Embedding Efficiency

The average number of embedded bits per unit distortion.

- LSB: $2 = 1/0.5$.
  - 1 bit: for a uniform distribution binary sequence.
  - Change: 50% of chance to change.
  - Efficiency:

$$\frac{1}{0.5}.$$

→ embed 1个bit

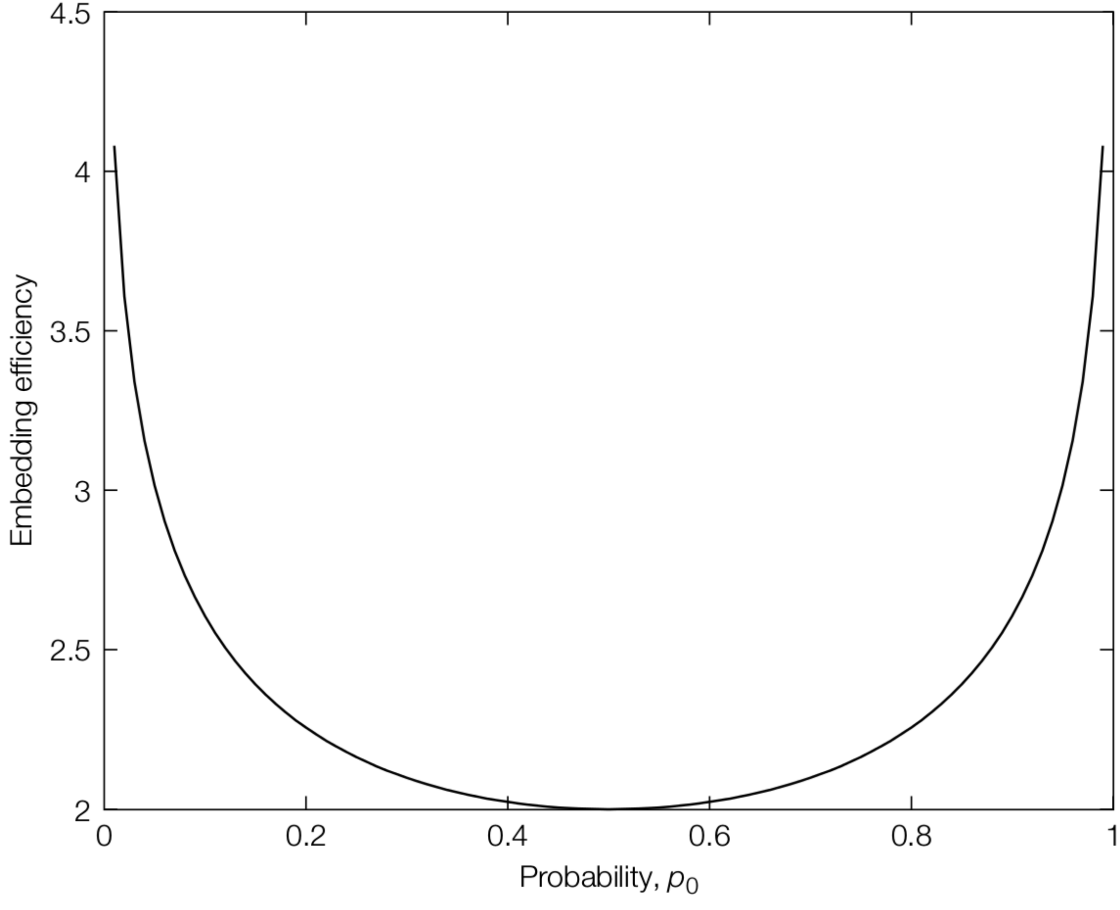
→ 50%的概率改动

Embedding Efficiency

The average number of embedded bits per unit distortion.

- LSB: $2 = 1/0.5$.
- Model Based:
  - Information:
  $$H(p_0) = -p_0 \log_2 p_0 - (1 - p_0) \log(1 - p_0).$$
  - Change:
  $$p_0(1 - p_0) + (1 - p_0)p_0 = 2p_0(1 - p_0).$$
  - Efficiency:
  $$\frac{-p_0 \log_2 p_0 - (1 - p_0) \log(1 - p_0)}{2p_0(1 - p_0)}.$$

Illustration

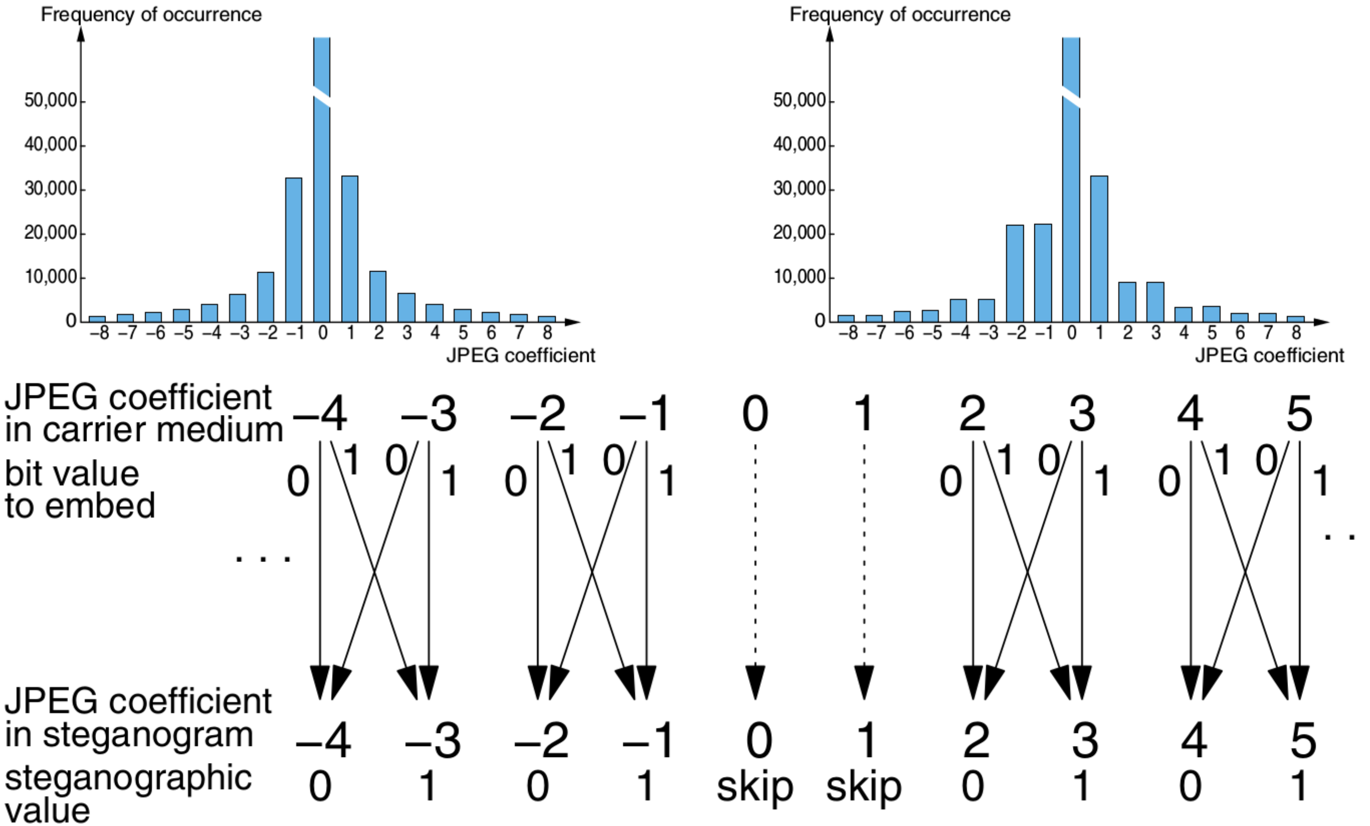


Notice the arithmetic decompress.

Masking Embedding as Natural Processing

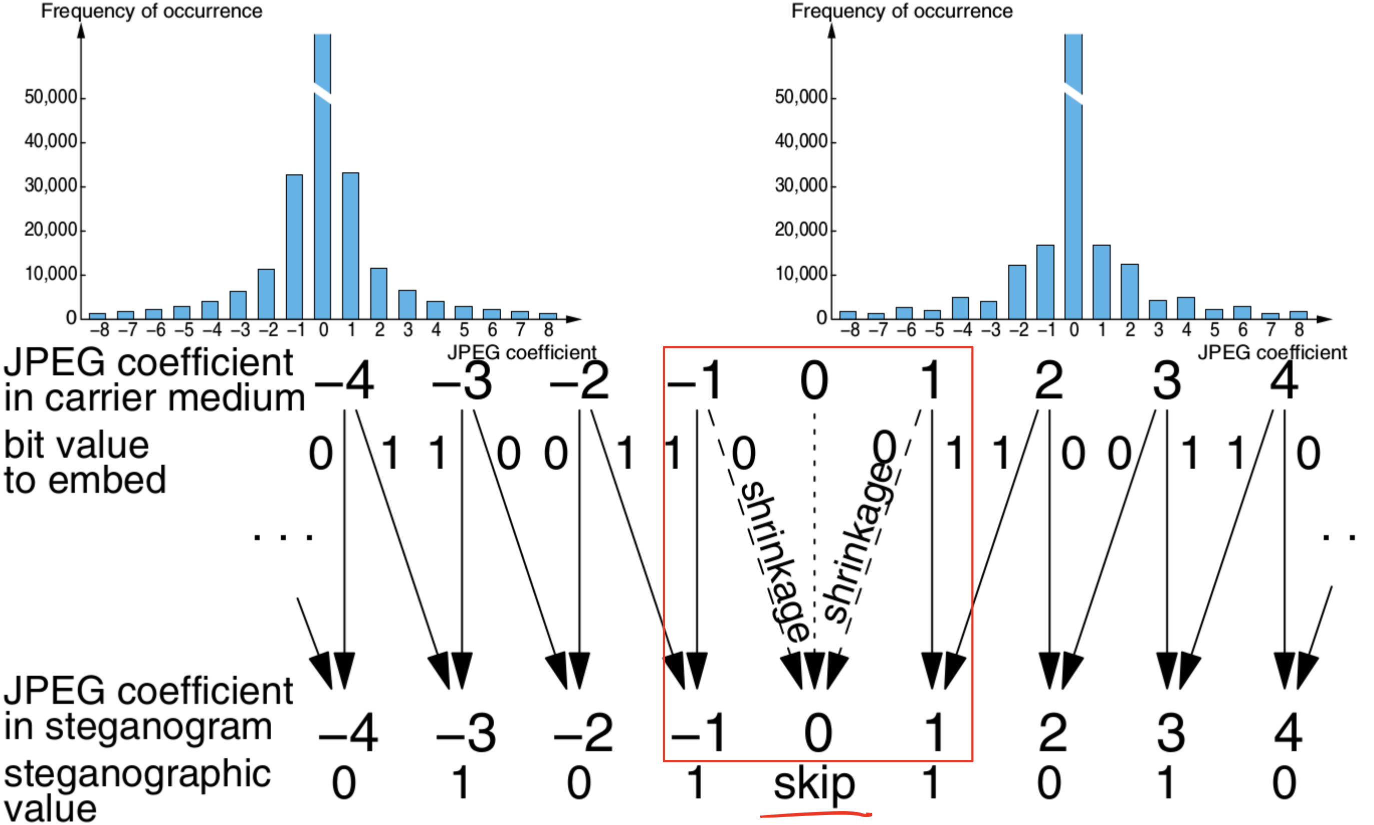
- Preserving statistics
  - Losing capacity.
- Mimicked some natural process
  - F3, F4, F5, ...

Jsteg



Frequency of occurrence
50,000
40,000
30,000
20,000
10,000
0
−8 −7 −6 −5 −4 −3 −2 −1 0 1 2 3 4 5 6 7 8
JPEG coefficient
Frequency of occurrence
50,000
40,000
30,000
20,000
10,000
0
−8 −7 −6 −5 −4 −3 −2 −1 0 1 2 3 4 5 6 7 8
JPEG coefficient
JPEG coefficient in carrier medium
−4 −3 −2 −1 0 1 2 3 4 5
bit value to embed
0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1
. . .
. . .
JPEG coefficient in steganogram
−4 −3 −2 −1 0 1 2 3 4 5
steganographic value
0 1 0 1 skip skip 0 1 0 1

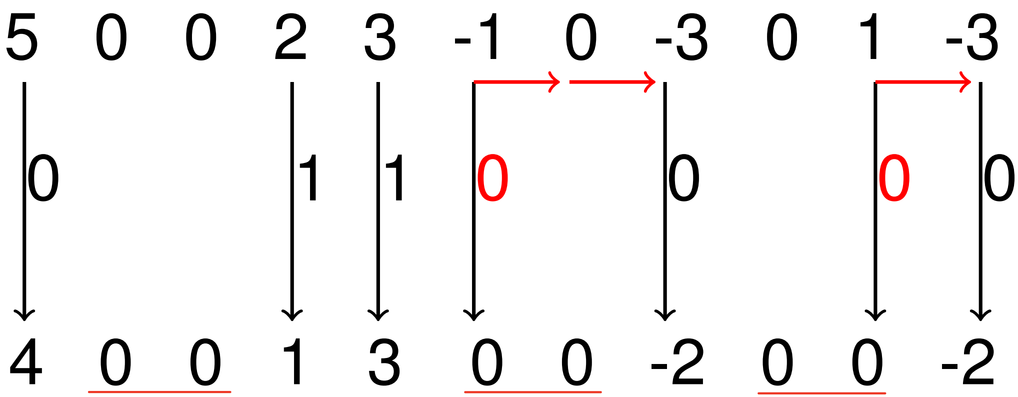
F3



解决 → 难以区分0是本身为0还是1/-1

F3 Algorithm

考试可能会考



Embedding 01100.

跳过

What Is the Problem in F3?

In normal work

- Decreasing

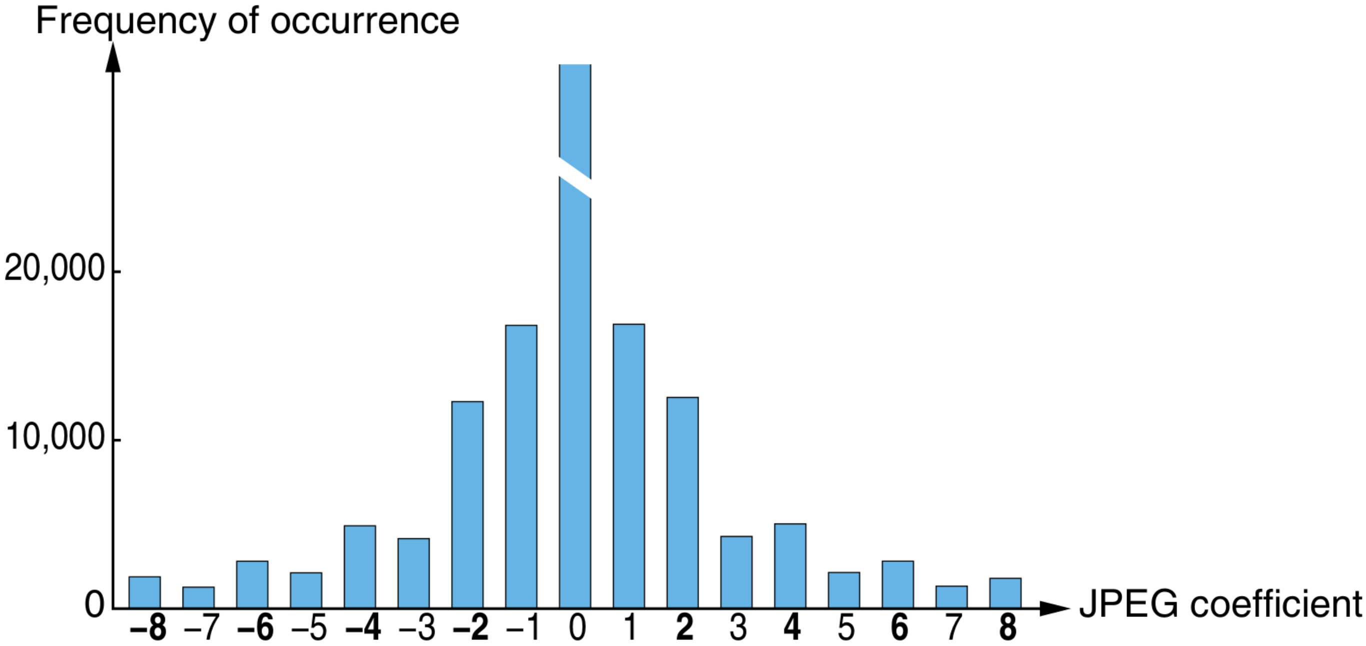
$$P(2i-1) > P(2i).$$

In Steganographic work

- More on even.

$$P(2i-1) < P(2i).$$

Defects of F3

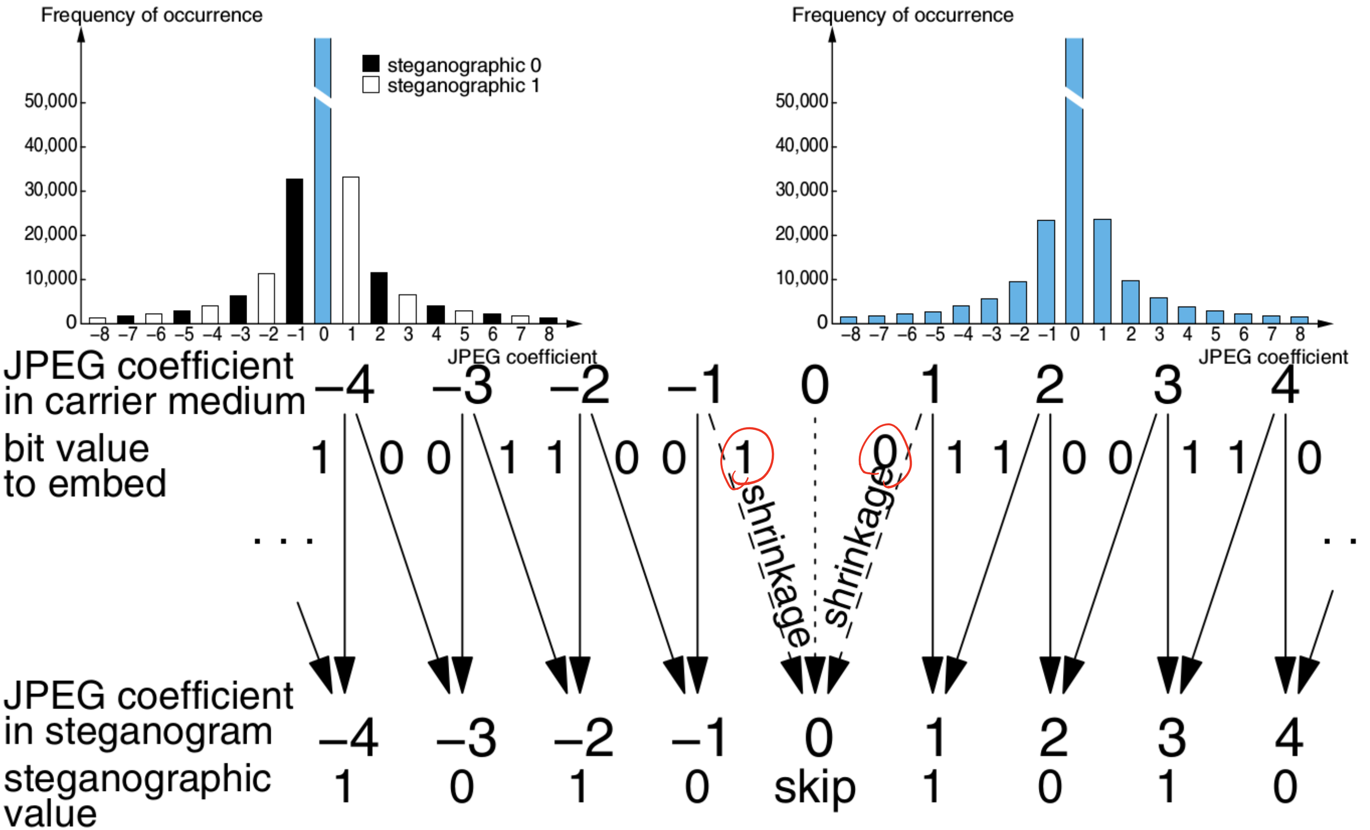


Reason

Repeated embedding after shrinkage. 0会变多.

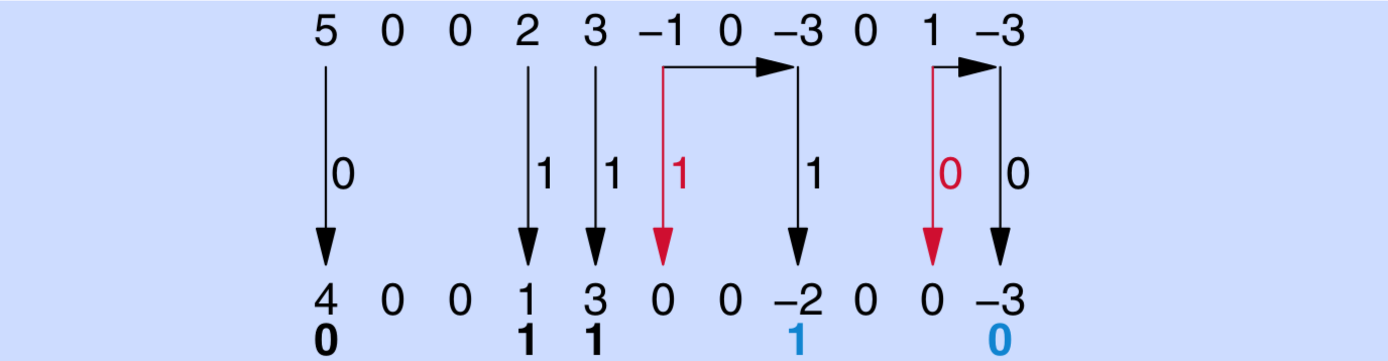
- Happens for embedding 0 only.
- Equivalent to add more 0 into the message code.

F4

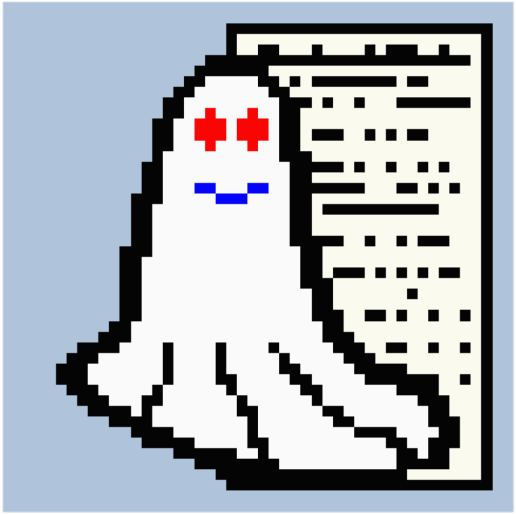
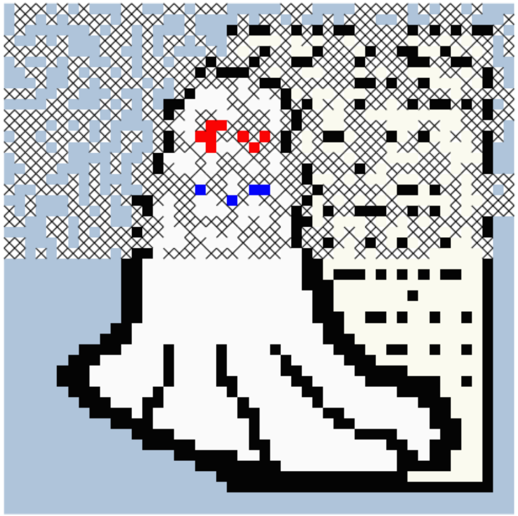

Frequency of occurrence
50,000
40,000
30,000
20,000
10,000
0
–8 –7 –6 –5 –4 –3 –2 –1 0 1 2 3 4 5 6 7 8
JPEG coefficient
steganographic 0
steganographic 1
Frequency of occurrence
50,000
40,000
30,000
20,000
10,000
0
–8 –7 –6 –5 –4 –3 –2 –1 0 1 2 3 4 5 6 7 8
JPEG coefficient
JPEG coefficient in carrier medium
–4 –3 –2 –1 0 1 2 3 4
bit value to embed
1 0 0 1 1 0 0 1 0 1 1 0 0 1 1 0
shrinkage
shrinkage
. . .
. . .
JPEG coefficient in steganogram
–4 –3 –2 –1 0 1 2 3 4
steganographic value
1 0 1 0 skip 1 0 1 0

F4 Algorithm

- Steganographic interpretation
  - Positive coefficients: LSB
  - Negative coefficients: inverted LSB
- Skip 0, adjust coefficients to message bit
  - Decrement positive coefficients
  - Increment negative coefficients
  - Repeat if shrinkage occurs

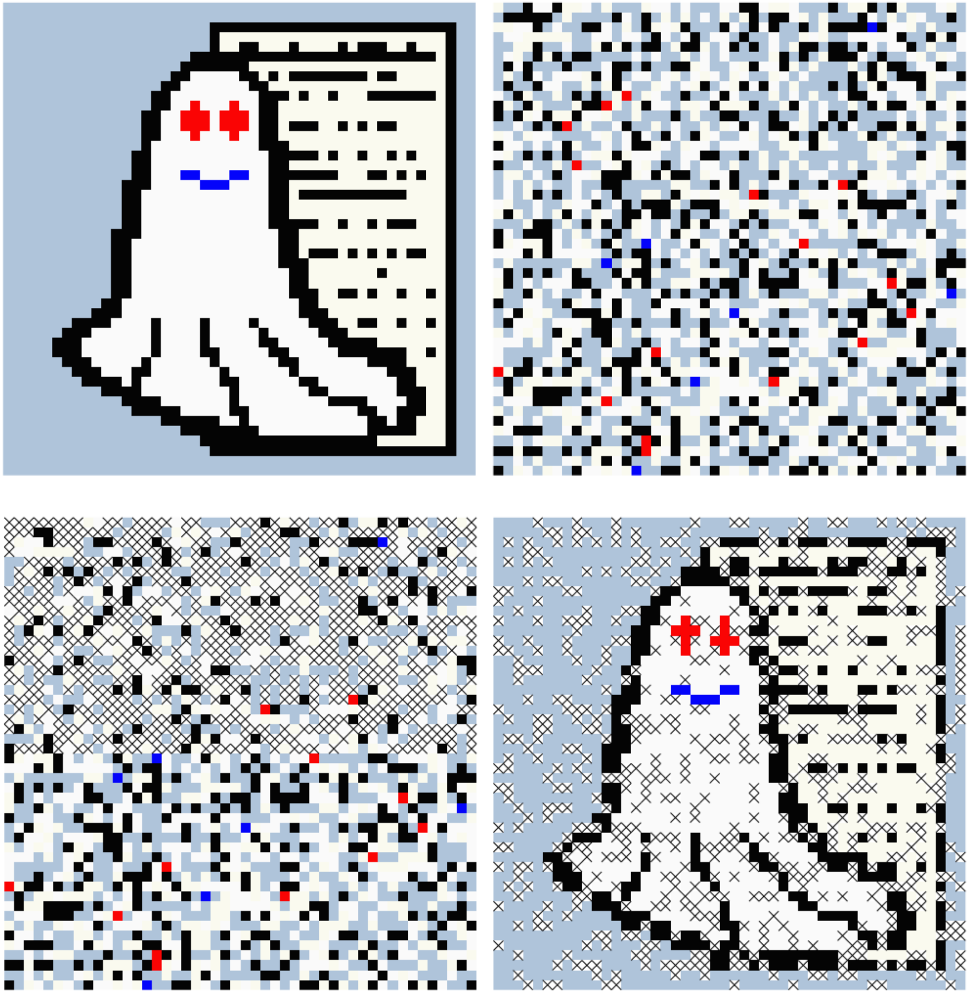


F4 Defects

Compare similar blocks or reverse fitting GCD.

Random Walk



More Efficiency?

考：5% shrinkage 那实际上会有多少 change

A：$5\% + \frac{95\%}{2} = 52.5\%$.

Example: Embedding 1736 bits

- F4: 1157 changes.
- F5: 459 changes by matrix encoding.
  - Embedding efficiency: 3.8 bits per change.

可能考F5.

Matrix Encoding

Embedding b_1, b_2 to x_1, x_2, x_3 with at most 1 change.

$$b_1 = LSB(x_1) \text{ XOR } LSB(x_2)$$
$$b_2 = LSB(x_2) \text{ XOR } LSB(x_3)$$

- Four equal probability cases.
- Change x_i accordingly.

Example

$$b_1 = LSB(x_1) \text{ XOR } LSB(x_2)$$
$$b_2 = LSB(x_2) \text{ XOR } LSB(x_3)$$

0,0	1,0	0,1	1,1
/	$\bar{x}_1$	$\bar{x}_3$	$\bar{x}_2$

Efficiency:

$$2/(3/4) = 8/3 > 2.$$

A Hamming Code

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Presentation: Matrix Embedding

- The idea of parity matrix.
- Efficiency.

Upper Bound on Embedding Efficiency

For a message set $\mathcal{M}$, in a n-pixel image, what is the minimal number of change R (in the sense of expectation).

- The bound of $\frac{\log_2 |\mathcal{M}|}{R}$: efficiency.
  - Larger means better efficiency.
  - The upper bound indicates the optimal situation.

Just Some Math 不考

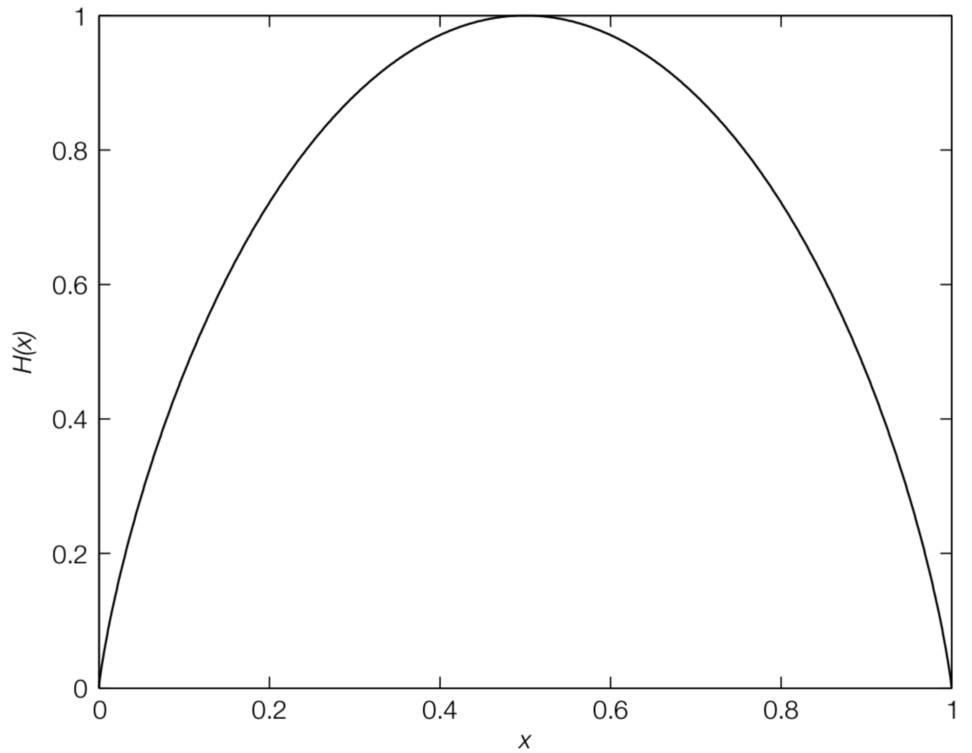
$$\log_2 |\mathcal{M}| \leq \log_2 \sum_{i=0}^{R} \binom{n}{i} 2^i$$

$$\leq nH(R/n)$$ information theory

$H(x)$

Binary entropy function

$$H(x) = -x \log_2 x - (1-x)\log_2(1-x).$$

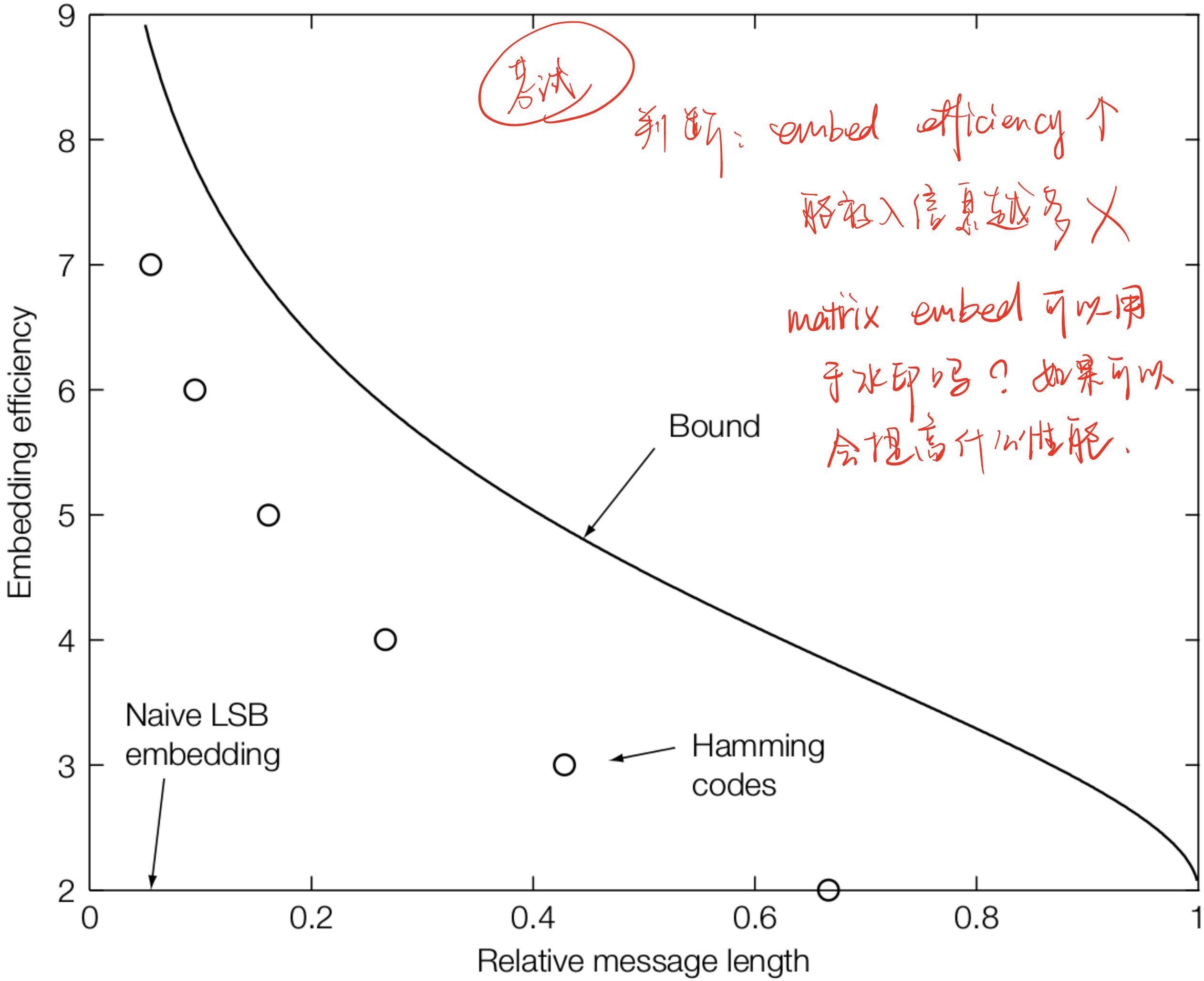


Continue the Math

$$
\begin{aligned}
\alpha = \frac{\log_2 |\mathcal{M}|}{n} &\le H(R/n) \\
\frac{n}{R} &\le 1/H^{-1}(\alpha), \quad H^{-1} \in [0, 0.5] \\
\frac{\log_2 |\mathcal{M}|}{R} \frac{n}{\log_2 |\mathcal{M}|} &\le 1/H^{-1}(\alpha) \\
e = \frac{\log_2 |\mathcal{M}|}{R} &\le \frac{\alpha}{H^{-1}(\alpha)}.
\end{aligned}
$$

- α: relative message length.
- e embedding efficiency.

Illustration



考试

判断：embed efficiency ↑ 嵌入信息越多 ×

matrix embed 可以用于水印吗？如果可以会提高什么性能.

Selection Rule

Choose the parts/locations to change.

- Known for both side: shared.
- Only known for sender: nonshared.

Nonshared Selection Rule

Motivation:

- In JPEG compress:
  - DCT: float value.
  - Round into integer.
- To minimize the change:
  - Choose values have largest rounding error to change, e.g. 5.47:
    - to embed 0: $5.47 \to 6, +0.53$.
    - to embed 1: $5.47 \to 5, -0.47$.
- More like normal compress procedure, but
  - How recipient detect the message?

Other Cases

- Adaptive steganography
  - If the neighborhood has certain property ...
  - But embedding may change the property.
- Eg. using the pixels with largest neighbor variance.