

画像・音声の表現と処理

増原 英彦

視覚や聴覚を通して得られる情報の表現には、大きく分けて2通りの方法がある。1つは、これらの情報を音声や画像のようなアナログ情報として表現するもの、もう1つは、楽譜などのように「音の高さや長さ」を記号化したり「図形の形・位置・大きさ」を記号化して表現するものである。

1 アナログ情報のデジタル化

音声や画像などのアナログ情報を計算機で処理するためには、情報をデジタル化する。音声や画像をデジタル化する方法や、その際の性質は共通部分が多いので、ここでは簡単のために音声情報について説明する。

音声情報は、時間とともに連続的に変化する音圧だと考えることができる。このような情報を計算機で扱うためには、標本化(sampling)と量子化(quantization)という2つの変換が必要である。

(2つの耳から入る音声情報は、そのような情報が2つあると考えればよい。また、モノクロの平面画像は位置とともに連続的に変化する明るさだと考えればよい。カラー画像は、モノクロ画像に相当する情報が3原色ごとにあると考えればよい。)

標本化: 音圧は連続的な時間とともに変化しているので、それを適当な間隔で測る。このときの間隔を標本化周波数という。例えば音楽 CD を作るには音圧を $1/44,100$ 秒間隔で測っているため、標本化周波数は $44,100\text{Hz}$ ($= 44.1\text{kHz}$) となる。

標本化周波数は、元の情報に含まれている情報をどの周波数成分まで記録できるかを決める重要な要素である。音声情報における時間周波数は「音の高さ」に相当するので、標本化周波数によって「どれだけ高い音まで記録できるか」が決まることになる。

標本化周波数と記録できる周波数(音の高さ)の間には、標本化定理という関係がある。この定理は

$$(\text{記録できる情報の周波数}) \times 2 \leq (\text{標本化周波数})$$

という関係を示している。音楽 CD の例に戻れば、 44.1kHz 標本化周波数で記録されているということは、 22.05kHz より高い音は記録できないということである。人間の聴覚は約 20kHz より高い音は良く聞きとれないと言われているので、CD が記録する音は「人間が良く聞こえる音をだいたいカバーしている」と言える。(音叉などが出す高さの「ラ」の音は 440Hz であるが、 22.0kHz はそれよりも5オクターブ以上音が高である。)

量子化: 標本化された情報はまた、連続的な量である。これを離散的な値で近似することを量子化と言う。例えばある瞬間の気温を計算機で扱うためには、「気温を -50 から $+50$ まで、 0.1 単位で区別する」ような量子化を行うのであれば、気温を $-50.0, -49.9, -49.8, \dots, 49.8, 49.9, 50.0$ の 1001 段階だけで代表させることになる。実際の気温が $29.87654\dots$ であったときは、 29.9 だとして扱うことになる。

音声情報であれば、標本化された各時点での音圧は連続的な量であるので、音圧を離散化して扱うことになる。この際に、量子化された値の情報量は量子化ビット数と言われる。例えば 256 段階に区別するのであれば、 $256 = 2^8$ より 8 ビットに量子化されたとなる。量子化されたビット数が多いということは、音圧の強弱に関してより詳細に表現できることになる。例えば音楽 CD では 16 ビット量子化を行うので、65,536 段階の音圧を表現できる。

問題 8-1: (音声情報のデジタル化) デジタル化して音声情報を記録する機器には、CD の方に DVD や MD などがある。これらの機器でのデジタル化で使われている標本化周波数と量子化ビット数を調べ、比較せよ。

問題 8-2: (画像情報のデジタル化) 写真や絵のような平面の画像をデジタル化して表現する場合には、平面を細かなマス目に分け、各マスの明るさを記録する。このような方法における標本化周波数、量子化ビット数は何を決定しているかを考えよ。

問題 8-3: (画像情報のデジタル化の比較) 計算機のディスプレイ装置、デジタルカメラ、ファクシミリなどはいずれも画像情報をデジタル化して扱っている。これらの装置における標本化周波数、量子化ビット数を調べ、比較せよ。

2 音声や画像の記号的表現

前節ではアナログ情報として与えられた音声や画像をデジタル化して表現する方法を説明した。一方、音楽や図は楽譜や言葉で表わすこともできる。例えば図であれば、

- 座標 (0,0) を中心に半径 10 の円を描き
- 座標 (10,0) と座標 (20,8) を対角線とする長方形を描き
- 座標 (0,0) から座標 (15,4) へ直線を引く

といった定規やコンパスを使って図を描く手順によって表わすこともできる。音楽であれば、出すべき音の音色、音階、音長を順に指定してゆくことで表わすことができる。いずれの表現方法も、座標、音色、音階、音長などを整数によって表わすことができるので、計算機によって扱うことができる。

実際に、計算機でこのような表現を使うことは珍しくはない。

製計図を作成する場合のように、人間が計算機を使って図を描く際には、図形の形状を記号化して表わすことが行われている。このような表現方法は図形のベクター表現とも言われる。(一方、アナログ情報としての図形を標本化・量子化して表わすものはラスタ表示と言う。)このような表現方法では、一度描いた図形を拡大・縮小・回転などの変形をしたり、消去・複製をすることが容易である。

音楽を記号によって表わす方法の 1 つには、MIDI (Musical Instruments Digital Interface) と呼ばれる形式がある。これは音階、音色、強さ、長さをそれぞれ整数で表わしている。

3 デジタル化された情報の圧縮

アナログ情報をデジタル化したデータは時として非常に大きなものになる。前回の資料で見たように、符号化の方法を工夫したり、繰り返し表われるデータに短い表現を与えることによってデータを

圧縮することは可能である。

しかし、文章を圧縮するための方法が音声や画像の圧縮にも有効とは限らない。例えば音楽の演奏をデジタル化したデータを考えてみよう。数分の1秒から数秒間という時間単位では「同じ音」が続いている。しかし、音の標本化を行う時間単位(数千分の1秒)では音圧もっと非常に複雑な変化をしている。(この音圧が変化する形を「音の波形」という。この波形が「音色」や「音階」を決めている。)さらに「同じ音」を2度演奏した場合でも、正確に同じ波形になることはまず無いので、「同じ値」が繰り返させることはほとんどない。

そこで、アナログ情報を圧縮する場合には、「人間が気付かない情報」を減らすことによって圧縮をすることが行われている。これは、これらの情報が最終的に人間が見たり聴いたりする目的で使われることを利用している。例えば、人間の聴覚は非常に低い音や非常に高い音は、中位の高さの音に比べると鈍感であるので、低音や高音に関する情報は精度の低い表現を使う。また、ある高さの強い音があった場合、その近くの高さの弱い音は聞き分けられなくなるので、その音を消してしまう、といったことを行う。これらの操作は「音圧の時間変化」である波を、フーリエ変換(の一種である変換)によって「様々な音階(周波数)の波の強さ」に置き換え、それらの「波の強さ」に対して精度を下げた表現を用いるなどの方法で圧縮する。

実際 MP3 と呼ばれる音声の圧縮方式や、JPEG と呼ばれる画像の圧縮方式はこのような方法の代表例である。

問題 8-4: (圧縮率) 音声や画像の表現方式の圧縮率を調べよ。例えば iTunes というソフトウェアを使えば、CD などに記録されている音声を MP3 形式に変換することができる。CD に記録されているデータの量と、MP3 形式に変換されたデータの量の比から圧縮率を求めることができる。

3.1 可逆圧縮と不可逆圧縮

データを圧縮する方法には、元のデータを完全に復元できるかどうかで2つの種類に分けることができる:

- 完全に復元できない(不可逆圧縮といわれる)方法では、展開後のデータは圧縮前とほぼ同じデータにしかならない。画像や音のように、完全に同じでなくても問題がない情報に用いられる。上述の MP3 や JPEG などの圧縮はこれに相当する。
- 完全に復元できる(lossless 圧縮といわれる)方法では、展開後のデータは圧縮前と全く同じデータになる。あらゆる種類の情報に用いることが可能だが、前者と比べて圧縮率(小さくなる度合い)も悪い。前回説明したランレングス圧縮や Lempel-Ziv 圧縮は完全に復元できる圧縮である。

4 実習: 画像の処理

練習 8-5: (ベクター表現による図の作成) HWB 「23.1. Tgif 入門」に従って、簡単な説明図を描け。描いた図はまず普通に保存した上で、PNG あるいは JPEG 形式で出力せよ。

練習 8-6: (ラスター表現による図の編集) HWB 「23.2. gimp 入門」に従って、練習 8-5 で出力した PNG あるいは JPEG 形式の図に色をつけてみよう。

5 課題

課題 8-1 および課題 8-2 のレポートを1つの HTML ファイルにまとめ、以下のような URL によって公開せよ:

`http://user.ecc.u-tokyo.ac.jp/~gXXXXXX/jreport8.html`

ただし、

- 課題 8-1 は必須とする。課題 8-2 もできるだけ提出することが望ましい。
- 課題を完了するのに要した時間と授業に対する感想も書くこと。
- 引用する場合はレポート中のどこが引用範囲であるか、また、出典がどこであるかが明確になるようにせよ。断りなく他人の言説を書き写す行為は盗作になるので注意せよ。

提出期限 2004年6月21日(月) 23時59分

課題 8-1: (情報交換システムの調査) 計算機を使った情報交換のためのシステムには、授業では紹介していないものも沢山ある。そのようなシステム1つをとりあげ、どのような仕組みで機能しているかを調べ、説明せよ。

情報交換システムとしては、ネットニュース、チャット(インスタントメッセージ)、メリーリングリスト(メールマガジン)、掲示板(WWWを使うもの・使わないもの)、オークション、電子商取引、ファイル交換(Napster, WinMX, Winny, etc.)、weblog、IP電話などが考えられる。

仕組みの説明には、できるだけ図などを使い、簡潔にわかりやすくせよ。

課題 8-2: (情報交換システムの比較) 課題 8-1 で調査した情報交換システムを、電子メールシステムあるいは WWW システムと比較せよ。

比較にあたっては「サーバどうしが通信をするか」「利用者どうしは即座に情報を受け取るか」「情報発信者の特定は可能か」などの比較項目を考え、共通点や違いがわかるように整理せよ。