

# サウンドエフェクトの数値実験

2009SE283 寺川智博

指導教員：小藤俊幸

## 1 はじめに

サウンドエフェクトとは、アナログの音をプログラムを用いてデジタル信号に変換して加工した音に変える技術のことであり、今回はその中でもディストーションという種類のものを取り上げ、加工の原理を解説しながら、実際にプログラムを用いてアナログ信号からデジタル信号に変換し、Windows Media Player を使って出力する。

## 2 LPF

LPF とは Low-Pass Filter の略で、低域通過フィルタと訳すことができる。[1]

これは、標本化する際に生ずるエイリアス歪みという現象を抑えるためのものであり、エイリアス歪みとは、アナログ信号に標本化周波数の  $1/2$  を超える周波数成分があった時に、その成分が標本化周波数の  $1/2$  以下の周波数成分としてデジタル化してしまう現象である。[1]

これは後述する A-D 変換よりも前に通す必要がある。

## 3 A-D 変換

A-D 変換とは、Analog-to-Digital 変換の略で、実際に音として聴こえるアナログ信号をコンピュータなどで扱えるようにデジタル信号に変える処理のことであり、「標本化」と「量子化」に分けられる。

A-D 変換前のアナログ信号の波形はすべての時刻で値を取るのに対し、A-D 変換後のデジタル信号は標本化・量子化を行うため、離散的な時刻でしか値を取らない。そのため A-D 変換後の波形が本来のアナログ信号により近付くためにはこの標本化・量子化をより細かい時間間隔で行う必要がある。

### 3.1 標本化

標本化は、1 秒あたりの標本化の回数として定義される標本化周波数という時間間隔でアナログ信号を計測するものであり、一回の標本化の周期は標本化周期と呼ばれる。標本化周波数は単位  $Hz$  を用いる。標本化周期は標本化する時刻であるため時間  $s$  となり、標本化周波数を  $f_s$ 、標本化周期を  $t_s$  とすると、両者の関係は以下の式を用いて表すことができる。[1]

$$t_s = 1/f_s$$

このことから周波数が高い場合は低い場合に比べてアナログの信号を観測する間隔が短くなるのがわかる。実際に市場で売られている音楽 CD は  $44.1kHz$  で標本化されているが、これはシャノンの標本化定理により標本化周波数の  $1/2$  以下の周波数成分については、アナログ信号とデジタル信号が数学的に等しくなると証明されていることと、人間の可聴領域が  $20Hz$  から  $20kHz$  までであると言われていることに基づいている。これは音楽な

どにおけるアナログ信号をデジタル化する際に、元の音データの品質を損なうことなく、データ量が膨大になるのを防ぐために考えられたものである。そのため、電話などの音データの量子化は  $8kHz$  で行われるのが一般的である。[1]

### 3.2 量子化

量子化とは、標本化したアナログ信号をある基準のもとで数値化するものであり、量子化をする際に必要となるステップ数を二進数で表した時に必要となる桁数によって定義される「量子化精度」によって表され、単位は  $bit$  を用いる。[1]

例えばステップ数が 4 の場合は  $4 = 2^2$  より、 $2bit$  で量子化されたことになり、ステップ数が 8 の場合は  $8 = 2^3$  より、 $3bit$  で量子化されたことになる。

#### 3.2.1 ダイナミックレンジ

量子化と関連するものとしてダイナミックレンジというものがある。これは量子化精度によって表される尺度であり、CD などの媒体に記録することができる最も大きい音と最も小さい音の比で定義される。[1] つまり、ダイナミックレンジが大きい媒体は大きい音に埋もれてしまうような小さい音も精度よく記録できるということになる。ダイナミックレンジの大きさは量子化精度によって以下のように求めることができ、単位は  $dB$  を用いる。[1]

$$dB = 20 \log_{10} \left( \frac{2^{\text{ビット数}}}{1} \right)$$

この式によると一般的に  $16bit$  で量子化される音楽 CD のダイナミックレンジは以下のようになる。

$$20 \log_{10} \left( \frac{2^{16}}{1} \right) \cong 96dB$$

このように音楽 CD のダイナミックレンジはおよそ  $96dB$  であることがわかるが、CD が発売される前に普及していたレコードやテープなどの媒体はおおよそ  $60dB$  程度のダイナミックレンジで記録されていたと言われている。このことから現在普及している CD は昔に比べてはるかに高い精度で音を記録できることがわかるわけであるが、例えばクラシックなどの音源を完璧に記録しようとする  $100dB$  以上のダイナミックレンジが必要とされるため、CD に記録されているクラシックの音は、実際に演奏されるクラシックを聴いた時の音よりもかなり劣化した音になっているわけである。しかし、DVD を用いて音を記録した場合、DVD は  $24bit$  で量子化されるため、ダイナミックレンジの大きさは  $144dB$  となり実際に聴くクラシックの演奏と音質的に大差のない精度で音を記録することができる。[1]

## 4 クリッピング

クリッピングとは、振幅が増幅しすぎて予め設定しておいた上限、下限(本来はアンプの出力制限)を超えた場合に超えた部分を打ち切る処理のことである。このクリッピングを行うと、音が割れて歪がかったものになるわけであるが、この歪がかった音というのがディストーションというサウンドエフェクトの音になるのである。

クリッピングには「ハードクリッピング」と「ソフトクリッピング」の2種類がある。どちらのクリッピングの方法も制限を超えた部分をカットするという点では同じであるが、ハードクリッピングは  $\sin$  波と制限の境目が角張っているのに対し、ソフトクリッピングは  $\sin$  波と制限の境目が丸みをおびているのがわかる。これはハードクリッピングが制限に達すると同時に急激にクリッピングを行うというものであるのに対して、ソフトクリッピングは制限に近付くにつれて緩やかにクリッピングを行うという処理だからである。

この二つのクリッピングの方法は見た目では明らかに違いがあるが、どちらも矩形波に似た形となり、本来の音以外の周波数成分を発生させるという点では同じである。そしてクリッピングを行った際に生ずる基本音とは異なった周波数成分というのが倍音である。倍音とは、音を出した時に元となる音と同時に発生するその音の整数倍の周波数を持つ音のことであり、例えば元の音の周波数が  $220\text{Hz}$  だった場合、2倍音は  $440\text{Hz}$ 、3倍音は  $660\text{Hz}$ ... といった具合になる。

すなわち、 $i$  倍音の周波数  $h_i$  は、基本音の周波数  $f_0$  に対して

$$h_i = if_0 \quad (i \geq 2)$$

と表すことができる。[1] 先述したように、 $\sin$  波はクリッピングを行うと矩形波に似た形をとる。矩形波は、基本音に対してその奇数倍の音を足し合わせることでできる波形であるため、 $\sin$  波を

$$s(n) = A \sin\left(\frac{2\pi f_0 n}{f_s}\right)$$

のように定義すると

$$s(n) = A \sin\left(\frac{2\pi f_0 n}{f_s}\right) + \frac{A}{3} \sin\left(\frac{2\pi h_3 n}{f_s}\right) + \frac{A}{5} \sin\left(\frac{2\pi h_5 n}{f_s}\right) + \dots + \frac{A}{i} \sin\left(\frac{2\pi h_i n}{f_s}\right)$$

のように表すことができる。[1] このように矩形波は基本音の奇数倍となる音が含まれていることがわかるわけであるが、上述したようにクリッピングを行うとその波形は矩形波と似たものになる。矩形波は、基本音に対してその奇数倍の音を組み合わせたものであるため、倍音を含むことは明白である。つまり、クリッピングを行うと本来の音にはない周波数成分を持つ音である倍音が発生し、この倍音はディストーションというサウンドエフェクトの最も重要な音の特徴である。また、上記の二つのクリッピングの方法では、ハードクリッピングのほうが倍音を多く含んでおり、硬い音になるため、一般的にディ

ストーションの加工をする際は、ハードクリッピングが用いられることが多い。

### 4.1 ハードクリッピングによるディストーション

ハードクリッピングを用いたディストーションは出力信号を  $s_1(n)$ 、入力信号を  $s_0(n)$ 、入力信号の増幅率を  $gain$  として以下の式で定義される。[1]

$$s_1(n) = \begin{cases} 1 & (gain * s_0(n) \geq 1) \\ gain * s_0(n) & (-1 < gain * s_0(n) < 1) \\ -1 & (gain * s_0(n) \leq -1) \end{cases}$$

また、この式は入力信号の増幅の限界値を 1 で正規化し、 $|s_1| \leq 1$  としている。この式を用いると入力信号のクリッピングが可能であるが、このままだと出力した時の音が大きくなりすぎてしまうため、増幅させた入力信号に  $level$  という係数を掛けて、出力の大きさを調節をする。すると、最終的な式は以下ようになる。

$$s_1(n) = \begin{cases} level & (gain * s_0(n) \geq 1) \\ level * gain * s_0(n) & (-1 < gain * s_0(n) < 1) \\ -level & (gain * s_0(n) \leq -1) \end{cases}$$

つまり、この式が入力信号をクリッピングし、実際にアンプから音をだせる信号にする式である。

## 5 D-A 変換

上記の方法でディストーションへ変換する処理が終わったら最後に実際にアンプから出す信号にしなければならない。そこで初めに行った A-D 変換の逆の作業を行うのだが、その際に標準化周波数の  $1/2$  よりも大きい周波数成分が発生するのを防ぐために初めと同じように LPF に通さなければならない。[1]

## 6 実験結果

以上の方法を実際のプログラムを用いて増幅率 ( $gain$ ) を 100、 $level$  を 0.5 に設定して実験したところ、加工する後は倍音が発生しており、音が歪んでいた。また歪みの発生と同時に音の伸びが長くなっているのがわかった。

## 7 おわりに

ディストーションのような歪んだ音を作るためには倍音を含ませる必要があり、そのためには適当な波形を矩形波にする必要があることがわかった。

アナログ信号を加工するためには標準化、量子化をし、連続的な波形から離散的な波形にする必要があることがわかった。

## 参考文献

- [1] 青木直史 編：『C 言語ではじめる音のプログラミング—サウンドエフェクトの信号処理』、オーム社、東京、2008。
- [2] FFT アナライザ入門  
[http://www.aandd.co.jp/adhome/products/nvh\\_analysis/fft/fft01.html](http://www.aandd.co.jp/adhome/products/nvh_analysis/fft/fft01.html)