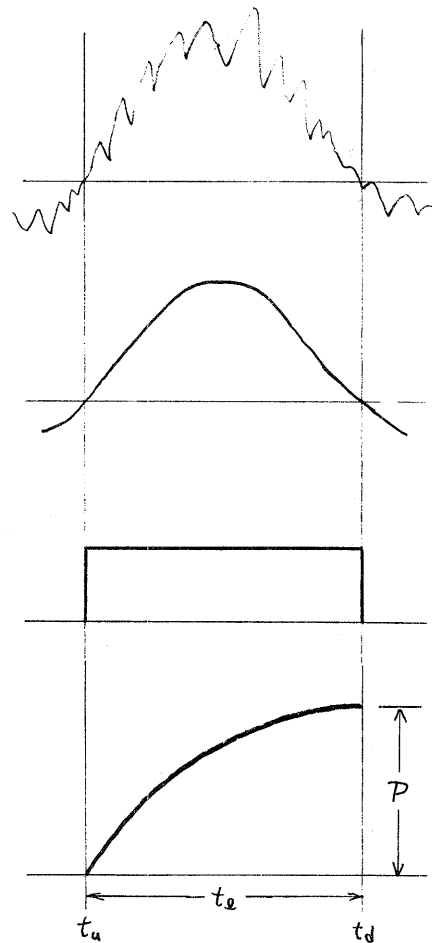


図4. ピッチ抽出

$$V_t = \frac{1}{n} \cdot \sum_{i=1}^n V_{t-n}$$

で算出される。この場合、平均する時間の長短により、図5に示される様なスムージング結果を得る。平均化時間が長くなると、除去したい成分のみならず、音高、音量の立ち上り、立ち下り変化をも弱めてしまう。この結果、本来は、ジャンプしている二音の間にみせかけの経過音が存在する事になってしまう。

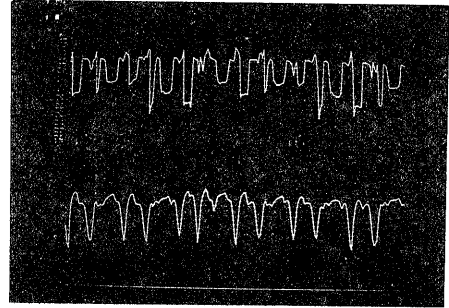
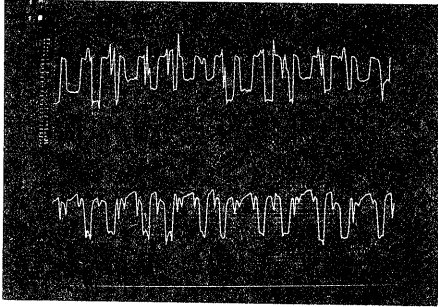


図5. 平均法によるスムージング

クリッピング法は、短時間に起こるパルス性の変化成分を除去し、音高、音量変化は保存する方法である。

for $j=1, \dots, n-1$

$$V_{t+j} = \begin{cases} V_t & \text{if } V_{t+j} > V_t \text{ and } V_{t+n} \leq V_t \text{ and } n < T \\ V_{t+j} & \text{otherwise} \end{cases}$$

T : constant

で算出される。パルス性の変化成分は、楽器音では比較的少く、不安定な人の歌声、特にピッチの定まらない子音部に多く発生する。図6は歌声の原データと、クリッピング後の波形を示している。

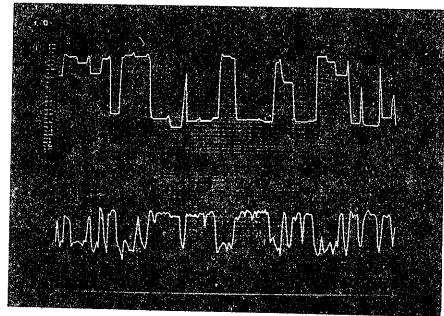
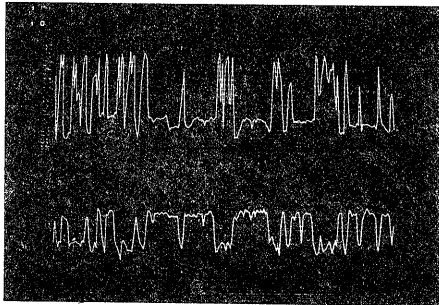


図6. クリッピング法によるスムージング

・音の同定

スムージングにより得られた音高波形、音量波形により、音高、音価を決定し音の同定を行う。音高は $A_4 = 440 \text{ Hz}$ とした平均律を用いている。

音高データが、各音高に相当する閾値内に連続的に存在する時、それらのデータ列は、一つの音を表わしている可能性がある。また、音量波形が急激に変化する所は、音と音との境目を示しており、閾値以下の場合、休符の存在を示していると考えられる。(図7)

まず 音高波形より、各音高の閾値を通過する時点をもとめ、さらに必要に依

じて、音量波形より、変化時間をもとめれば、これらの時点で区切られた区間が、その相等する音高を持つ音の音価となる。この様にして、各音の音高・音価が決定され、音符コードに記号化される。

。採譜操作

採譜は、端末装置を介して、インタラクティブにおこなわれる。採譜操作には、

- 。演奏時間
- 。速度規準
- 。サプレスレベル値
- 。最短有効音価値

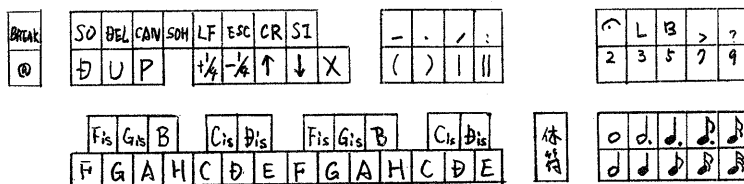
を指示する。速度規準は、音の絶対的な時間と採譜の基本音符を関係づける。通常の五線譜の記譜においては、リズムパターンや視覚的見地より、拍の基本となる音符が決定されている。そして、*Andante*, *vivace* などの表現で、実時間の音と関係づけられる。本指示は、この速度記号の逆操作に相当するものである。サプレスレベル値は、音信号と雑音とを弁別するレベルの設定にもちいられる。また、最短有効音価値は、楽譜表現された時に必要とみなされる最短値をあたえる。節まわしなどを採譜する場合は短く設定し、微細な音の動きをとらえ、旋律の大まかな構造をもとめる時は長く設定される。

。採譜例

上述の操作により、楽器音を採譜した例を、図8, 9, 10に示す。

5. ミュージック・キーボード

音符を入力するためのキー配列をもつ鍵盤も、音楽入力の方法として用いている。この鍵盤は、ピアノ等の鍵盤楽器と同型の並びを持つ音高キーと、音価を表わすキー、および制御キーにより成り立っている。音高キーは、基本としては2オクターブの幅を持っている。そしてオルガンにおけるストップの様に、UPシフト、DOWNシフトキーを並用する事により、高音域、低音域を表現する事を可能にしている。(図11)



6. 記譜

計算機内部で用いられている音符コードは、計算機処理には適しているが、楽譜としてのイメージからかけ離れている。そのため、理解する人間側に負担を加え、マンマシンコミュニケーションの立場からは適当ではない。

そこで、従来からの表現形である五線譜を用いる事により、理解しやすい形の出力としている。印刷はレーザービームプリンタを用いておこなっており、プリンタは、ネットワークオペレーションにより、本体計算機で制御される。

印刷の手続を以下に記す。

i) 計算機の主記憶上に、イメージパッドを準備する。

印刷のためのプリントイメージ展開は、本体計算機上で行っており、展開されたイメージを、プリンタに送出する方式をとっている。パッドとしては、1728 x 2048のビットプレーンとしており、A4紙の大きさに相当させている。

ii) 五線に相当するパターンを書き加える。

iii) 音符のパターンおよび位置を決定する。

楽譜は、五線一行に音(休)符17位置を持ち、8行1頁としている。

各音符は、音符コードを元にして、音符テーブルを引く事により、印刷コードに変換され取り扱われる。印刷コードは、

(パターン番号、楽譜上の位置、五線上の位置、臨時記号の有無種別)

の形であらわされている。

iv) 五線上に音符パターンを書く。

iii) で得られた位置に音符コードに相当するパターンを書く。この時、音符パターンの向き、上下加線、#・bなどの臨時記号が判断され、書き加えられる。

v) プリンタシステムにイメージを転送する。

イメージパッド上に出来上がった印刷イメージをプリンタシステムに転送する。計算機間は、ネットワーク機能によりタスク対タスクで結合されており、印刷機が動作可能かどうかを交信後、2048バイト単位でデータ転送をおこなっている。通信路としては、1Mボー/秒、半二重方式を用いている。転送データは、プリンタシステムのイメージメモリーに順に格納される。

vi) プリンタを起動し、印刷する。

転送されたデータを所定のメモリーに格納した後、プリンタのモードセットや紙送り等を指示し、実行を開始させる。プリンタは、逐次メモリーを読み出す事により楽譜を印刷する。

図12は、印刷出力例である。



図.12. 印刷出力例 (部分)

7. 演奏

格納されているデータや分析結果を、より直接的に、音として出力できれば、結果の理解や把握が容易となる。このために、計算機制御により音の発生をおこなっている。音を作り出す装置としては、シンセサイザーを用いている。

民族音楽では、 $\frac{1}{2}$ 音階のみならず、 $\frac{1}{4}$ 音階、 $\frac{1}{8}$ 音階などでなりたっているものも多く、また、こぶしなどに代表される細かな音の動きが重要なニュアンスとなり得る。シンセサイザーは、

- ・連続的に無段階の音高を作り出す事ができる。
- ・電圧制御の形式であるため、制御が容易。

などの特徴を有しているため、上述の再生に適している。

このシンセサイザーは、附属するコントローラに、

1. 音高、音量等の制御データ (CV)
2. ステップ時間データ (ST)
3. ゲート時間データ (GT)

を送出する事により制御される。(図13)

シンセサイザーの発振器は、印加電圧により、その発振周波数を変化させる事ができる構造になっている。この電圧を変化させ、必要音高を得る。一方、音価は、ステップ時間データによりあたえられ、発振器の後段に置かれた増幅器の、OFFからONへのタイミングにより作り出される。また、スタカートやテヌートなどの表現は、このステップ時間内での増幅器ONの時間の割合により作りだし、ゲート時間データとしてあたえられる。(図14)

音符コードより変換算出された上記データは、図15の形式で、300ボーで結合されたコントローラ(マイクロプロセッサ8080)に送出される。転送されたデータは、コントローラ内部のメモリーに格納されており、制御プログラムにより、A/D変換され、制御電圧が得られている。

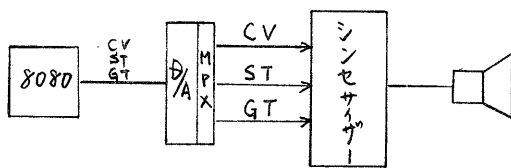


図.13. シンセサイザーの制御

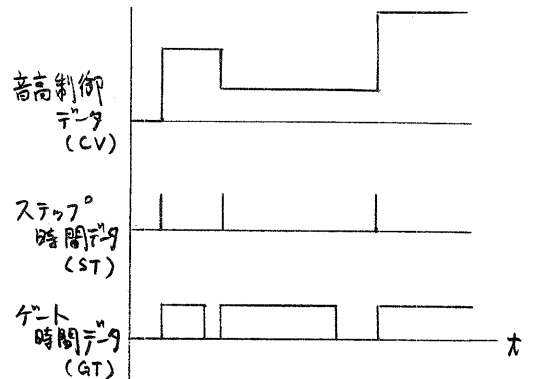


図.14. CV, ST, GTの時間チャート

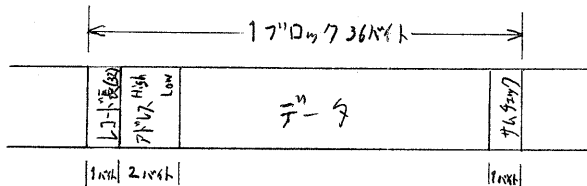


図.15. 転送フォーマット

8. おわりに

民族音楽研究を支援する事を目的としたシステムにつき述べた。本システムでは、音による音楽の入力、音および譜による出力、旋律情報の蓄積、分析をおこなう事ができる。

特に採譜に関しては、取り出されるデータ量をなるべく小さくおさえるために、ハイブリッド方式とし、長時間演奏に対応した。また、採譜手続きを、演奏時間のオーダーで実行する事ができた。

しかし、本システムでは単旋律の音楽をとり扱うことができるが、複旋律や和音の採譜は、今後の問題として残されている。また、同様に、

1. 採譜時に、より精密な情報や補間のための情報を抽出する。
2. 演奏にアゴークグが生じている時、そのテンポに同期した採譜を行うか、現象記録の立場から、時間で割り出した音符をもちい、時間的にずれは、そのまゝ記録するかを選択決定。
3. 採譜対象として、楽器音は比較的安定した音であるが、歌声は、音高音量ともに非常に不安定であり、厳密な意味においてピッチの決定に困難がともなう。
4. 記譜に関しては、たが五線の形式のみを用いているため、表現上のみやすさが必ずしも充分ではない。
5. 音色情報の取り扱い。

などが残されており、引き続き考察を重ねてゆきたい。

最後に、音楽面から国立民族学博物館藤井知昭助教授、情報工学面から、京都大学長尾真教授に御助言、御検討いただいた。また、筑波大学中山和彦教授に、レーザビームプリンタ印刷技術の御助力をいただいた。ここに記し、感謝の意を表します。

参考文献

- Anderson, V.C., Melograph Model B/T, Operation & Maintenance Manual, Interocean Systems Inc. San Diego, 1975.
- Buxon, W., et al., The Use of Hierarchy and Instance in a Data Structure for Computer Music (Computer Music Journal, Vol.2, No.4, 1978)
- Lincoln, H.B., コンピュータと音楽, カワイ楽譜, 1970.
- Moorer, J.A., Signal Processing Aspects of Computer Music - A survey (Computer Music Journal, Vol.1, No.1, 1977)
- Moorer, J.A., On the Transcription of Musical Sound by Computer (Computer Music Journal, Vol.1, No.4, 1977)
- Taylor, H., SCORTOS: Implementation of a Music Language (BYTE, Vol.2, No.9, 1977)