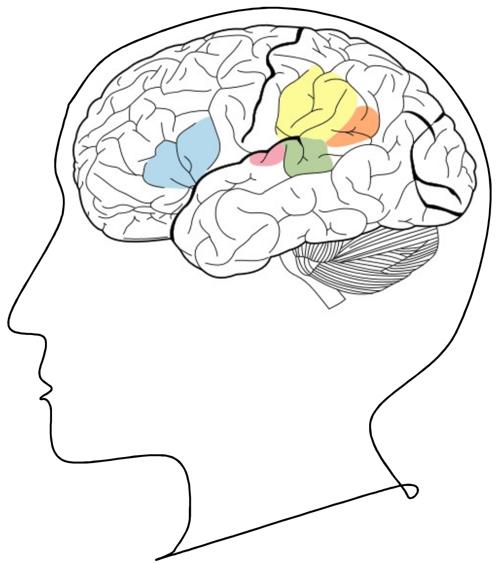
Introduction to Multimedia Compression

National Chiao Tung University
Chun-Jen Tsai
9/15/2014

Data is Everything, and Nothing

- Everyday, we are bombarded by all kinds of information (data)
  - textbooks, news papers, movies, songs, conversations, lectures, preaching, …
- What are the purposes of all the information?
  - To cause chemical reactions in your brains
  - To “duplicate” chemical reactions from one brain to the other
  - To …



Three Aspects of Data

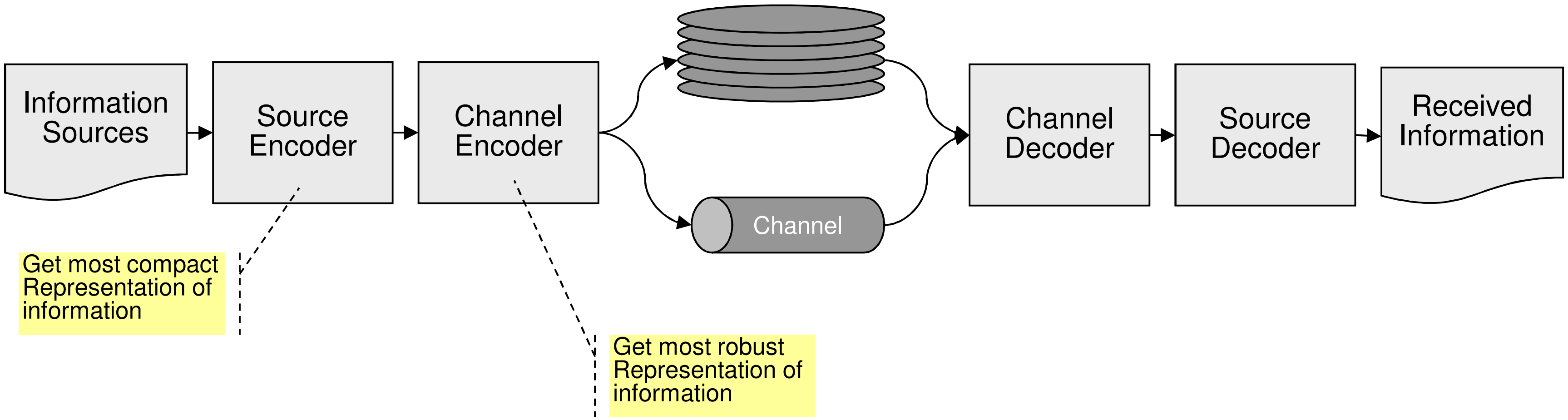
- To distribute information (data) around to serve your purposes, there are three aspects one should worry about:
  - Quantity
  - Reliability
  - Security
- Example: “knowledge of multimedia compression”

The Paper That Starts It All …

- In 1948, Claude E. Shannon published the revolutionary paper, “A Mathematical Theory of Communication.”
  - Later, in 1949, a book was published based on this paper, but the first word of the title was changed from “*A*” to “*The*”
- The paper provides many insights into the essence of the communication problem
  - In particular, Shannon perceived that all communication is essentially digital !

Data Distribution Systems

- ❑ Shannon was the first person to partition a communication system as follows:



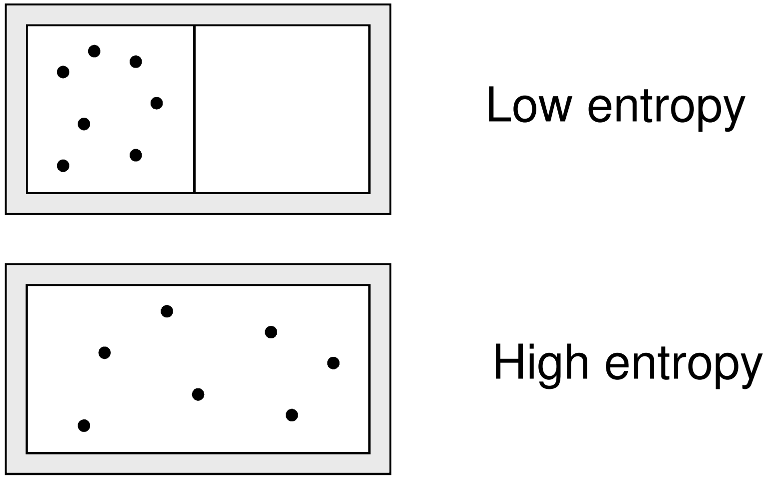
→ The information content of a source and the information capacity of a channel can be identified using the concept of ***entropy***

The Origin of Information Theory

- The term “entropy” was first used in thermodynamics and in statistical mechanics
  - Some people think that information theory grew out of statistical mechanics because L. Szilard applied an idea of information to solve a physical problem in 1929
- However, Shannon’s work evolved from the field of electrical communication
  - “entropy” was used in information theory merely due to its mathematical analogy with the entropy of statistical mechanics

Entropy in Thermodynamics

- In thermodynamics, entropy is a measure of thermal energy of a body of gas



- Statistical mechanics says that an increase in entropy means a decrease in predictability

Linking Back to Information Theory

- The complexity of a system depends on our knowledge of the system; the more we know about the system, the less words we need to “describe” the system
- In information theory, the amount of information conveyed by a message increases as the amount of uncertainty as to what message actually will be produced becomes greater

Some “Information”

- Check the “entropy” of the following messages
  - My dog cannot fly
  - My dog runs faster than a chicken
  - My dog is a lady dog
  - My dog runs slower than a chicken
  - My dog can sing
- It seems that, a rare message carries more information than a common message

Frequency-based Coding

- Morse code
  - Invented in 1838 by Morse for electrical telegraph, and expanded by Vail in 1844
  - To shorten the transmission of messages, English text was coded based on relative frequencies of occurrence
  - The efficiency of Morse code can only be improved by 15% using modern theory†
  - Questions: efficient for all languages?

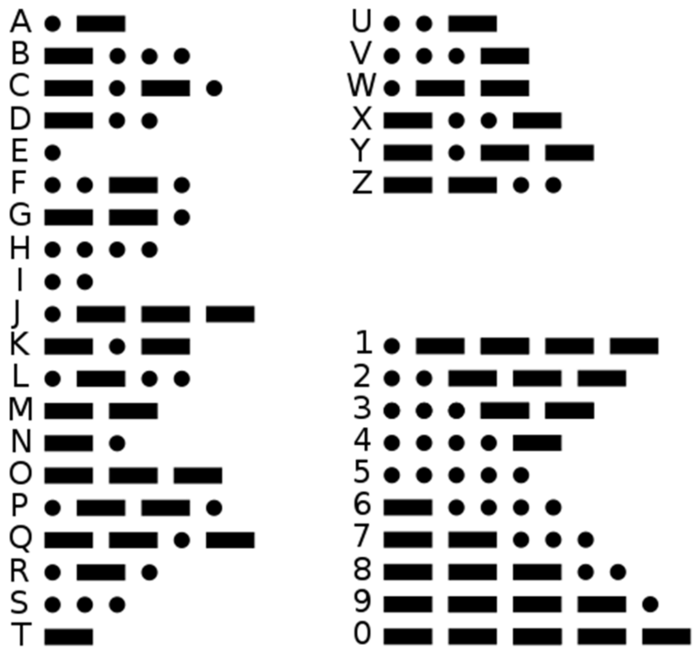
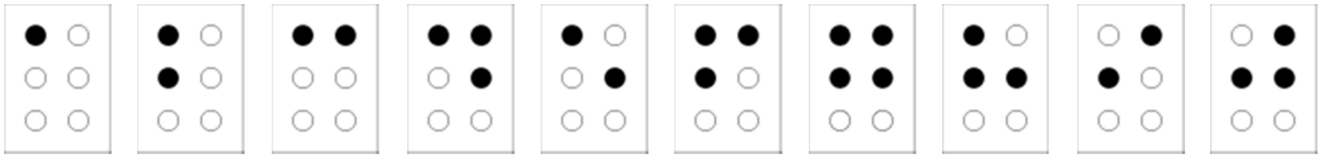


fig. ref.: wikipedia

† J. R. Pierce, *An Introduction to Information Theory*, *2nd. Ed.*, Dover Publications, 1980.

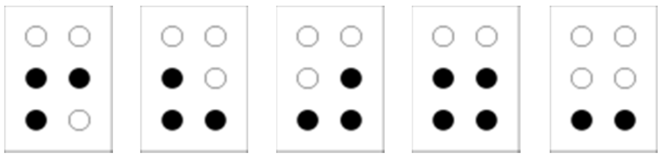
Context-based Coding

- Braille code, by Louis Braille in 1825
- Grade 1 Braille
  - Letters and numbers



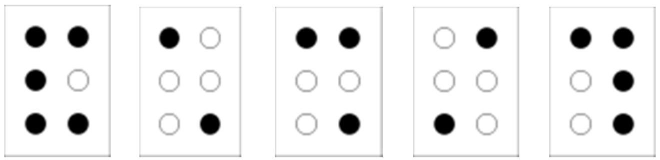
A or 1 | B or 2 | C or 3 | D or 4 | E or 5 | F or 6 | G or 7 | H or 8 | I or 9 | J or 0

  - Symbols



! | “ or ? | ” | (or) | –

- Grade 2 Braille



AND | CH | SH | ST | TH

Model-based Coding

- Statistical structure is not the only way of compression. Describing “things” using “models” is usually less wordy
- For example, what is the minimal precise description of π?
  - Shannon’s idea – the unpredictability of patterns of digits in π
  - Kolmogorov’s idea – the size of a program that computes π

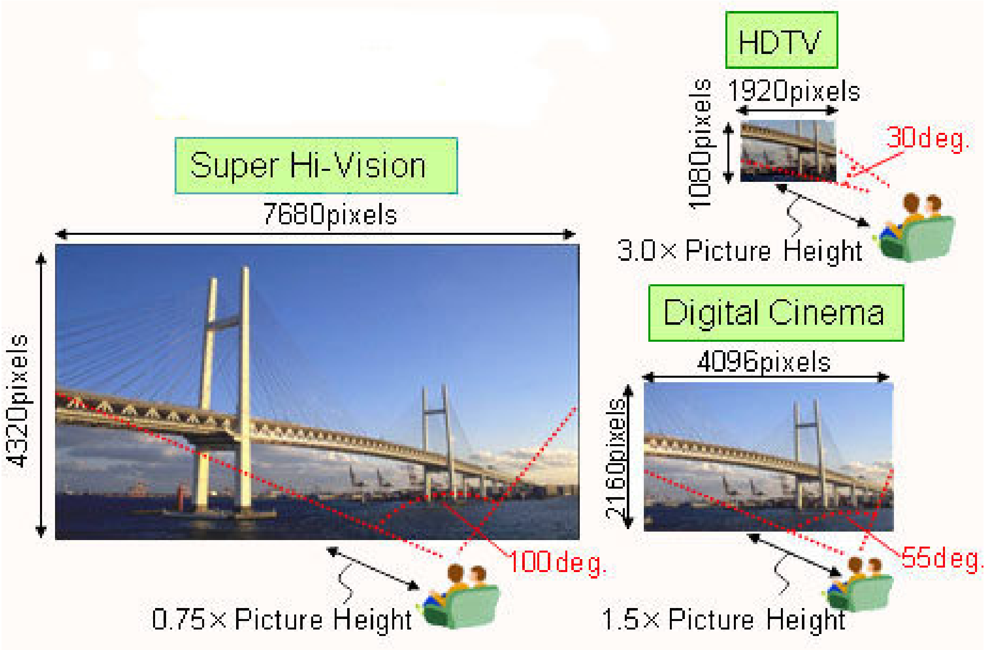
How Large Is the Amount of Data?

- 1 second of CD audio:
  - 44100 samples $\times$ 2 channels $\times$ 16 bits/sample
    = 1,411,200 bits
- 1 second of 1080p HD video:
  - 1920 $\times$ 1080 pixels $\times$ 3 color channels $\times$ 8 bits/color sample
    $\times$ 30 frames
    = 1,492,992,000 bits

Sometimes, large data amount is a technique against piracy

The Future Is Here Already

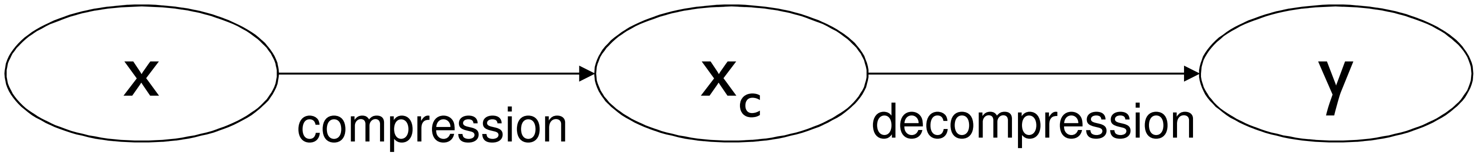
- Super Hi-Vision† (8K system)
  - 7680 × 4320 = 33 Mega pixels per frame
  - Physical data rate (video-only): 180-600 Mbps



† http://www.nhk.or.jp/digital/en/super_hi/

Data Compression Concept

- X – original data, X_c – compressed representation, y – reconstruction



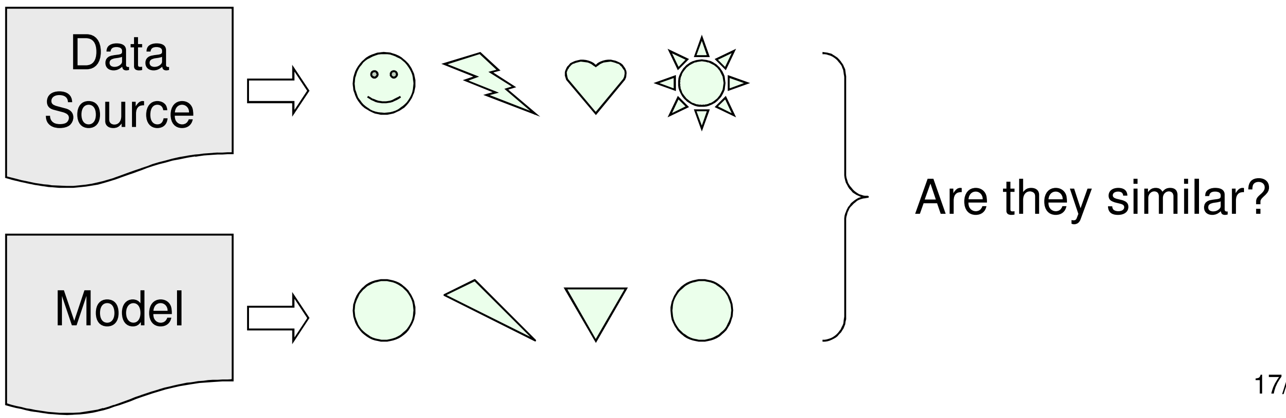
- Lossless compression: when y is equal to x
- Lossy compression: when y is different from x
- Compression ratio: $|X| : |X_c|$ or $(|X| – |X_c|/ |X|) *100\%$
  - For example, $|X| = 65536$ bytes, $|X_c| = 16384$ bytes, the compression ration is 4:1 or 75%.
- Data rate: for time-varying data, the number of bits per second (or sample) required to represent the data

Lossless and Lossy Compressions

- Text compression techniques are often lossless
  - Any counter examples?
- Image, audio, video compression techniques are often lossy
  - Any counter examples?
- Distortion: the difference between the original and the reconstruction
  - If the distortion is small, we say the “quality” or “fidelity” is high. Or, we say the reconstruction is a “high-definition” copy of the original

Modeling and Coding

- One of the most powerful tools in data compression is called “data modeling”
  - Model – a systematic way to describe data
- A common data compression scheme is to “encode” a description of the model, and a description of how the data differ from the model (aka, residual)
  - By encode, we mean to put the data in binary digits



Example 1: Linear Model

- Data sequence y_i,

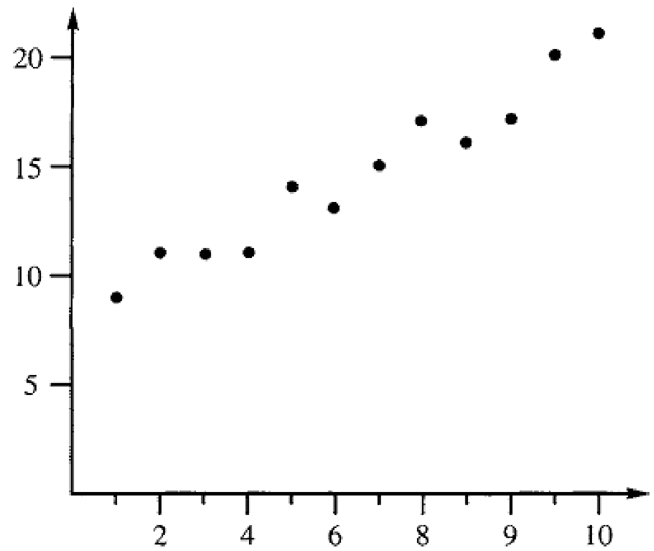
9	11	11	11	14	13	15	17	16	17	20	21	$\cdots$

- Model: $\hat{y}_n = n + 8, n = 1, 2, \ldots$

  Model parameter: $1, 8$

- Residual:

  $e_n = y_n - \hat{y}_n$
  $= 0, 1, 0, -1, 1, -1, 0, \ldots$



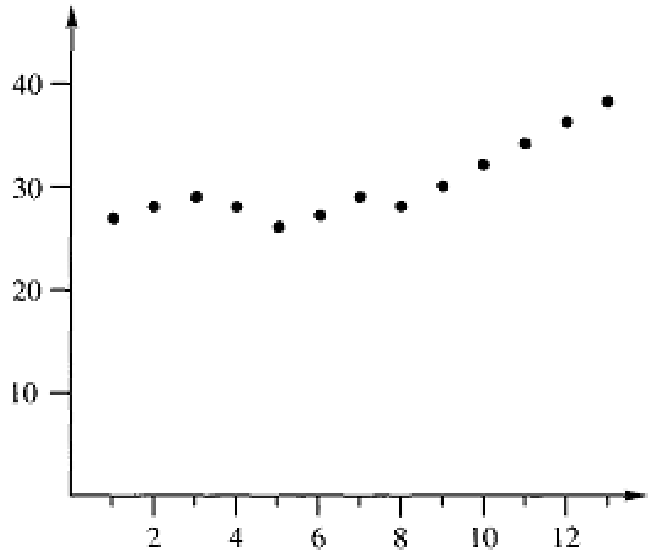
Example 2: Differential Model

- Data sequence y_i,

27	28	29	26	27	29	28	30	32	34	36	38	. . .

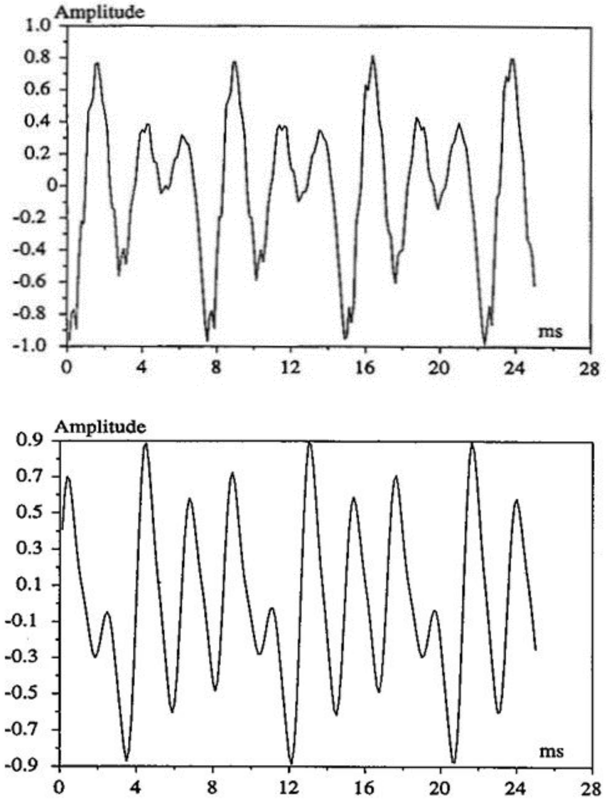
- Model: $\hat{y}_1 = 0, \hat{y}_n = y_{n-1}, n = 2, 3, \ldots$
- Residual:

$$e_n = y_n - \hat{y}_n$$
$$= 27, 1, 1, -1, -2, 1, 2, \ldots$$

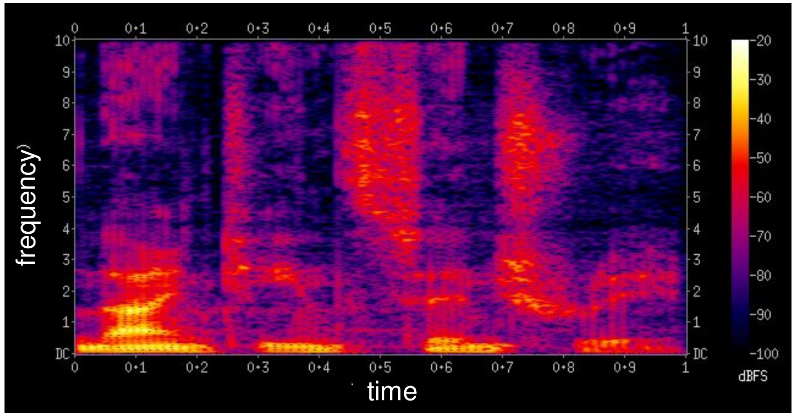


What About Speech Models?

- Typical speech signals:



waveform

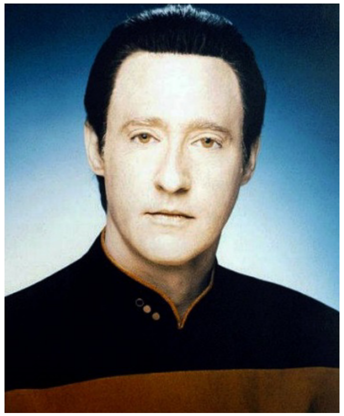


spectrogram

What About Image Models?

- Typical image signals:






Example 3: Variable Length Coding

- Given a sequence of symbols:

aƀarrayaranƀarrayƀranƀ farƀfaarƀfaaarƀaway

  - If fixed length coding (FLC) is used: 3 bits per symbol
  - If variable length coding (VLC) is used: 2.58 bits per symbol
    → 1.16 : 1 compression ratio

TABLE 1.1 A code with codewords of varying length.

a	1
n	001
ƀ	01100
f	0100
n	0111
r	000
w	01101
y	0101