

Chapter 5

オーディオ処理

佐近田展康

本章ではMax/MSPによるオーディオ処理について解説を行う。コンピューターの処理能力の向上と、MSPのリリースにより、Maxは音そのものをプログラミングの対象にできるようになった。これはコンピューター音楽家やサウンド・クリエイターにとって長年の夢であり、最近になってようやく踏み込むことができるようになった待望の領域でもある。オーディオ処理プログラミングでは、デジタル・オーディオについて、よく知り、よく理解することが求められる。したがって、どうしても理論的な説明とその学習を避けて通るわけにはいかない。本章では、ここで必要となる知識や理論、技術をMax/MSPによるプログラミングを通じて具体的に紹介する方法をとる。さらに深く理解したい読者はデジタル信号処理関連の書籍など……特に、コンピューター音楽に関する最も網羅的で専門的な日本語文献であるカーチス・ローズ(Curtis Roads)著の『コンピューター音楽—歴史・テクノロジー・アート—』(東京電気大学出版局)を参考にしていきたい。

5-1 デジタル・オーディオの基礎

音とは何だろうか？ Maxを使う目的は、作曲するため、ライブで活用するため、斬新なサウンドをクリエイトするため……人それぞれの目的があるだろう。しかし、それはすべて最終的に物理現象としての音、つまり実際に耳に聞こえる音に関わる。Chapter 5で取り扱うオーディオ処理は、音のデジタル表現、つまりデジタル・オーディオに関する処理のことである。このデジタル・オーディオは、物理現象としての音をデジタルに置き換えたもので、もともと音が持っている法則性や特徴を引き継いでいる。したがって物理現象としての音について、最低限の知識は持つておく必要がある。加えて、デジタル・オーディオが物理現象としての音とどのように違うのか、どのような特殊性を持つのかについても理解が必要だ。

ここでは、音とは何かから始めて、デジタル・オーディオの諸特徴まで、その概説を行う。物理の用語、耳慣れない用語が多く出てくるかもしれないが、本章で紹介するMax/MSPによるオーディオ処理を理解するためには、しっかりと把握してほしい。

音とは何か？

何かの物体が振動すると、その物体は周囲の空気を押ししたり引いたりする。ここに微細な空気圧の変化が生じ、次から次へと周囲の空気へ伝えられる。

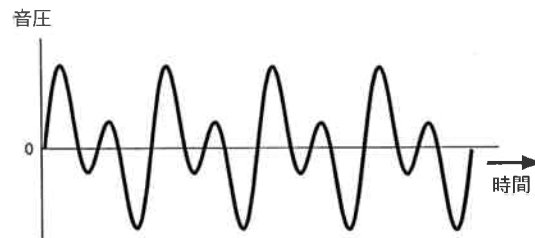
■5-1-1 振動が波として空気を伝える



圧力の変化が連続して伝えられる状態を一般的な言葉で言えば波である。ここで波として伝えられる空気圧の変化は、何かの物体に当たるとそれを押ししたり引いたりしようとする。もしその物体が巨大な岩だったら、微細な空気圧の変化くらいではビクともしないだろう。しかし、もし人の耳の鼓膜だったら、鼓膜は空気圧により押されたり引っ張られたりして振動する。この振動が神経電流の変化となり脳で知覚されたもの、それが音である。つまり音とは最初の物体(音源)の振動が、波(音波)として空气中を伝わり、耳の鼓膜に圧力(音圧)をかけて振動させ、知覚されたものなのである。

この音圧の変化と時間の関係を図で表すと5-1-2のように描くことができる。ここに描かれたグラフは、結局は音波を図示したものと同一であり、音の波形と呼ばれる。

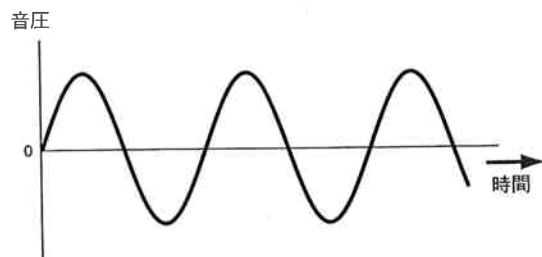
■5-1-2 音の波形



① 周期的波形と非周期的波形

5-1-2の音の波形をよく見ると、一定の同じ形が繰り返されているのが分かる。この同じパターンを繰り返す波形を周期的波形と呼ぶが、こうした音を聞くと、人は音の高さを感じる。楽器の音や歌声は、おおむね周期的な波形になっており、それで人は音程を感じたりメロディを聞き分けたりできるのだ。周期的波形のうちで、最も単純な波形は、同じ形の山と谷が規則正しく繰り返されるだけのサイン波である。具体的には、音叉の音や、電話の受話器を上げたときに聞こえるツーンという電子音が、サイン波に近い波形を持つ。

■5-1-3 サイン波の波形



一方、繰り返しのパターンがどうしても見つけられない音の波形もある。これは非周期的波形と呼ばれ、耳に聞こえる音は、音の高さが感じられないノイズ(雑音)になる。

■5-1-4 非周期的波形の例



この散らばり方が最も激しいノイズを、特にホワイト・ノイズと呼ぶ。具体的にはFMラジオから放送以外に聞こえるノイズがホワイト・ノイズに近い。

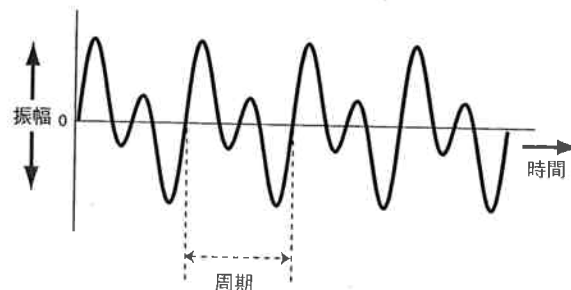
現実には聞こえる無数のさまざまな音は、このサイン波とホワイト・ノイズという両極端の間に存在すると考えることができる。

② 音の時間領域表現と波形

先ほどの例にもある通り、音を視覚的に表現する場合、縦軸に音圧、横軸に時間の流れをとると、波形のグラフを描くことができる。こうした表現方法を、音の“時間領域表現”と呼ぶ。周期的波形の場合、同じ波形パターンが繰り返されるわけだが、その繰り返し1回分を1周期と呼び、1周期の長さを波長あるいは周期と言う。1秒間に繰り返される周期の数は周波数と呼ばれ、ヘルツ(Hz)という単位で表される。周波数は音の高さを示す値として非常に重要だ。音楽用語で、音高はドレミファソラシといった音名で表されるが、例えば中央ラ(A)がいったいどれくらいの音高なのかを示すピッチは、最終的に周波数の値で定義される。中央ラ(A)のピッチは国際的に440Hz(場合により442Hz)と決められている。

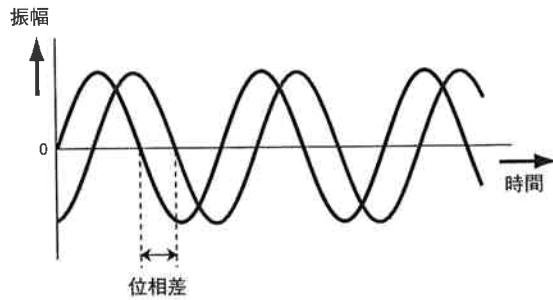
音の時間領域表現において、縦軸は音圧だといったが、波形に関する用語で言えば、波の変化の大きさは振幅と呼ばれる。音圧にしろ振幅にしろ、その大小は聴感上の音の大きさを左右する。

■5-1-5 音の時間領域表現



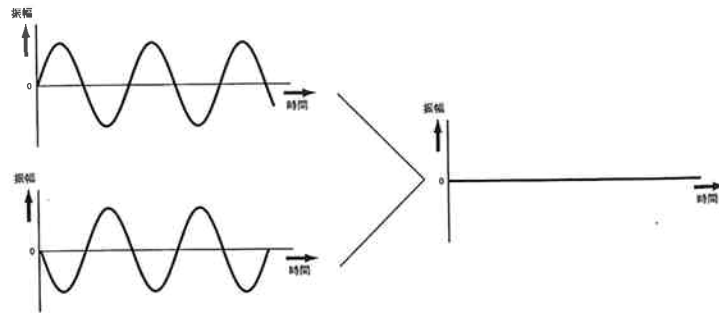
もう1つ、周期的波形を考える場合に重要なのが位相である。5-1-6のように波形を左右に並行移動することを位相をずらすといい、ずれの大きさを位相差と呼ぶ。位相をずらしても、周波数、周期には何も影響を与えない。

■5-1-6 位相をずらす



しかし、周波数が同じ2つのサイン波を同時に鳴らすとき、5-1-7のようにそれぞれの波形が半周期ずれていると、山と谷の部分が互いに打ち消し合って、実際には音が聞こえない、あるいは聞こえにくいといった事態がありえる。スタジオやステージなどの音響現場ではこの現象がトラブルの原因になる場合がある。また、逆に位相のずれを積極的に音作りに利用する可能性もある。

■5-1-7 半周期位相がずれた場合の打ち消し合い

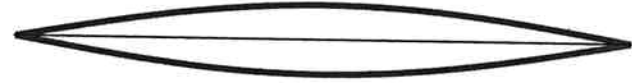


2 部分音と音色

音の高さは周波数として表現でき、音の大きさは音圧変化の大きさとして表現できる。では音色はどのように表現できるのだろうか。ギターや人間の声ではその波形が異なる。波形の違いが音色の違いとなって聞こえるわけだ。では、その波形の違いは何によってもたらされるのだろうか。

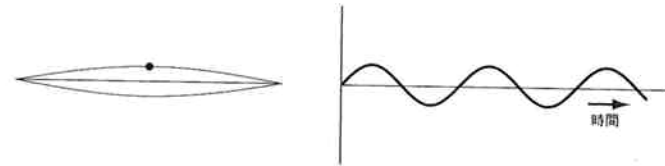
ここに1本の弦が両端を固定されてピンと張られているとする。これをギターを弾くのと同じように指で弾くと弦は一定の周期で振動し、それが空気中を音波として伝わって耳の鼓膜を振動させ音が聞こえるだろう。この音を基音といい、その周波数を基本周波数と呼ぶ。

■5-1-8 弦を弾いた振動



この弦のある部分に印を付け、それが時間とともにどのように上下に運動するのかを図示すると、音の時間領域表現と同様の周期的波形を描くことができる。

■5-1-9 弦の振動の時間領域表現

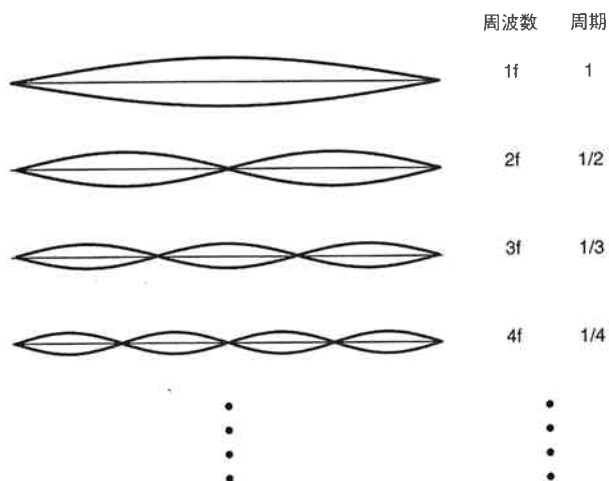


tableが差外

この波形はサイン波であり、これが音波として空气中を伝わって耳に聞こえたときには、サイン波の音として聞こえるはずだ。しかしギターの音色は、電話のツーン音とは異なっている。

実際の弦の振動は、5-1-8の大きな振動以外に、目には見えなくても次のように長さを2、3、4、5…… n 等分した細かな振動を多少とも含んでいる。弦は両端が固定されているので、 n は必ず整数になるだろう。そして、それぞれの振動の周波数は、基本周波数の2、3、4、5…… n 倍になる。つまり、弦を弾いて聞こえる音は基音だけでなく、多数の異なる周波数の音も一緒に聞こえていることになる。これらの音を部分音、あるいは倍音と呼び、このバランスが音色を決める。

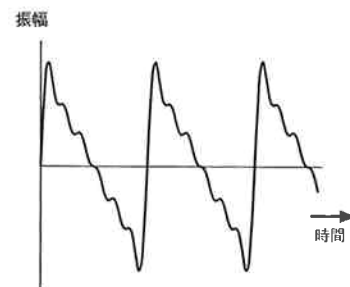
■5-1-10 弦の振動に含まれる細かな振動と周波数



部分音の波形は、基音の波形に対して周期が $1/2$ 、 $1/3$ 、 $1/4$ …… $1/n$ のサイン波だ。つまり音の波形は、周期の異なる複数のサイン波が混ざり合ったものだと言える。5-1-11は、周波数と振幅の異なる複数の部分音(サイン波)が混ざり合った波形の例だ。

エンベロープ

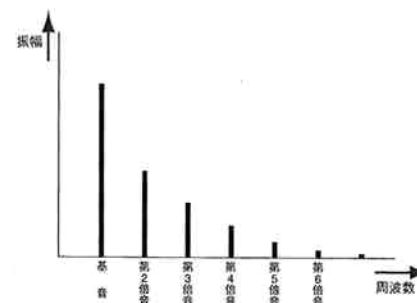
■5-1-11 サイン波が混ざり合った波形



音の周波数領域表現とスペクトル

ところで、時間領域表現とは異なり、基音および部分音の音圧の大きさを縦軸に、その周波数を横軸にとると、音を表現するもう1つのグラフが描ける。これを音の“周波数領域表現”といい、ここで描かれるグラフをスペクトル(あるいはスペクトラム)と呼ぶ。これはある瞬間をつかまえて、その一瞬の音がどのような部分音から構成されているかを示したものである。部分音は倍音とも呼ばれるが、それぞれ周波数の低い方から第2、第3、第4、第5……第 n 倍音となる。

■5-1-12 音の周波数領域表現(スペクトル)



先ほどの弦の振動では、各部分音の周波数は基本周波数の整数倍だった。これを特に整数倍音といい、楽器音色の多くは整数倍音を多く含んでいるので、濁りやうねりの少ない澄んだ音色として聞こえる。広く世界のさまざまな地域の伝統音楽では、古くからこの整数倍音に基づいた音階を見だし、音楽の基礎にしている例が多い。

一方、基本周波数と各部分音の周波数が整数倍でない場合、例えば2.15倍や9.77倍といった場合、これらは非整数倍音と呼ばれる。金属的な音色、音程感の明瞭でない音色などは、非整数倍音を多く含んでいる。

純粋なサイン波の音を除いて、われわれが耳にする音は部分音を含んだ複合音である。これは音の構造を分析してその特徴を知る上で重要なだけでなく、本章で解説していく音響合成にとっても非常に大切なポイントになる。

1 音の大きさとデシベル値

人間にとって聞こえる音の大きさ(ラウドネス)は、多くの要因が作用し合った複雑なものである。まだ誰も客の来ないパーティ会場で、BGMの準備のためにCDをかけたとき。アンプのボリュームをちょうど適当な音量だと感じるものに調整しておいたはずなのに、客が集まりはじめるとその音量では物足りなくなり、ボリュームを上げるかもしれない。客が帰ってしまうと、今度はBGMがうるさく感じられるだろう。つまり、音の大きさは、ある音の大きさと別の音の大きさを比較することで成り立っている。また周波数によっても音の大きさの感じ方は変わってくる。例えば、小さな音量では低い周波数帯の音はいっそう小さく聞こえる。

人間の感覚は、ある音の音量が10倍になったときの増加量と、10倍から100倍になったときの増加量を同じように感じるといわれる。これは対数的な関係であり、音の大きさである音圧レベルは、

$$20 \times \log_{10} \left(\frac{V_2}{V_1} \right)$$

$$20 \times \log_{10} (\text{ある音圧} / \text{基準音圧}) \quad \begin{matrix} \downarrow \\ 20 \times \log_{10} 2 = 20 \times \log 2 \\ = 20 \times 0.3010 \\ = 6 \text{ (dB)} \end{matrix}$$

という式で表され、その単位はデシベル (dB) になる。基準音圧は通常は耳に聞こえる最小音量の音圧を指すが、ここで重要なのは2つの音圧の比で音圧レベルを考えている点だ。ある音源、音響機器が出せる最小音圧と最大音圧の比をダイナミック・レン

ジと言うが、フル・オーケストラの場合およそ100dB、CD再生装置で約96dBになる。

デシベル値は、実際の音響機器を扱う場合に頻繁に目にする値になる。例えばオーディオ・ミキサーは数多くの入力端子から入ってくるオーディオ信号の音量を調整し、ミックスして出力する装置だ。この音量調整のために多くのフェーダーと呼ばれるスライダーが付いており、フェーダーを上げると、音量が大きくなる。通常、フェーダーには目盛りが振られていて、これらの値はデシベル値を示している。

ここで現在聞こえている音の音圧を聴感上で2倍にする状況を考えて、先ほどの式に当てはめて、フェーダーを現在の位置からおよそ6dB分上げる。1/2倍にするには6dB分下げ、1/4倍にするには、6+6=12dB分フェーダーを下げればよい。

このように音の大きさについてはそれが比較に基づいており、しかも大きく人間の感覚に依存するものであるため、音の高さを表す周波数のように厳密な定義は難しい。

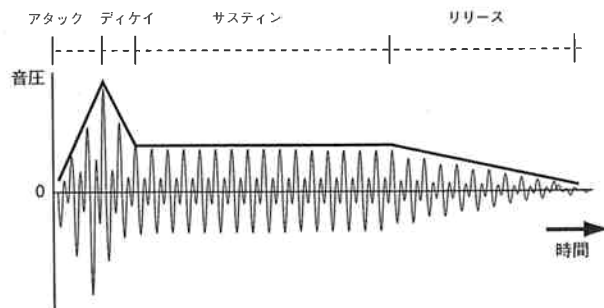
2 音の時間的变化とエンベロープ

音は刻々と変化する現象である。スペクトルはある特定の瞬間の部分音の構成について教えてくれるが、別の瞬間にはスペクトルの形は変わっているだろう。この刻々と変化する特性は、実際に音響合成を行う際に、極めて重要なものだということが分かる。例えばバイオリンの音の波形から、ある1周期分を取り出し、それを単純に繰り返す周期的波形を作ったとする。この音を実際に聞いてみると、単調で、原音とはほど遠いものとして聞こえるだろう。これはまったく同じ波形が繰り返されていることによる単調さのせいで、逆に言えば、現実の音は刻々と波形を変化させていることになる。

音量についても同様で、例えば打楽器の音は、たたいた直後から次第に小さくなって自然に聞こえなくなる。一方バイオリンの音は弓を一定の力で引き続けている限り持続する。

ある一定時間の波形の振幅の最大値(ピーク)を線で結んでいくと、その音の時間的な音量変化について、おおよその形をつかむことができる。この形をエンベロープと呼ぶ。各種の楽器は、その楽器特有のエンベロープを持っており、例えばフルートの波形にピアノのエンベロープを付けると、その音はもはやフルートには聞こえない。エンベロープはその特徴的な部分に名前が付けられている。楽器音の判別には特にアタック部分のエンベロープの形が重要になる。

■5-1-13 エンベロープと各部の名称

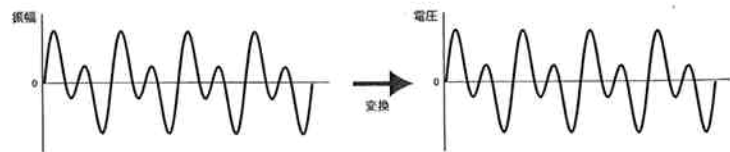


音のデジタル表現

もともと物体の振動だったものが、空気の圧力変化に形を変え、鼓膜の振動に形を変え、神経電流の変化に形を変え……さまざまな変換を経て音は知覚される。このプロセスに機械が参加してくると、また別の変換が行われる。例えば、楽器の音をマイクで拾い、それをスピーカーから拡声して聞く、といった状況を想像してほしい。

ここでは空気の圧力変化としての音は、マイクによって電圧の変化へと変換される。つまり音波は電気信号に変換されるわけだ。

■5-1-14 音波の電圧への変換



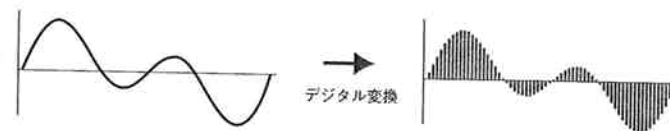
変化する電圧は、ある瞬間だけを捉えると1個の数値(xボルト)として表すことができる。また別の瞬間には違った数値(yボルト)として表すことができるだろう。この2つの数値の間はどうなっているかといえば、連続した滑らかな線で結ばれ、ジャンプしたりはしない。いかに短時間の激しい変化であっても、その変化は数値から数値へと跳び移るようなものではなく、滑らかに連続したものなのだ。電圧に限らず、振動、波、空気の圧力変化、およそ物理的な現象は、すべて“連続して変化する量”という共通した特徴を持っている。耳慣れたアナログという言葉は、こうした特徴に対して付けられた名前だ。したがって、電気信号はアナログ信号であり、電気信号に変換された音は、音のアナログ表現と呼ばれる。

一方、アナログに対立するものとしてデジタルという言葉がある。これはもともと指を表す言葉から派生し、指を折って数を数える様子を示している。数はそれぞれがバラバラで独立しており(離散的と言う)、1個ずつ数えたり、取り出したり、入れ替えたりすることができる。アナログvsデジタルは、連続vs離散、量vs数、測るvs数えるといった違いとして考えることができる。

アナログ信号とデジタル信号もまさに同じような違いを持っている。アナログ信号は、切れ目なく連続しているため、1個、2個といった数え方をすることはできない。一方、デジタル信号は非連続で、バラバラの数が並んだものなのだ。

こうして音のデジタル表現を理解することができる。5-1-15は、音のアナログ信号をデジタル信号に変換している様子だが、アナログ信号は常に1本の滑らかな線として表現されるのに対し、デジタル信号は多数の棒グラフとして表される。1本、1本の棒は、ある瞬間におけるアナログ信号の電圧値を示しており、その他の棒とは独立している。

■5-1-15 アナログ信号のデジタル変換

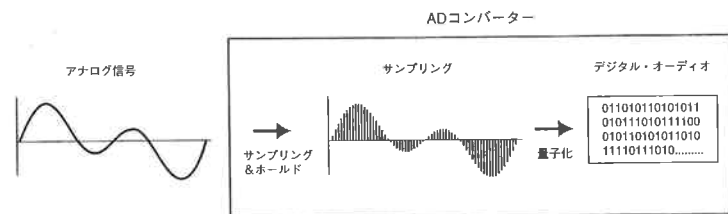


5-1-16 サンプルングとデジタル・オーディオ

コンピューターをはじめとするデジタル機器は、受け取ったデータをデジタル・データとして扱う。音も同様で、デジタル信号に変換された音のことを、デジタル・オーディオ(あるいはデジタル・サウンド)と呼ぶ。

さて、デジタル・オーディオが得られる過程を順を追って見ていこう。マイクで拾った音はまず連続的なアナログ信号に変換される。次に、このアナログ信号はADコンバーター(Analog to Digital Converter、ADC)という装置に渡される。これは、AD変換、つまり連続的なアナログ信号を離散的なデジタル信号に変換する装置である。具体的に言えば、ADコンバーターは、一定の時間ごとにアナログ信号の値を計測し、それを2進数の数値に変換する。コンピューターはこの2進数の数値として音を扱い、記録し、保存する。

■5-1-16 デジタル・オーディオが得られる過程



こうして連続する信号のある瞬間の値を計測して数値として取り出すことをサンプルングと言い、取り出された個々の数値をサンプルと言う。したがって、デジタル・オーディオとは多数のサンプルの集合だと言える。

一方、コンピューターに記録されたデジタル・オーディオは、再びアナログ信号に変換され、最終的に耳に聞こえる音波へと変換されなければならない。DA変換、つまりデジタル信号からアナログ信号を再構成する装置をDAコンバーター(Digital to Analog Converter、DAC)と言う。

5-1-17 サンプルング・レートと量子化ビット数

“よい音”という感覚は人によっても時代や文化によっても異なる。しかし、その1つの基準として原音に対する忠実度、つまり本来の音をいかに変えることなく記録したり、伝えたり、再現したりできるかが、オーディオ機器の性能の目安となる。

デジタル信号はアナログ信号を写し取ったものであるが、決して同じものではない。連続した信号を、ある時間間隔でサンプルングしている以上、あるサンプルングから次のサンプルングまでの間にアナログ信号がどう変化しているかをつかまえることはできないからだ。

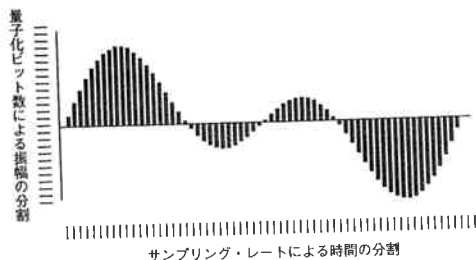
デジタル信号を元のアナログ信号にできるだけ近いものにし、デジタル・オーディオの忠実度を高めるためには、まずサンプルングの時間間隔をできるだけ短くしていく必要がある。1秒間に行うサンプルングの回数を示す値としてサンプルング・レート(サンプルング周波数)がある。サンプルング・レートが大きければ大きいほど、デジタル信号は元々のアナログ信号に対して忠実なものになる。

例えばコンパクト・ディスク(CD)の場合、そこに記録されるデジタル・オーディオのサンプルング・レートは44.1kHz(=44,100Hz)で、これは1秒間に44,100回のサンプルングを行って得られたものであることを意味する。

もう1つ、個々のサンプルを数値に置き換えることを量子化と言うが、この精度もまたデジタル・オーディオの忠実度を決める重要な要素だ。つまり、サンプルングの際に、アナログ信号の値を計測する物差しの目盛りが細かければ細かいほど忠実度は高まることになる。デジタル・オーディオを構成する各サンプルは、2進数の数値であり、010100111010……といった0か1のビット列で表されるが、1つのサンプルの値を何個のビットで表現するかを量子化ビット数と言う(単にビット数と呼ぶ場合も多い)。通常、ADコンバーターでは、アナログ信号の振幅の最大値を1、最小値を-1として変換を行うが、この-1~1の範囲をどれくらいの細かな精度で分解して数値化するかを決めるのが量子化ビット数の大きさである。

例えば、コンパクト・ディスクの場合、そこに記録されるデジタル・オーディオの量子化ビット数は16ビットになる。これは1つのサンプルが2の16乗=65,536通りの値を取り得ることを示している。つまり、-1~1の範囲を65,536分割した物差しで、個々のサンプルの計測を行うわけである。

■5-1-17 サンプルング・レートと量子化ビット数による忠実度の考え方



サンプルング・レートと量子化ビット数を大きくしていけば、理論的にはデジタル・オーディオの忠実度はどんどん向上し、さらに“よい音”を目指すことができるだろう。しかし、デジタル機器には、それだけ膨大なデータ量を極めて短い時間で処理しなければならない負担が生じる。よくCDレベルのクオリティという言葉が使われる。これは広く流通しているオーディオ記録メディアとしてのコンパクト・ディスクの量子化ビット数とサンプルング・レート(16ビット44.1kHz)を指しており、その再生音が、原音に対する忠実度という点で、取りあえず現段階で満足できるデジタル・オーディオの現実的な基準になっていることを示している。

③ サンプルング定理

あるサンプルングから次のサンプルングまでの間に、元々のアナログ信号がどう変化しているかをつかまえないことはデジタル・オーディオの宿命だ。つまり、デジタル・オーディオでは、2つのサンプルの間は滑らかに繋がっていると予測するしかない。ひょっとしたら、2つのサンプルの間にさらに細かな波形の動きが含まれているかもしれないが、それは見落とされる。サンプルング・レートを大きくして、できるだけ元のアナログ信号に忠実なデジタル変換を目指すことはできるが、表現できる周波数の大きさには限界がある。

ナイキストはこれに関して、次の定理を発見した。

デジタル・オーディオが表すことのできる最高周波数は、サンプルング・レートの1/2の周波数になる

これはサンプルング定理あるいはナイキストの定理と呼ばれ、またサンプルング・レートの1/2の周波数はナイキスト・レートとも呼ばれる。例えばサンプルング・レートが44.1kHzの場合、もし30kHzの周波数の音をサンプルングすると $44.1 - 30 = 14.1$ kHzという異なった周波数の音に化けてしまう。これをエイリアシング現象と呼ぶ。ADコンバーターはこのような現象を最小限にするために、変換を行う前にアナログ信号をフィルターに通し、ナイキスト・レート以上の成分をカットしている。

人間の耳で聞き取れる音の周波数帯域は、およそ20Hz~20kHz(20,000Hz)だとされている。このため、デジタル・オーディオの標準的なサンプルング・レートは20kHzの倍の40kHz以上になっている。コンパクト・ディスクの場合は44.1kHz、デジタル・オーディオ・テープ(DAT)の場合は48kHzという具合だ。

20kHzを超える周波数の音は超音波であり、コウモリには聞こえても人間には聞こえないとされている。しかし、聞こえないはずの超音波が、実際には音色感や音の距離・方向感といった人間の感覚に影響を与えている可能性があり、サンプルング・レートをより大きなものにする試みも行われ始めている。

④ デジタル・オーディオの可能性

アナログとデジタルではどちらが音がよいか?という議論がよく聞かれる。アナログの音、例えばビニール・レコードを再生した音は、艶やかで暖かみがあり、それに引き替え、デジタルの音、例えばCDの再生音は、ぎすぎすして冷たいといわれる場合がある。実際、プロのレコーディング・スタジオでも、デジタル・オーディオ・ワークステーションを基本に作業しながらも、独特の音質を得るため多数のアナログ機器を使って音作りをしている場合もある。

これとは別に、アナログ信号には常にノイズの問題が付きまとい、これはコピーを重ねるごとにますます顕著になる。例えばテープ・レコーダーで録音した音を何度もテープからテープへとダビングしていくと、テープを再生する際に必ず発生するヒスノイズもまた何度もコピーされ、増幅される。一方、デジタル信号は数値の集合であり、原理的

に何度コピーしても劣化しない。1という数字がコピーによって徐々に0に近づくことはないからだ。

アナログ信号は連続的で、デジタル信号は離散的で、その基本的な性質が違っている。デジタル信号にはサンプリング時に見落としが必ずつきまとう以上、アナログとデジタルの音は聴感上も微妙な点で違っていることは確かであろう。しかし、先ほども書いたように、よい音の感覚は極めて個人差があるし、時代や文化によっても異なる。どちらが音がよいか?という議論をここで行っても意味がない。

それよりも、ここで重要だと考えるのは、音がデジタル・オーディオに変換されることによってもたらされる、音へのアプローチのさまざまな可能性だ。

ここでは、音とは何か?という問いからスタートして、波形、振幅、周期、周波数、スペクトル、部分音、エンベロープといった基本的な概念で音のさまざまな側面を説明してきた。初めて触れる読者にはたくさんのなじみのない言葉の羅列に見えるかもしれない。しかし、逆に言えば、たったこれだけの基本的な概念で音は説明できるのである。説明できるだけではない、自分の求める音、今まで聞いたことがない音を生み出し、操作することが可能になるのだ。

これまで音に関する膨大なアナログ機器が開発されてきた。それらは、こうした音の基本的な概念に基づいて、音を合成し、加工し、記録し、その都度新しいサウンドの可能性、音楽の可能性を切り開いてきた。

デジタル技術もその延長線上にある。しかし、決定的に異なっているのは、デジタル・オーディオが、サンプルという数値の集合だという点だ。数値は1個ずつバラバラに操作することができる。計算したり消去したり新しく追加することもできる。数値をコンピュータで生成させて、何もないうところから新しい音を作る音響合成(サウンド・シンセシス)の可能性が開ける。また与えられたデジタル・オーディオの数値を変更することで音を自在に加工することもできる。つまり数値を生み出し、操作することで、音について想像し得るすべての可能性を試すことができるのである。

音に関するアナログ技術が切り開いて来た膨大な財産は、今1台のコンピュータの中で、デジタル・オーディオ処理という1つの技術により受け継がれ、さらに未知の領域に発展しようとしている。可能性の具体的な中身は、本章でMax/MSPによるオーディオ処理を解説していく中で理解されるだろう。

し
05/3/22

5-2 MSPの概要

MSPプログラミングは従来のMaxプログラミングと表面的には大きな違いはない。しかし、重要な点で両者の考え方には違いがあり、以前からMaxに親しんでいる人には多少とも頭の切り替えが必要だ。そこで、従来のMaxプログラミングと比較しながら、MSPとは何かについておおまかに眺めてみることにする。

🌀 MSPとは何?

MSPは独立した1つのアプリケーション・ソフトウェアではない。MSPはMaxにリアルタイム・オーディオ処理機能を追加するエクステンション(機能拡張)である。1999年にCYCLING '74より初めてリリースされたMSP1.0は、当時すでに広く使われていたOP CODEのMax3.5に、あとからオーディオ処理機能を追加する画期的なものであった。

MSP1.0は、オーディオ処理を行う多数のオブジェクトと、それらをMax3.5上で動作できるようにするためのライブラリから構成され、まさにエクステンションという性格が明瞭であった。しかし、2001年6月のCYCLING '74によるMax4/MSP2の同時リリースにより、MSPの機能はMax本体に最初から組み込まれるようになり、個別のインストール作業も不要になった。今やMSPはMaxと一体化しており、Maxのオーディオ処理機能を指す名前だと考えた方がよいだろう。ちなみにCYCLING '74では、このオーディオ処理機能を外した単体のMax4も発売している。

Maxは当初から音楽に力点を置いたプログラミング環境であったが、コンピュータの処理能力がリアルタイムにデジタル・オーディオを扱うには荷が重すぎたため、実際のサウンド生成を外部のシンセサイザーやサンプラーにゆだね、もっぱらMIDIを介したコントロールを行ってきた。しかしコンピュータの処理能力の向上とMSPの登場により、音そのものをプログラミングの対象にするという新しい段階に入ったのである。

MSPで可能になるのは、リアルタイムのデジタル・オーディオ処理である。これは一般的なコンピューター用語では、音に関するリアルタイムなデジタル信号処理(DSP)と呼ぶことができる。“音に関する”と付け加えるのは、音以外に映像などでもDSPが行われるためである(『Chapter 6 画像処理』参照)。また“リアルタイム”と付け加えるの