

第6講 文字・音・画像デジタルデータの特徴

教材のダウンロードは以下URLにアクセスし、印刷して受講すると効果的です。

赤字を書き込んでいくとより効果的です！！！！

<https://joho-tarou.com/>



JOHO 太郎「情報 I」スーパー講義

6-1 文字データの表現

文字コード

… 一つの文字に割り当てた番号のこと (例) 「3042」はコンピュータで「あ」を表示

重要! 番号体系のセットは決めた団体と歴史があって、複数できてしまった!

1963年

ASCIIコード

… 米国規格協会ANSI で制定 (7ビット) → 英数字・記号・制御記号

1969年

JIS X 0201

… 日本産業規格JIS で制定 → ASCII+カタカナ (8ビット=1バイト)

~1987年

JIS X 0208

… 日本産業規格JIS で制定 → ASCII+カタカナ + 漢字
(漢字は16ビット → 2バイト文字)

~1982

重要!

シフトJIS

… アスキー・マイクロソフト等で制定 → ASCII+カタカナ + 漢字
(漢字は16ビット → 2バイト文字)

~1991

重要!

EUC-JP

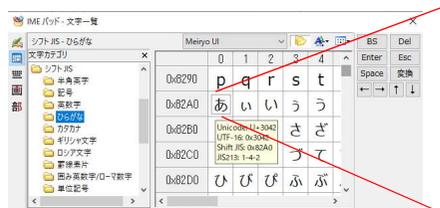
… Open Software Foundation 等で制定・昔のWebで利用

1991年

重要!

Unicode

… ISOの文字コード規格委員会 → 世界中の文字を扱う (~21ビット)



あ	い	い
Unicode: U+3042 UTF-16: 0x3042 Shift JIS: 0x82A0 JIS213: 1-4-2		



- ✓ 同じ「あ」であっても、文字コードセットが異なれば、数値は異なる
- ✓ 文字コードの選択を誤ると、表示できずに「文字化け」が起きる

6-2 音データの表現

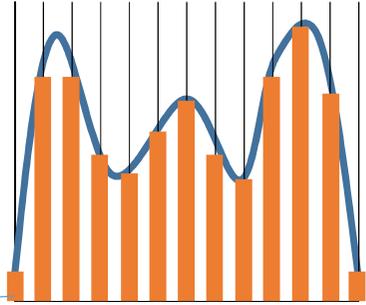
自然界 = アナログ



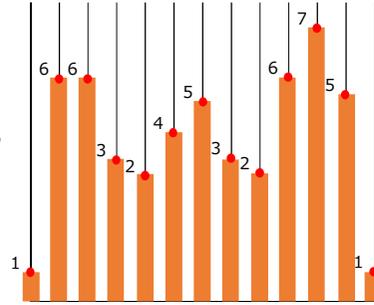
A/D変換



標本化



量子化



符号化

1 6 6 3 2 . . .
001 110 110 011 010 . . .

標本化周波数

... 1秒間に標本化する回数のこと。単位は

Hz

重要!

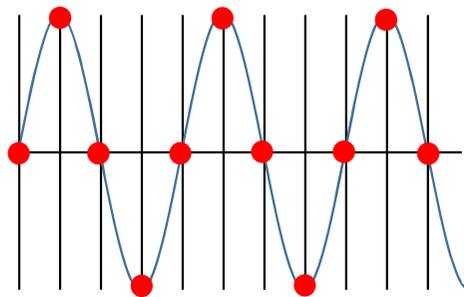
波形が高いほど大きい音

波形がせまいほど高い音

標本化定理

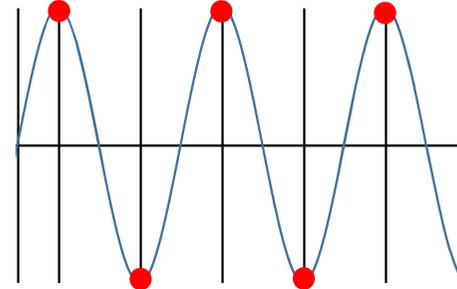
1周期

1周期の1/4で標本化すれば音を再現できる



1周期

1周期の1/2で標本化うまくいけば再現（ギリギリ!）



重要!

1周期の1/2より細かく標本化を取るすなわち、最大周波数の2倍の周波数を超えてサンプリングする必要がある

6-2 音データの表現 ~身近な例~

ジャンル | J-POP 作詞 | - 作曲 | -

このシングルのハイレゾ音質はこちら

 アイドル
ハイレゾ FLAC | 24bit/96kHz
550円

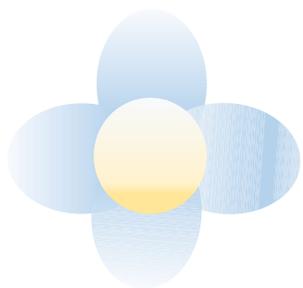


ハイレゾ音質は高音質と言われているけど、どうして？！

96kHz → 1秒間に96000回も標本化している！
(ちなみにCDは44.1kHz)

6-3 画像・動画データの表現

自然界 = アナログ



標本化

ドット / 画素(ピクセル)

量子化

0	2	0
2	5	2
0	2	0

階調

色の濃淡を示す段階値のこと。大きいほど多くの色の濃淡を表せるため、正確な色に近づく

符号化

0 2 0 2 5 . . .
000 010 000 010 101 . . .



ドットを細かくし、1ドット当たりの階調を大きくすると現実に近い画像になるがデータ量は増える

動画データ → 画像を連続に表示



フレーム

... 1枚1枚の画像のこと

フレームレート

... 1秒間に表示するフレームの数 単位は fps

✓ ドットの細かさを表すものとして「解像度」が使われる。dpi 1インチ当たりのドット数が使われる。高いほど、高画質

6-3 画像・動画データの表現（身近な例）

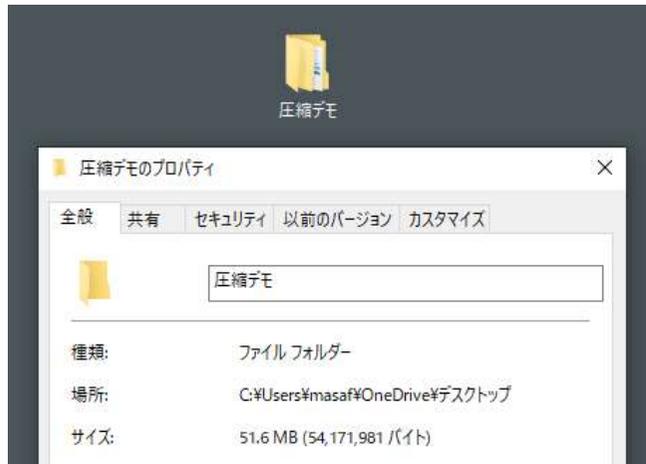

PIXUS TS6330
充実の基本機能。5色独立インクのスタンダードモデル。
5色・独立型・BCI-381シリーズ
3.0型TFT タッチパネル
✓
2WAY給紙：前面カセット給紙・背面給紙
自動両面プリント（はがき非対応）・自動電源ON
コピー、スキャナー(1200dpi CIS)
—
約18秒

スキャナーの商品説明に、1200dpiと書いてある

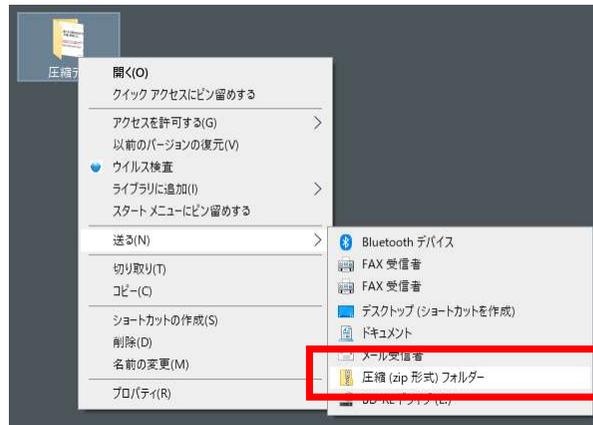


→ 1インチの中に1200ドット標本化。
大きいほど、高画質！

6-4-1 データの圧縮 (可逆圧縮)

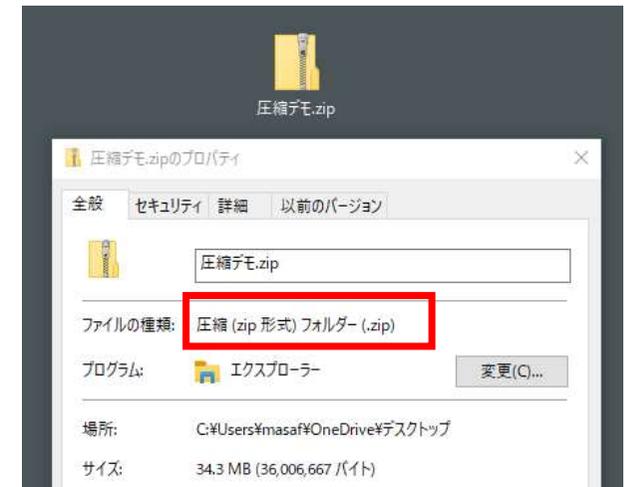


51.6MB



圧縮

展開 / 解凍



34.3MB

重要!

圧縮後のデータ量が元と比較してどのくらい減ったのか

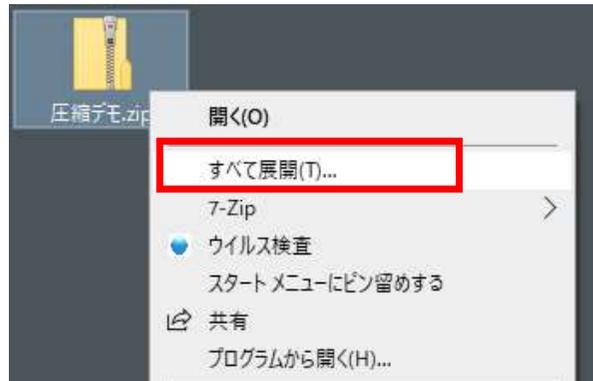
$$\text{圧縮率} = \frac{34.3}{51.6} \doteq 66.5\%$$

重要!

よく使われる圧縮形式

ZIP

RAR



6-4-2 データの圧縮 (非可逆圧縮)

非可逆圧縮

… 圧縮前の状態に戻せない → 戻す必要もない！そのまま利用できるもの
圧縮する事でデータ容量を稼ぐことができる (データの質より利便性を求める)

音声

MP3

… スマホでの音楽視聴で使われている音楽ファイルで用いられる形式

画像

JPEG

… ジェイペグと読む。Web・スマホのカメラ・デジカメで使われている画像の形式



Webで使われるその他の画像

GIF → 可逆圧縮方式、そもそも256色以下の表現なので軽い。

PNG → ピングと読む。可逆圧縮方式のため、JPEGより大きいサイズ

動画

MPEG

… DVDで使われている動画形式(MPEG2), インターネットで使われる(MPEG4)

6-4-3 デジタルデータの特徴

情報をデジタルデータで扱うことの長所・短所を理解しておこう！

デジタルデータの長所

情報は劣化しない

情報の加工が容易

管理が容易

デジタルデータの短所

デジタル化でわずかに情報を失う

著作権者の権利が侵害されやすい

情報は流出しやすい

データ量が膨大になる

第6講 確認問題 問1

コンピュータで使われている文字コードの説明のうち、適切なものはどれですか。

- ① ASCIIコードは、英数字・記号・制御文字およびカタカナに関する規定がある
- ② Unicodeは、多言語を共通して利用できるよう考案された
- ③ EUCは、アルファベット、数字、特殊文字及び制御文字からなる
- ④ シフトJISコードは、UNIXにおける多言語対応で制定され、初期のWeb上において非常によく使用されてきた。

→ 正解は②です。

- ① はJIS X 0201の説明
- ③ はASCIIの説明
- ④ はEUCの説明

第6講 確認問題 問2

音データのデジタル化に関する記述のうち、適切なものはどれか。

- ① 量子化の段階でデータは0と1だけになる
- ② 標本化で使用する単位は、KB (キロバイト)を使用する
- ③ 標本化周波数を11kHzから22kHzにすると、データ量は2倍になる
- ④ 標本化周波数を11kHzから22kHzにすると、音程は高くなる

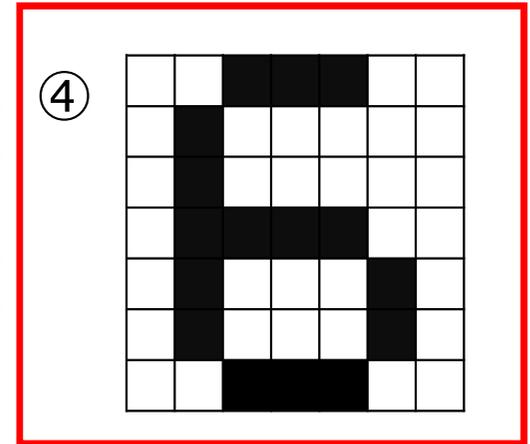
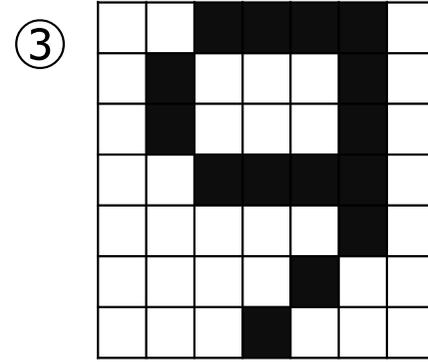
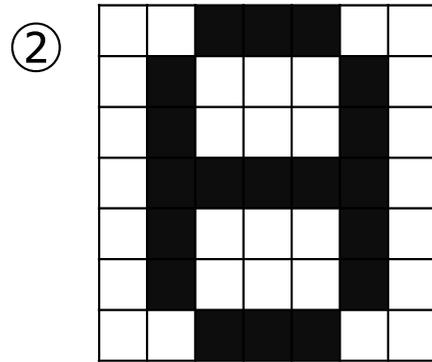
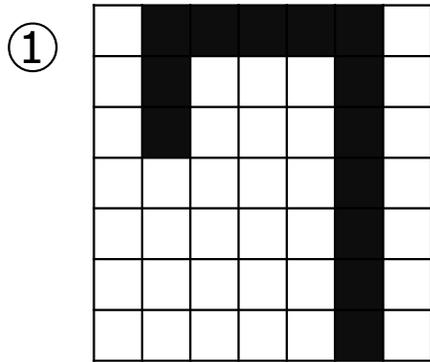
→ 正解は③です。

標本化周波数は、アナログの音源から1秒間にデータを取得する回数を表します。11kHzは11000回、22kHzは22000回を表し、データ取得の回数が2倍となり、データ量も2倍になります。

- ① : 符号化が正しい
- ② : Hz (ヘルツ)が正しい
- ④ : 標本化周波数と音程は関係ありません

第6講 確認問題 問3

画像データを圧縮せずに出力した場合、22文字目から28文字目が、「0111100」となるデジタル画像はどれか。ただし0は白で出力、1は黒で出力とする。



→ 正解は④です。
右図のように格子にナンバーを付け、
22文字目から28文字目に着目し、
白黒黒黒黒白白
0 1 1 1 1 0 0
と一致するものは④のみとなります。

1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31	32	33	34	35
36	37	38	39	40	41	42
43	44	45	46	47	48	49

第6講 確認問題 問4

映像データや音声データで 사용되는、圧縮方式はどれですか。

① JPEG

② RAR

③ ZIP

④ MPEG

→ 正解は④です。

MPEGは動画の圧縮フォーマットで、MPEG1、MPEG2、MPEG4などの規格があります。

①のJPEGは、画像ファイル形式

②、③のZIP・RARは、データ圧縮のファイルフォーマットのことです。

第6講が完了しました！

Good job!

チャンネル登録をお願いします！