

レポート

データのデジタル化とコンピュータの原理

平成25年4月16日

愛知県立大学外国語学部

県立 太郎

目次

1. はじめに.....	3
2. デジタルデータとは.....	3
2.1. 標本化.....	3
2.2. 符号化.....	4
2.3. 量子化.....	5
3. デジタルのよさ.....	5
3.1. 完璧な保存とコピー.....	5
3.2. 完璧な伝達.....	5
3.3. データ圧縮.....	5
4. コンピュータの基礎と歴史.....	5
4.1. コンピュータを構成する要素.....	5
4.1.1. CPU.....	5
4.1.2. メモリ.....	5
4.1.3. HDD.....	5
4.1.4. メモリとHDDの比較.....	6
4.2. コンピュータの歴史.....	6
5. おわりに.....	7

1. はじめに

デジタル(digital)方式とは、“すべての情報を 0 と 1 だけで表す”方式であるということは、近年、一般にもよく知られている。ここで言う情報とは、文字、音声、画像などである。これらの情報を 0 と 1 だけで表現することにより、次のようなメリットがある^[1]。

情報の完璧な保存とコピーが可能 … デジタル方式では、コピーによる劣化が生じない。アナログ方式ではコピーを重ねると雑音の付加が避けられず、品質が劣化してしまう。

完璧な伝送 … デジタル方式では通信によって情報の品質が劣化することがない。アナログ方式では雑音の付加が避けられない。

データ圧縮 … デジタル方式では、大量のデータを小さなサイズのファイルに圧縮することが可能である。画像の JPEG 方式はその一例である。

これらの各項目の詳細については、3 章で詳述する。

しかし、例えば音声のような情報は、一体どのように 0 と 1 だけで表されるのだろうか。このレポートは、その疑問からスタートし、情報のデジタル化とともに、それを取り扱うコンピュータの原理と歴史も併せてまとめる。

2. デジタルデータとは

本章では、音声のデジタル化を例にとり、音声という連続信号が 1 と 0 だけで表現されるまでの過程を説明する。

図 1 にその概要を示す。音声とはそもそも、空気の振動である。マイクロフォンは、音声によって生じた空気の振動を、電圧の変化に変換する。この電圧が、標本化・量子化・符号化の 3 つのステップを経て、0 と 1 のみからなるデジタルデータに変換される。それぞれの詳細について、以下に説明する。

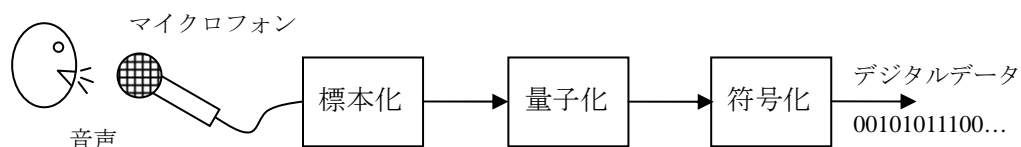
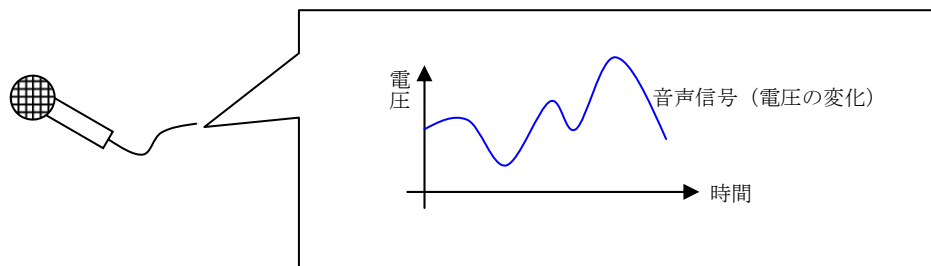


図 1 音声のデジタル化の全体像

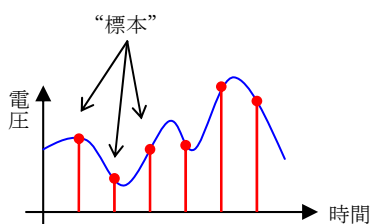
2.1. 標本化

標本化の概念を図 2 に示す。まず、マイクロフォンの出力は、同図 (ア) に示すような電圧の波形として表される。標本化は、同図 (イ) に示すように、この波形の電圧を一定の間隔で調べ、その

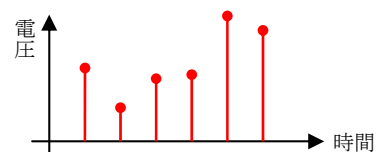
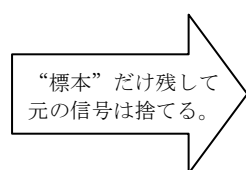
値のみを記録する。この値を標本(sample)と呼ぶ。そして、同図(ウ)に示すように、元の音声信号は破棄してしまう。残るのは標本だけである。



(ア) マイクロフォン出力の電圧のイメージ



(イ) 標本化のイメージ



(ウ) 標本化の結果得られた結果

図 2 標本化の概念

2.2. 符号化

図 1 では、標本化の次に来るのは量子化である。しかし、量子化を理解するためには符号化について知っていることが近道である。このため、本節ではまず、符号化について先に説明する。符号化(coding)は、例として表 1 に示すような符号表を用いて、標本の電圧を符号語、すなわち 0 と 1 だけからなる系列に変換することである。ここで、これらの 0 と 1 の系列を符号語と呼ぶが、“符号”とは、符号語の集合を指す。

表 1 符号表の例

電圧[V]	符号語
1.0	101
0.8	100
0.6	011
0.4	010
0.2	001
0.0	000

しかし、この符号化において、次のような問題が生じる。この符号表にある電圧の値をよく見ると、例えば 1.0[V]の次は 0.8[V]となっていて、その間の値は表のなかにない。しかし実際には、標本はマイクロフォンの出力の電圧を一定の間隔で調べたものであるのだから、1.0[V]と 0.8[V]の間の値となることがあるはずである。この問題に、どのように対処すればよいだろうか。その対処方法が、次節に説明する量子化である。

2.3. 量子化

量子化(quantization)は、前節で述べた問題に対処するために行う、次のような方法である。例えば、標本の電圧が 0.986[V]という値であったものとする。このとき、この値は 1.0[V]と 0.8[V]の中間に位置する値であるが、1.0[V]の方により近い。このため、0.986[V]を 1.0[V]であったものとしてしまう。このように量子化は、すべての標本に符号表による符号化を可能にするために、標本の値を、符号表に掲載されている電圧のうちもっとも近いものに丸めてしまう操作である、と言える。これにより、符号化が可能となり、音声データが 0 と 1 だけで表現され、音声のデジタルデータが得られた。

しかし、量子化は量子化誤差という問題を生む。音声をデジタル化すると、この量子化誤差の影響を避けることができない。量子化誤差は量子化雑音とも呼ばれ、量子化は元の信号に雑音を加えているのに他ならない。このため、音声をデジタル化したのと同時に、音声品質の劣化は避けられない。

量子化雑音の影響は避けることができないが、軽減することは可能である。そのためには、符号表の電圧の間隔を狭くすればよい。しかしそのためには符号語がたくさん必要となる。このため、符号語の桁数を増やすことで、符号語の数を増やし、すべての電圧に符号語を割り当てる。したがって、符号語の桁数が量子化誤差の大きさを決める大きな要素となっているということが出来る。このため、符号語の桁数には特に用語が与えられており、これを“符号の解像度”と呼ぶ。単位はビットである。

3. デジタルのよさ

(省略)

3.1. 完璧な保存とコピー

(省略)

3.2. 完璧な伝達

(省略)

3.3. データ圧縮

(省略)

4. コンピュータの基礎と歴史

(省略)

4.1. コンピュータを構成する要素

(省略)

4.1.1. CPU

(省略)

4.1.2. メモリ

(省略)

4.1.3. HDD

HDD(Hard Disk Drive)は、ハードディスクと呼ばれる金属製の円盤の表面にデータを記憶する装置である。その外観を、図 3 に示す。半導体でできたメモリと異なり、電源を切ってもデータが失われないという特長がある反面、データの読み書きはメモリに比べて低速である。また、ハードディスクは衝撃に弱いという問題がある。これは、次のような理由による。図 4 に示すように、ハードディスクは、アーム先端に設置されたヘッドと称するデータの読み書きを行う部分との間にわずかに隙間を保っていないとてはならない。しかし、衝撃を与えると、ヘッドがハードディスクと接触し

てしまうことで、ハードディスクに傷が生じてデータが失われたり、ヘッドが損傷したりする。このため、ハードディスクの取り扱いには注意を要する。

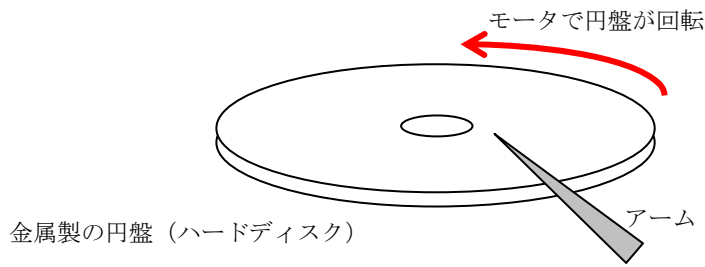


図 3 ハードディスクの外観

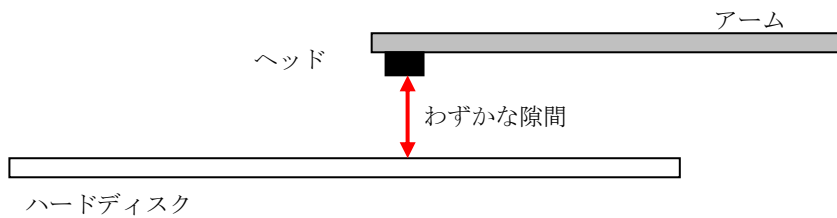


図 4 ハードディスクのアームを横から見た図

4.1.4. メモリと HDD の比較

(省略)

4.2. コンピュータの歴史

世界最初のコンピュータは、1945年に米国・ペンシルバニア大学で開発された ENIAC (the Electronic Numerical Integrator and Computer) である^[2]。このコンピュータは、第2次世界大戦に関連した弾道計算を行うために開発された。図5に示すように非常に多数の真空管から構成され、故障が非常に多かったと言われている。そこから始まるコンピュータの歴史を、2にまとめる。

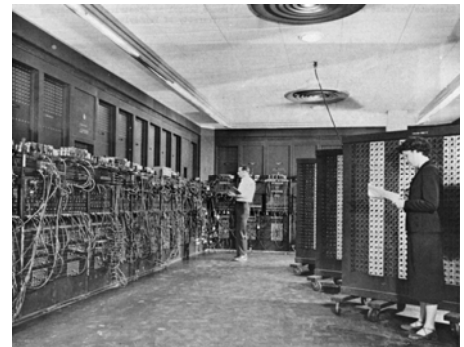


図 5 ENIAC の概観^[3]

2 コンピュータの歴史^[4]

年	コンピュータの歴史	インターネットの歴史
1945年	世界最初のコンピュータ“ENIAC”	
1980年代	OS (operating system)の登場	
1990年代	Windows (マイクロソフト) の出現	WWW と HTML の出現
2000年代		Yahoo や Google の登場

5. おわりに

本レポートでは、情報のデジタル化とは具体的にどのようなことをすることなのか、という疑問から出発して、音声のデジタル化についてまとめた。次に、それを取り扱うコンピュータの原理と歴史をまとめた。このレポートにより、情報のデジタル化からコンピュータの構成についての一連の流れが明らかとなった。

参考文献

- [1] 神谷幸宏, “MATLAB によるデジタル無線通信技術”, コロナ社, pp. 14-15, 2008 年.
- [2] <http://www.seas.upenn.edu/about-seas/eniac/> 2013 年 3 月 31 日 閲覧.
- [3] <http://upload.wikimedia.org/wikipedia/commons/4/4e/Eniac.jpg>
- [4] 愛知太郎, “コンピュータの歴史”, 県立書店, pp. 68-95, 2010 年.