Steganography

Jenny Hong   Stephen Boyd

EE103
Stanford University

November 30, 2016

Steganography

- goal: send a secret message embedded in an image (or text, audio, video, . . .)
- sender modifies the image to incorporate the secret message
- modified image should look like the original one
- message recipient decodes message from the modified image
- we'll look at some simple methods

The message

- let's send one byte, which is an integer between 0 and 255
- one byte can encode a character, like 'c' or '!'
- one byte is represented by its 8-bit boolean expansion, *e.g.*,

$$00101011 \sim 2^5 + 2^3 + 2^1 + 2^0 = 32 + 8 + 2 + 1 = 43$$

- we'll represent one byte as an 8-vector s, with each s_i 0 or 1
- so, *e.g.*, $s = (0, 0, 1, 0, 1, 0, 1, 1)$ represents byte 43

Encoding and decoding

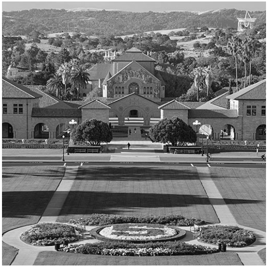
- sender changes original image x (an N-vector) to $x + As$
- A is an $N \times 8$ matrix known to sender and receiver
- entries of A are small enough that x and $x + As$ look very similar
- receiver recovers message using a left inverse B of A:

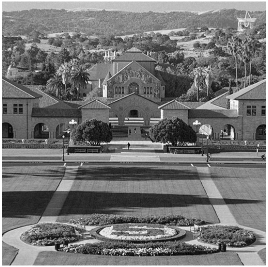
$$\tilde{s} = B(x + As) = Bx + BAs = Bx + s$$

- the first term Bx is 'noise', and we hope it is small compared to s
- our final guess of the sent message is $\hat{s} = \text{round}(\tilde{s})$
- as long as $|(Bx)_i| < 1/2$, we won't make an error, and $\hat{s} = s$
- want B small, so we'll use $B = A^\dagger$
- trade-off: if A is small, then $x + As$ is very near x, but then B is large, and Bx can be large enough to give errors in decoded image

Example

- image is 100×100, so $N = 10000$
- matrix $A = \alpha \tilde{A}$ where $\tilde{A}$ has random entries in $\{-1, 1\}$
- α is a scaling factor
- message represented by vector $s = (0, 0, 1, 0, 1, 0, 1, 1)$
- original image (left) and image with message (right), $\alpha = 0.01$

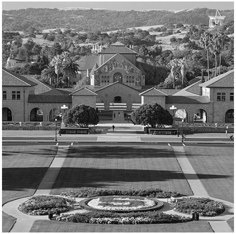




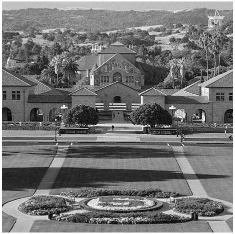
- $\tilde{s} = (-0.10, -0.06, 0.87, -0.12, 1.1, 0.11, 1.0, 1.0)$
- rounding recovers message: $\hat{s} = s$

Scaling α

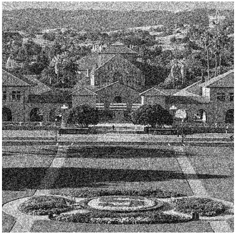
track how many errors are made in recovering the message



$\alpha = 0.001$
1 error



$\alpha = 0.01$
0 errors



$\alpha = 0.1$
0 errors

Some raw images

Images with message encoded