

2001年度 卒業論文

指導教授 山本孝一

圧縮・解凍

浜松短期大学

2000H131 山崎聖子

## 目次

- 1 . 圧縮？解凍？
- 2 . なぜデータを圧縮するのか
- 3 . 圧縮の種類
  - 1 . 可逆圧縮
  - 2 . 非可逆圧縮(不可逆圧縮)
- 4 . 可逆圧縮の原理
- 5 . 最後に

## 参考文献

## 1 . 圧縮？解凍？

単に「圧縮」といっても、色々な意味がある。

### あっしゅく【圧縮】

- (1) 物体や気体に圧力を加えて、容積を小さくすること。圧搾。
- (2) 文章などを、縮めて短くすること。
- (3) コンピューターで扱うデータを一定のアルゴリズムを用いて変換し、小さくすること。
- (4) 二つ以上の観念や事象が融合して一つになる、夢などにみられる心理現象。

ここでは、(3)について取り上げようと思う。

逆に、圧縮したデータを元のデータに復元する処理のことを「解凍」または「展開」「伸張」「減圧」「抽出」などという。

## 2 . なぜデータを圧縮するのか

データを圧縮し容量を小さくすることで、ネットワーク上でデータの送受信にかかる時間を短縮したり、ハードディスクなどの記憶装置により多くのデータを記録するために利用する。

## 3 . 圧縮の種類

### 1 . 可逆圧縮

圧縮されたデータを解凍しても、圧縮前のデータと比べて欠落がまったく起こらない圧縮方式。

この可逆圧縮方式を使うと、圧縮前のデータを、完全に復元できる。

拡張子の例：

.lzh (圧縮ファイル)	} 圧縮ファイルの圧縮・解凍には、専用のソフトが必要
.zip (圧縮ファイル)	
.gif (静止画ファイル)	
.PNG (静止画ファイル)	

### 2 . 非可逆圧縮(不可逆圧縮)

多少のデータの欠落を許容する代わりに、劇的に圧縮効率を高めた圧縮方式。データを削ってサイズを小さくするため、同一のデータに完全に復元することが出来ない。

これらのデータはプログラムや文字データと違って、多少のデータの欠損や改変があっても、まったく意味が変わってしまうということはない。このため、非可逆圧縮方式では、人間の視聴覚特性を元に圧縮前のデータに含まれる人間が認識

しにくい成分を削ることにより、極めて効率のよい圧縮を実現している。  
画像や映像、音声などの圧縮に用いられる。  
またデータを削ってサイズを小さくするため、保存を繰り返すとどんどん画質が劣っていく。

拡張子の例： .jpg (静止画ファイル)  
.jpeg (静止画ファイル)  
.mpeg (動画ファイル)  
.mp3 (音声ファイル)

## 4 . 可逆圧縮の原理

基本的には冗長性を利用して圧縮をする。  
例えば、

「ララララー」

をという言葉の場合、「ラ」が5回続くので、

「ラ×5回」

と表現する。

5回繰り返すという定義をどうするかという問題はあるが、原理的にはこうして圧縮をする。またこれをランレングス法という。

5文字連続だと圧縮効果も出てくるが、2、3文字だと圧縮効果は低くなる。  
そこで繰り返しの単位を1文字単位ではなく複数の文字の組み合わせまで拡張する事にする。

「This\_is\_a\_pen.」 (スペース = ( \_ : アンダースコア))

という文で考えると。よく見ると「is\_」が2回繰り返している。  
そこで

「Th(is\_)×2回 a\_pen.」

と表現する。

こうすると繰り返しは1回だが3文字分圧縮できる。

しかしこのようにうまく繰り返しの単位が続けば良いのだが、そうも行かないため「辞書」という物を導入する。今度は、

「This\_is\_his\_pen.」

という文。

やはり「is\_」が3回見られるので、

```
「Th###h#pen.」(辞書：#="is_")
```

と表現する。

今度は繰り返しが3回になるため大分圧縮できた。

長いデータになればこういう組み合わせもたくさん出てくるため、圧縮効率は更に上がると考えられる。そしてこれを **LZW 法** という。

また、モールス符号をルーツとした別の圧縮方法に、**ハフマン法**もある。

## 5 . 最後に

「圧縮・解凍」とは、インターネットを利用していれば、必ずと言っていいほど行き当たる問題だと思う。

現に私自身も行き当たったことがある。

まだ「圧縮・解凍」という言葉がわからなかった時.lzh や.zip のファイルの開き方がわからなくて苦労した経験がある。そしてそこから調べていくうちに「圧縮・解凍」の意味を知り圧縮・解凍ソフトの存在を知り、無事解決することができた。

今では家でも授業でも、普通に圧縮・解凍を使っている。

しかし、その種類と原理までは考えたことはなかった。

今回、「圧縮・解凍」について興味を持ったのは、そういう何気なく使っていたものだからこそ、その中身を知りたいと思ったからだった。また、私達が常に見慣れている.jpg や.mp3 までもが圧縮ファイルだとは知らなかった。こんな身近なところにまで使われていたのかと思うと、いかに汎用的なものだったのかを実感した。

何気なく使用していたものが実は複雑で、さまざまな構造を持っている。これは私達が気づかないところに考えられないほどたくさん置かれている。

パソコンは特にそうだ。

便利なものこそ、その裏側には複雑で精密な構造がある。もちろん日常生活の上ではその構造は知らなくてもいいのだが、あえて少しだけ覗いてみる。

すると意外に、私達が理解できることがたくさん出てくる。もちろん完成形に至るまでは応用に応用を重ねなければいけないのだが、ばらばらにしてみれば、思ったほど理解に苦しむことはない。そしてそこでわかったことは、私達が難しいと思っていることでも、考えて組み合わせて洗練すれば、いくらでも答えは導き出せるということ。そしてそこにはいくらでも可能性があるということ。

難しいからわからない。ではなく、難しいから基礎に戻ってみる。

これは私達が最も忘れがちであり、実は一番の解決方法なのではないだろうか。

## 参考文献

情報・通信事典 e-Words

<http://www.e-words.ne.jp/>

便利ツール[国語辞典]

<http://dictionary.goo.ne.jp/cgi-bin/jp-top.cgi>

松ちゃん本舗

<http://www.geocities.co.jp/SiliconValley-SanJose/1892/index.html>

hideのページ

<http://www2s.biglobe.ne.jp/%7Ehkimura/index.htm>

圧縮&解凍

<http://popup2.tok2.com/home/newbie/technic/comp.html>

データ圧縮

<http://www.matsusaka-u.ac.jp/~okumura/compression/>