情報処理 II 第 8 回

情報の電子化 (3) 画像、音声

桂田 祐史 (かつらだ まさし)

2005 年 6 月 16 日

ホームページは http://www.math.meiji.ac.jp/~mk/syori2-2005/

1 連絡事項

- 課題 3 のレポートは昨日 (6/15) までが一応の締め切りでした。このレポートについては「正解」を次回発表するつもりです。まだ出していない人はお早めに。
- 前回準備が整っていなかったのでパスした、換字暗号を解いてみよという研究課題 http://www.math.meiji.ac.jp/~mk/syori2-2005/jouhousyori2-2005-07/node7.html の説明から始めます。

2 マルチメディア情報の電子化

文書はテキスト・ファイルとして電子化できることを見てきたが (なお、付録 A にも目を通して欲しい)、画像や音声についてはどうだろうか？

2.1 画像の電子化

コンピューターのディスプレイは、色 (明暗込み) をつけた点 (ピクセルあるいはドット) の集まりとして構成されている (カラー・ディスプレイ上の 1 ピクセルは、R, G, B (red, green, blue) に光る 3 つの小さな点で構成されることが多い)。各ピクセルの色情報を数値化して、それを集めることによって画像情報をデジタル・データとして表現できることになる。

例えば、実習に用いているパソコンのディスプレイ画面は、

- 縦横約 1000 ピクセルずつ[1]、つまりおよそ $1000^2 = 100$ 万ピクセルからなっていて、
- 各ピクセルは約 1600 万色の色がつけられる。

[1] UNIX 環境では、xdpyinfo コマンドを実行すると調べられる。例えば、この文章を書いているパソコンの画面は、横 1024, 縦 768 ピクセル。

ピクセル数は最近のパソコン、ワークステーションでは大体この程度である (やや高級なもので 1600 × 1200 程度)。色数は、1 ピクセルの色を 8 ビットで表わす場合 256 色で、16 ビットの場合は 65536 色、24 ビットの場合は約 1600 万色となる[2]。

この素朴な方法で 1 画面分の画像を記録したファイルのサイズはいくらになるか計算してみよう (すぐ後で実際にそういうファイルを作ってみて確認する)。すぐに分かることは

画像を記録したファイルはサイズが大きい。

これは動画では特に顕著である。テレビなどでは、1 秒間に約 30 の画像を次々に映すことで動画を表現している。パソコンの画像をテレビのように 1 秒間に約 30 回描き換えるとして、1 時間にどれだけの量のデータが必要となるか？ちょっと考えてみよう[3]。

そこで、

画像の記録にはデータの圧縮が不可欠

になる。デジタル・データの圧縮法は次のように大きく二つに分類できる。

デジタル・データの圧縮法

(1) 可逆な圧縮 オリジナルのデータが完全に復元できる圧縮法

(2) 不可逆な圧縮 オリジナルのデータが完全には復元できない圧縮法— うまくやると圧縮後のデータのサイズを非常に小さくすることができる

文書データやプログラム・データなどは可逆な圧縮をすることになる。UNIX や Windows の世界で普及しているプログラム gzip, compress, zip, LHa などは可逆な圧縮をする。

画像の場合は、オリジナルと多少違っていても、十分役立つことが多いため、

画像データでは不可逆な圧縮が採用されることが多い。

色々な圧縮法があり、画像データのフォーマットも多くの種類がある。静止画像データ・フォーマットの代表的なものの名前をあげると、**JPEG**, **PNG**, **TIFF**, **GIF**, ... (デジタル・カメラでデファクト・スタンダードになりつつあるエグジフ (Exif) は JPEG をベースにデジタル・カメラ向けの固有情報などを記録できるようにしたものである。)

画像データは不可逆な圧縮をすることでサイズをかなり小さくできる。

---

[2] $2^8 = 256$, $2^{16} = 65536 \fallingdotseq 6$ 万 5 千, $2^{24} = 16777216 \fallingdotseq 1600$ 万.

[3] 3MB × 30 × 3600=300GB 以上 !!! ビデオのデジタル技術がいかに圧縮技術のお世話になっているか良く分かる。

X が動いている場合に試せる実験

```
a308-06% xdpyinfo                                                    (結果は省略)
a308-06% xwd > image.xwd                                             →この後マウスの入力待ちになる。
                                                                       ウィンドウを選択して左ボタンを
                                                                       クリックする。
a308-06% ls -l image.xwd                                             ← サイズを見てみる。
(省略)                                                               → ルートウィンドウを選択した場合
                                                                       全画面が記録されて約 5 MB
a308-06% xwud -in image.xwd                                          ← 試しに画面に表示してみよう。
a308-06% cat image.xwd | xwdtopnm | pnmtopng > image.png             ← PNG というフォーマットに変換
a308-06% ls -l image.png                                             ← サイズを見てみる。
-rw-r--r-- 1 re00018 re00018 21723 Jun 15 09:27 image.png            → かなり小さくなった！
a308-06% display image.xwd                                           → 表示してみる。
                                                                       終了は、マウス右ボタン, Quit.
a308-06% rm image.xwd                                                ← ディスクの無駄使いなので削除
```

Windows でウィンドウを画像として保存する方法

Windows で、PrintScreen (PrtScr) キーをタイプすると画面全体をクリップボードにコピーできる。また Alt キーを押しながら PrintScreen キーをタイプするとその時にアクティブなウィンドウをクリップボードにコピーできる。それを画像を扱えるソフト、例えば「ペイント」(スタート→すべてのプログラム→アクセサリ→ペイント) に編集→貼り付けすれば OK。ペイントでは、ファイルをセーブするときにファイルの種類が選択できる。無圧縮であるビットマップ形式 (ビット数が選択できる) 以外に JPEG, GIF, TIFF, PNG が用意されている。

Windows で画面関係の情報を調べる方法

マウスカーソルをデスクトップにおいて、マウスの右ボタンを押し、「プロパティ」を選択し、「設定」をクリックして選択する。「画面の解像度」と「画面の色」を見るとよい。

かくして、画像圧縮は必要不可欠なもので、盛んに利用されているものだが、

画像圧縮技術は特許などがからんで「取り扱い注意」の部分がある。

例 (GIF の場合)

WWW ページでも良く使われている GIF 画像は、UNISYS という企業が特許を持っている[4]圧縮法 (LZW) を利用することが多い。UNISYS と契約した会社の作成したソフトで GIF 画像を作成する場合は問題ないが、そうでない場合は特許料を支払う必要がある。大した問題ではないと思われるかもしれないが、例えばフリーソフトは閉め出されることになる。今のところ、データを再生するプレーヤー (画像の場合はビューアー？) の開発を制限しているケースはないようだが、万一プレーヤーも制限されるようになると、「作成したデータは本当に作成者のものなのか？」心配になってくる (筆者の杞憂であれば良いのだが)。

[4]アメリカでは、2003 年 6 月 20 日に失効して、日本でも無事 2004 年 6 月 20 日に失効した。

2.2　音声の電子化

音声の電子化は、適当な時間間隔で空気の圧力変化を (離散的な) 数値で記録するのが基本である[5]。

音楽用 CD (CD-DA) では、44.1kHz のサンプリング・レートで[6]、圧力を 16 ビット (2 バイト) の数値で (普通は 2 チャンネルのステレオで) 記録する。特にデータの圧縮はされないため[7]、1 分間、ステレオで録音されたデータの容量は、

$$44.1 \times 10^3 \times 2 \times 2 \times 60\mathrm{B} = 10584000\mathrm{B} \fallingdotseq 10.1\mathrm{MB}$$

ものサイズになる (CD が登場した当時は、これは巨大なサイズであったが、今では大したことがないように感じられるようになった...)。

Windows で録音してみる

マイクロフォンを Windows パソコンに接続すれば、Windows に付属しているサウンド レコーダー(「スタート」→「すべてのプログラム」→「アクセサリ」→「エンターテイメント」→「サウンド レコーダー」) を用いて録音できる (ただしわずか 60 秒 — ささやかな仕様ですね)。データの形式は **G.723.1**, **PCM**, **MP3**, **Windows Media Audio** (V1, V2) など。

(音声に関する実験は、後日 Mathematica を体験するときに行なう予定。もしマイクを持っていたら、その時に持って来ると楽しい？)

音声データに関しても、画像データと同様のことが言える。

音声データも不可逆な圧縮をすることでサイズをかなり小さくできる。
(ただし音楽用 CD のデータは圧縮されていない!!)

音声データを圧縮する方法も特許がからむことが多い。

最近では著作権保護機能なども組み込まれるようになっている。

MD の場合

MD (MiniDisc) では、**ATRAC** (アトラック) (Adaptive TRansform Acoustic Cording) という圧縮技術を採用している。CD-DA の場合の約 1/5 程度のサイズになるということで、容量約 140MB の MD メディアに CD-DA とほぼ同じ長さの音声を記録できる。なお ATRAC3 (MD の LP モード (MDLP)) では、CD-DA の約 1/10 程度のサイズになる。

---

[5]手前味噌になるが、桂田研の卒研レポート松山周五郎『音の Fourier 解析』http://www.math.meiji.ac.jp/~mk/labo/report/open/2003-matsuyama.pdf

[6]要するに 1 秒間に 44.1 × 1000 回測定する。人間は通常 20kHz までの音を聴くことができると言われていて、20kHz までの音を記録するには、倍の 40kHz のサンプリング・レートがあれば十分 (Shannon の標本化定理による) ということから、この程度の値を採用した (らしい)。

[7]何と言っても、CD (compact disc) の誕生は NEC PC-9801 (主記憶 128KB, CPU clock 5MHz) が出たか出ないかの、1980 年代初頭だから、圧縮技術が気軽に使えたはずはない。CD については、http://www.sony.co.jp/Fun/SH/2-9/h1.html, http://www.oneoffcd.com/info/historycd.cfm など。

MP3 の場合

mp3 (MPEG 1 Audio Layer-3) という圧縮技術では、44.1kHz, 16bit ステレオという品質を採用した場合で、オリジナルのデータの約 1/10 程度のサイズ (つまり 1 分 1 MB 程度) に圧縮できる (例えば 3 分 44.39 秒の曲が 3,590,559 バイト ≒ 3.4MB)。圧縮・復元の方法は公開されているが、実際に圧縮するプログラム (エンコーダー) は特許料を払わずに作成することはできない (？)。

2.2.1 研究課題 2

(マイクを持っている人向き) パソコンにマイクを接続して、録音してみよ。無圧縮 PCM, 44.100kHz, 16 ビット, ステレオで 10 秒程度録音してみて、ファイルのサイズがどうなるか調べ、理論値 (自分で計算せよ) と比較せよ。ファイルの形式を MP3 (MPEG Layer-3) に変換するとファイルのサイズはどう変るか。

2.2.2 生データ募集

身近に楽器のある人にお願い。実験用 (もしかすると情報処理 II で、あるいは桂田研の卒業研究で利用) に色々な楽器の音を録音してもらえないでしょうか。なるべく高品質 (サンプリング・レートは 44.1kHz 以上, 量子化ビット数 16 ビット以上, ただしステレオである必要はない) でお願いします。使いやすいように一ファイルに一つの音が入っているようにすると嬉しい。弦楽器や太鼓などの音が欲しい...

3 レポート課題 5 のための説明の補足

3.1 最近の外部記憶媒体の容量

- (最近は使われなくなってきたが...) フロッピーディスクは、2DD が 720KB, 2HD が 1.44MB
- CD-ROM (CD-R, CD-RW なども同様) は 650 ~ 700MB[8]
- DVD-ROM/R/RW は CD-ROM の約 7 倍 (片面シングル・レイヤーの場合) 4.7GB
- パソコン用ハードディスクは数十 GB 程度で、200GB を超えるものも珍しくない。
- デジタル・カメラや携帯電話で使われている記憶媒体には色々あるが[9]、その一つである SD メモリーカードでは、32 ~ 512MB の容量のものが市販されている。

[8] 音楽用 CD では、当初記録時間が 74 分までだったが (伝説によるとカラヤン指揮のベートーベン第 9 交響曲が一枚に収まるように決められたとか)、その後 80 分のものも出て来たことに対応している。

[9] コンパクトフラッシュ (CF), メモリースティック, SD メモリーカード, スマートメディア, xD ピクチャーカード, マルチメディアカード, マイクロ・ドライブ, mini SD カード、... 覚えきれない。

3.2 テキスト・ファイルのサイズ

通常、日本語は 1 文字 2 バイトで表現されている。だから日本語のテキスト・ファイルのサイズは大ざっぱに言って、文字数× 2 バイトと考えれば良い (実際には空白や改行、また英数字など 1 バイトで表現される文字などもあるので、あくまでも概算である)。

文書の文字数を概算するのは簡単だから (一行○ 文字、一ページ○ 行、○ ページ)、文書の文字情報をテキスト・ファイルにした際のサイズの概算ができることになる。

4 レポート課題 5

以下のことを調べよ。〆切は 7 月 6 日とする。(ファイルのサイズについての感覚を身につけてもらうのが主旨であって、自分で計算すること。)

(1) フロッピー・ディスクを使ったことがあるか?(教えて下さい。アンケートのつもりです。)

(2) 自分が触れるコンピューター (情報科学センターの Windows 環境, UNIX (Linux) 環境, 自宅のパソコン) にあるファイル[10]のサイズについて調べよ。バイト数以外に、CD-R にどれくらい入るかを記せ。

(a) 文書ファイル
レポート、メイル、C プログラム、TEX のソース (.tex) など。
(単にサイズだけ書いてもあまり意味がない。長いものもあれば短いのもあるのだから。例えば「印刷して何ページくらいの文書が何バイトになる」等の情報を添えること。)
ワープロソフトの文書ファイルなどを調べてみるのも良い。

(b) 実行可能プログラム

- 自分が普段使っているプログラムをいくつか選び、そのプログラム・ファイル[11]のサイズを調べよ。(大規模なソフトウェアの場合、一つのプログラムから別のプログラムを呼び出し、全体として複数のプログラムが協調して働くこともあるので、結構難しい。それゆえ必修とはしないが、トライしてみること。)
- C で書いたプログラムを持っている場合、コンパイル前 (ソースプログラム) と後 (実行可能プログラム、あるいは機械語プログラム) でどうサイズが変わるか。

(c) 現在、自分が持っているファイルの総量。それは自分のホームディレクトリィのあるディスクの全容量の何 % に相当するか。

UNIX 環境で調べる場合

```
a308-06% du -ks ~
a308-06% du -ks ~/.snapshot
a308-06% df -k
```

[10]自分以外の持ち物でも構わない。例えば、私はホームディレクトリィ (~re00018) を開放している (読み出しを許可してある) ので、そこにある Gutenberg テキストや、解析概論 I の講義ノートの LATEX ファイル (~re00018/tex-sample/textbook/ にある) などを調べることも出来る。

[11]プログラムの実体は、UNIX ならば which プログラム名 として追跡する。Windows ならばアイコンを右クリックしてプロパティから追跡する。

Windows 環境から調べても良い (どうやれば良いかは自分で見つける)。

(3) 自分が持っている本を一冊選び、その文字情報を記憶するファイルを作った場合、サイズはどれくらいになるか計算せよ。CD-R には何冊分記憶できるか。(古い小説などの場合、実際に青空文庫で電子化されたファイルが探し出せるかもしれない。自分の計算と照らし合わせると面白い。)

(4) (もし出来れば) 画像ファイル、音声ファイルなど。これはパソコンに限らない。
(記録の形式、画像の場合は図の大きさ (ピクセル数) & 色数、音声の場合は時間等も分かる範囲で調べる。本来のサイズの何分の一に圧縮されているか概算せよ。)
カメラつき携帯を持っている人からのレポート求む (私に色々教えて下さい)。

注: 音声ファイルについては、時間を取ってもう少し詳しく説明することを考えている。

A　データの圧縮用コマンド

A.1　UNIX の場合

UNIX では、可逆な圧縮をするための一般的なコマンドとして、compress, gzip, bzip2 などがある。compress は以前はよく使われたが、やはり特許の問題であまり使われなくなってきた。最近は gzip が良く使われている。

```
a308-06% cd filter
a308-06% cp alice29.txt alice29.doc
a308-06% gzip alice29.doc                ← 圧縮する
a308-06% ls -l alice*                    → alice29.doc の代りに alice29.doc.gz という
                                           小さなファイルが出来る
a308-06% wc alice*                       ← ファイルのサイズを調べる。
a308-06% gzip -d alice29.doc.gz          → 復元する。
a308-06% ls -l alice*                    → alice29.doc.gz が無くなり、alice29.doc が現れる。
a308-06% diff alice29.doc alice29.txt    ← 変化がないかチェックする (変化はないはず)。
```

英文テキスト・ファイルなどでは 1/3 程度までにサイズが小さくなる。また `gzip -d` で decompress (圧縮の解除) すると、完全に元に戻る。

A.2　Windows の場合

Windows では、複数のファイルを一つにまとめてから圧縮をするコマンド (アーカイバー (archiver) と呼ばれる) が普及している (Lha, WinZip, Cab, ...) [12]。私の最近のお気に入りは、**Lhaca** デラックス版 (`http://www.au.wakwak.com/~app/Lhaca/` から入手可能) である。情報科学センターの Windows 2000 環境には、圧縮機能を除いた (解凍機能のみを備えた) **Lhasa** がインストールされているが、この選択はちょっと残念[13]。

[12] UNIX では、複数のファイルを一つにまとめるコマンドとして、**tar** が良く使われている。tar と (例えば) gzip を組み合わせることで Windows のアーカイバーと同様のことができる。このあたりはプログラムの設計思想の違いであろう (単一の機能を持ったプログラムを複数組み合わせて利用して複雑な仕事をしようという UNIX と、一つのプログラムで色々なことに対応しようとする Windows)。

[13] 率直に言うと... 見識を疑う。情報科学センターのユーザーは人からものをもらうとしか考えていないのかな？

B　数学版 Gutenberg プロジェクトができないか...

実は私は漠然と夢見ていることがある。数学版 Gutenberg Project, あるいは青空文庫が作れないだろうか？ということである。特に日本人にとって、数学の古典に触れる機会はかなり限られているが (よほど歴史とお金のある大学でないと、19 世紀以前の文献はほとんど持っていない)、それだけに実現できたらどんなに素敵だろうと思う。

権利の問題をどう解決するとか、入力をどうするとか、種となるテキストをどこで調達するとか、とても片手間に出来ることではないのだが...

必要となる技術 (例えば TEX はその一つ) は揃っているような気がするだけに何とかならないかな、と徒然に考える。