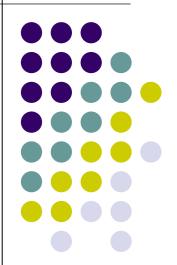
Image Steganography and Steganalysis

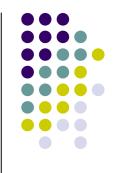


Outline



- Steganography history
- Steganography and Steganalysis
- Security and capacity
- Targeted steganalysis techniques
- Universal steganalysis
- Next generation practical steganography
- Conclusion

Steganography



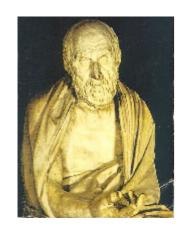
- Steganography "covered writing".
- For example (sent by a German spy during World War I),

Apparently neutral's protest is thoroughly discounted and ignored. Isman hard hit. Blockade issue affects pretext for embargo on byproducts, ejecting suets and vegetable oils.

Pershing sails from NY June I.

Ancient Steganography





Herodotus (485 – 525 BC) is the first Greek historian. His great work, The Histories, is the story of the war between the huge Persian empire and the much smaller Greek citystates.

Herodotus recounts the story of **Histaiaeus**, who wanted to encourage **Aristagoras of Miletus** to revolt against the Persian king. In order to securely convey his plan, Histaiaeus shaved the head of his messenger, wrote the message on his scalp, and then waited for the hair to regrow. The messenger, apparently carrying nothing contentious, could travel freely. Arriving at his destination, he shaved his head and pointed it at the recipient.

Ancient Steganography





Pliny the Elder explained how the milk of the thithymallus plant dried to transparency when applied to paper but darkened to brown when subsequently heated, thus recording one of the earliest recipes for invisible ink.

Pliny the Elder. AD 23 - 79

The **Ancient Chinese** wrote notes on small pieces of silk that they then wadded into little balls and coated in wax, to be swallowed by a messenger and retrieved at the messenger's gastrointestinal convenience.



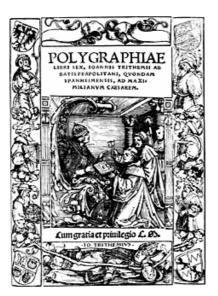
Renaissance Steganography





Johannes Trithemius (1404-1472)

1518 Johannes Trithemius wrote the first printed book on cryptology. He invented a steganographic cipher in which each letter was represented as a word taken from a succession of columns. The resulting series of words would be a legitimate prayer.





Renaissance Steganography



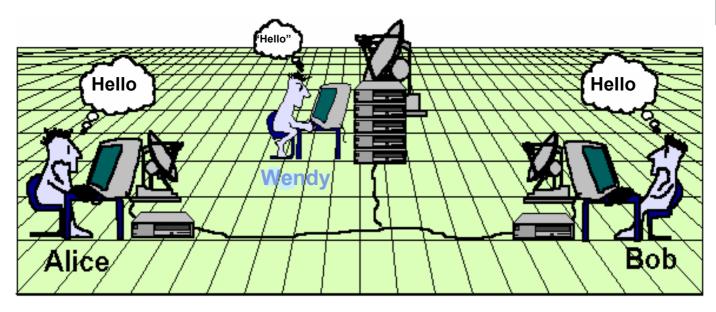


Giovanni Battista Porta (1535-1615)

Giovanni Battista Porta described how to conceal a message within a hardboiled egg by writing on the shell with a special ink made with an ounce of alum and a pint of vinegar. The solution penetrates the porous shell, leaving no visible trace, but the message is stained on the surface of the hardened egg albumen, so it can be read when the shell is removed.

Modern Steganography - The Prisoners' Problem

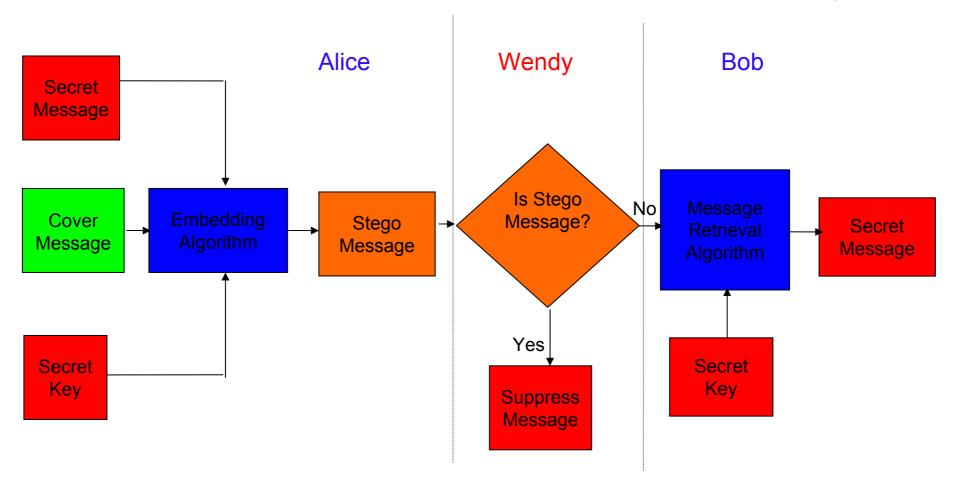




- Simmons 1983
- Done in the context of USA USSR nuclear non-proliferation treaty compliance checking.

Modern Terminology and (Simplified) Framework









- If system depends on secrecy of algorithm and there is no key involved – pure steganography
 - Not desirable. Kerkhoff's principle.
- Secret Key based steganography
- Public/Private Key pair based steganography

Active and Passive Warden Steganography



- Wendy can be passive:
 - Examines all messages between Alice and Bob.
 - Does not change any message
 - For Alice and Bob to communicate, Stego-object should be indistinguishable from cover-object.
- Wendy can be active:
 - Deliberately modifies messages by a little to thwart any hidden communication.
 - Steganography against active warden is difficult.
 - Robust media watermarks provide a potential way for steganography in presence of active warden.

Steganalysis



- Steganalysis refers to the art and science of discrimination between stego-objects and cover-objects.
- Steganalysis needs to be done without any knowledge of secret key used for embedding and maybe even the embedding algorithm.
- However, message does not have to be gleaned. Just its presence detected.

Cover Media



- Many options in modern communication system:
 - Text
 - Slack space
 - Alternative Data Streams
 - TCP/IP headers
 - Etc.
- Perhaps most attractive are multimedia objects -
 - Images
 - Audio
 - Video
- We focus on Images as cover media. Though most ideas apply to video and audio as well.

Steganography, Data Hiding and Watermarking



- Steganography is a special case of data hiding.
 - Data hiding in general need not be steganography. Example – Media Bridge.
- It is not the same as watermarking.
 - Watermarking has a malicious adversary who may try to remove, invalidate, forge watermark.
- In Steganography, main goal is to escape detection from Wendy.

Information Theoretic Framework



• Cachin [3] defines a Steganographic algorithm to be \mathcal{E} secure if the relative entropy between the cover object and the stego object pdf's is at most \mathcal{E} :

$$D(P_C||P_S) = \int P_C \cdot \log \frac{P_C}{P_S} \le \epsilon$$

- Perfectly secure if $\varepsilon = 0$
- Example of a perfectly secure techniques known but not practical.

Problems with Cachin Definition



Problems:

- In practice, leads to assumption that cover and stego image is a sequence of independent, identically distributed random variables
- Works well with random bit streams, but real life cover objects have a rich statistical structure
- There are examples for which D(X||Y)=0 but other related statistics are non-zero and might enable detection by steganalysis
- There are some alternative definitions but they have their own set of problems.

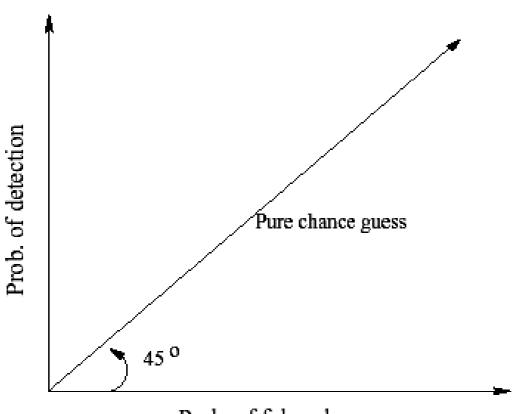
Another Way to Look at Security



- Chandramouli and Memon (2002)
- False Alarm Prob. P_{FA} = P(detect message | no message)
- Detection Prob. P_{Det} = P(detect message | message)
- If P_{FA}= P_{Det} then the detector makes purely random guess
- Therefore:
 - We call a steganographic algorithm γ secure (0< γ <1) if
 | P_{FA}- P_{Det} | ≤ γ
 - If $\gamma = 0$ then the algorithm is perfectly secure w.r.t. the detector.

Detector ROC Plane





Prob. of false alarm

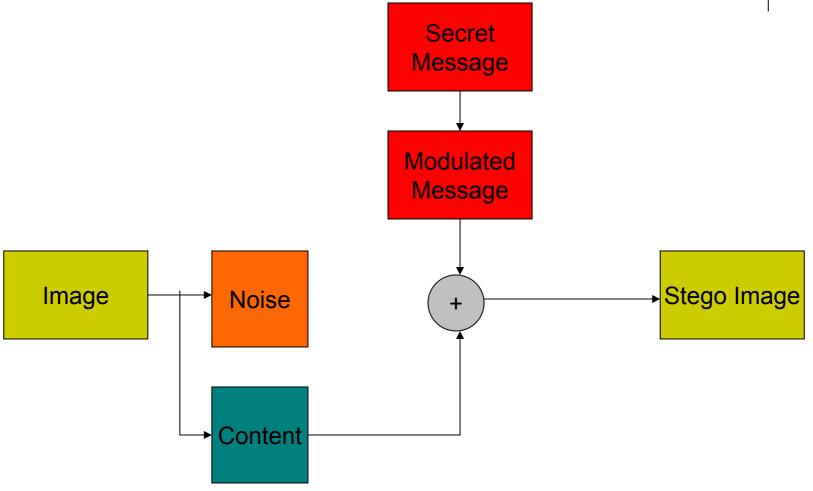
Steganographic Capacity



- By steganographic capacity we mean the number of bits that can be embedded given a level of security.
- This is different from data hiding or watermarking capacity.
- Specific capacity measures can be computed, given detector, and steganographic algorithm (Chandramouli and Memon, 2002)

Steganography in Practice



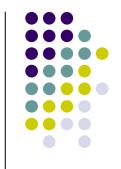


Steganalysis in Practice



- Techniques designed for a specific steganography algorithm
 - Good detection accuracy for the specific technique
 - Useless for a new technique
- Universal Steganalysis techniques
 - Less accurate in detection
 - Usable on new embedding techniques

A Note on Message Lengths



- Steganalysis techniques have been proposed which estimate the message length
- BUT:
 - An attack is called successful if it could detect the presence of a message.
 - So we mostly ignore message length estimating components.

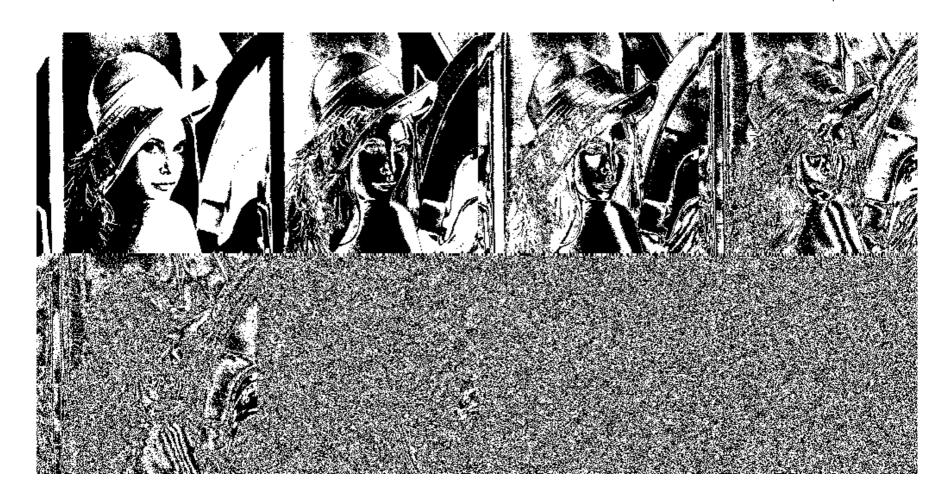
Simple LSB Embedding in Raw Images



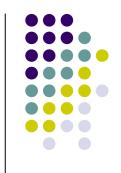
- LSB embedding
 - Least significant bit plane is changed. Assumes passive warden.
- Examples: Encyptic[9], Stegotif[10], Hide[11]
- Different approaches
 - Change LSB of pixels in a random walk
 - Change LSB of subsets of pixels (i.e. around edges)
 - Increment/decrement the pixel value instead of flipping the LSB

LSB Embedding





Steganalysis of LSB Embedding



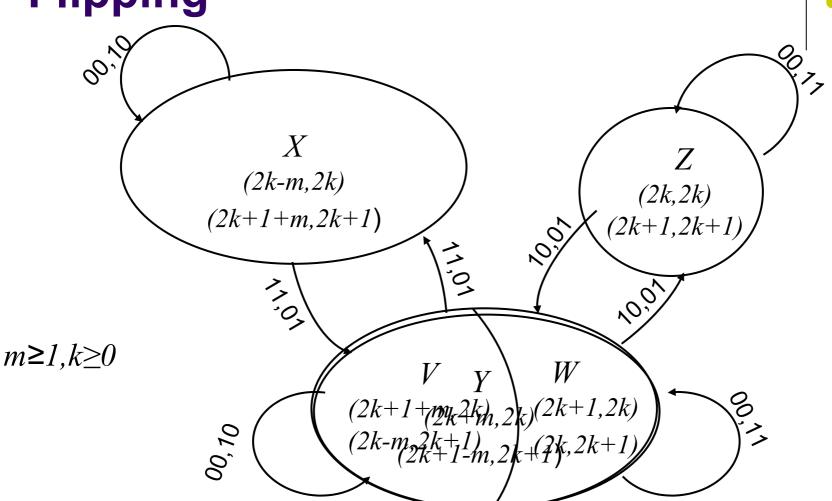
- PoV steganalysis Westfeld and Pfitzmann [12].
 - Exploits fact that odd and even pairs from "closed set" under LSB flipping.
 - Accurately detects when message length is comparable to size of bit plane.
- RS-Steganalysis Fridrich et. al. [14]
 - Very effective. Even detects around 2 to 4% of randomly flipped bits.

LSB steganalysis with Primary Sets



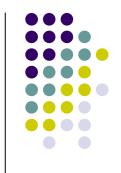
- Proposed by Dumitrescu, Wu, Memon [13]
 - Based on statistics of sets defined on neighboring pixel pairs.
 - Some of these sets have equal expected cardinalities, if the pixel pairs are drawn from a continuous-tone image.
 - Random LSB flipping causes transitions between the sets with given probabilities, and alters the statistical relations between their cardinalities.
 - Analysis leads to a quadratic equation to estimate the embedded message length with high precision.

State Transition Diagram for LSB Flipping



X, V, W, and Z, which are called **primary sets**

Transition Probabilities



 If the message bits of LSB steganography are randomly scattered in the image, then

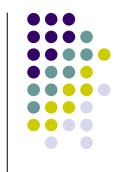
i)
$$\rho(00) = \left(1 - \frac{p}{2}\right)^2$$
,

ii)
$$\rho(01) = \rho(10) = \frac{p}{2} \left(1 - \frac{p}{2}\right)$$
,

iii)
$$\rho(11) = \left(\frac{p}{2}\right)^2$$
.

 Let X, Y, V, W and Z denotes sets in original image and X', Y'. W' and Z' denote the same in stego image.

Message Length in Terms of Cardinalities of Primary Sets



 Cardinalities of primary sets in stego image can be computed in terms of the original

$$|X'| = |X| \left(1 - \frac{p}{2}\right) + |V| \frac{p}{2}$$

$$|V'| = |V| \left(1 - \frac{p}{2}\right) + |X| \frac{p}{2}$$

$$|W'| = |W| \left(1 - p + \frac{p^2}{2}\right) + |Z| p \left(1 - \frac{p}{2}\right)$$

Assuming

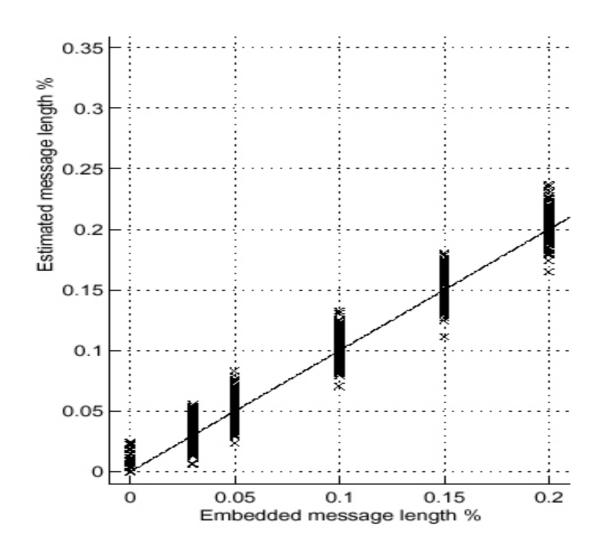
$$E\{|X|\} = E\{|Y|\}$$
 and some algebra, we get:

$$0.5\gamma \cdot p^2 + (2|X'| - |P|)p + |Y'| - |X'| = 0$$

• Where $\gamma = |W' \cup Z'| = |W \cup Z|$.

Simulation Results





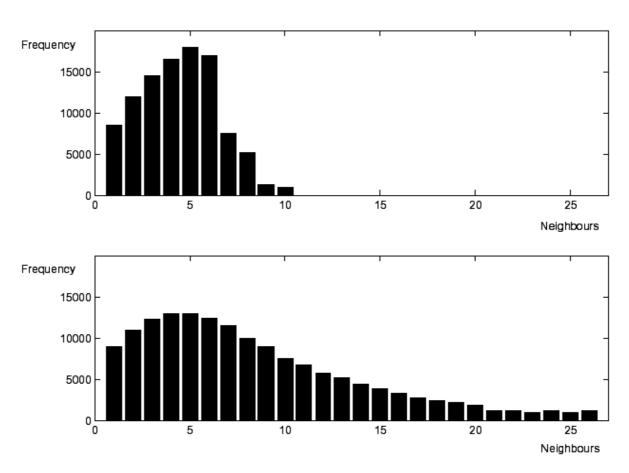
Hide



- Instead of simply flipping the LSB, it increments or decrements the pixel value
- Westfeld [16] shows that this operation could create 26 neighboring colors for each pixel
- On natural images there are 4 to 5 neighboring colors on average

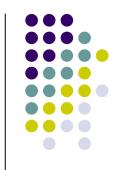
Hide





Neighborhood histogram of a cover image (top) and stego image with 40 KB message embedded (bottom)[16]

LSB Embedding in Palette Images



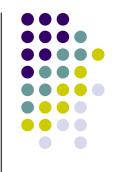
- Embedding is done by changing the LSB of color index in the palette
- Examples: EzStego[17], Gifshuffle[18], Hide and Seek[19]
- Such alteration result in annoying artifacts
- Johnson and Jajodia[20] look at anomalies caused by such embedding

EzStego



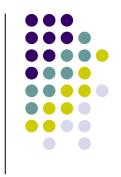
- EzStego [17] tries to minimize distortion by sorting the color palette before embedding
- Fridrich [6] shows that the color pairs after sorting have considerable structure
- After embedding this structure is disturbed thus the entropy of the color pairs are increased
- The entropy would be maximal when the maximum message length is embedded

Embedding in JPEG Images



- Embedding is done by altering the DCT coefficient in transform domain
- Examples: Jsteg[21], F5[22], Outguess[23]
- Many different techniques for altering the DCT coefficients

F5



- F5 uses hash based embedding to minimize changes made for a given message length
- The modifications done, alter the histogram of DCT coefficients
- Fridrich [6] shows that given the original histogram, one is able to estimate the message length accurately
- The original histogram is estimated by cropping the jpeg image by 4 columns and then recompressing it
- The histogram of the recompressed image estimated the original histogram

F5 plot



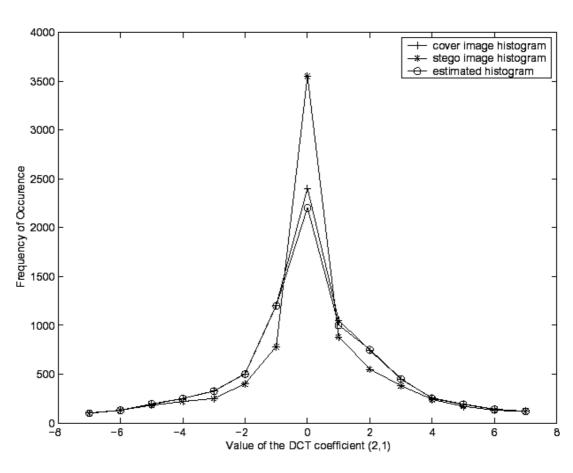
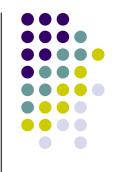


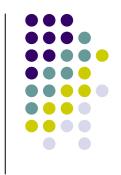
Fig. 5. The effect of F5 embedding on the histogram of the DCT coefficient (2,1).[6]

Outguess



- Embeds messages by changing the LSB of DCT coefficients on a random walk
- Only half of the coefficients are used at first
- The remaining coefficients are adjusted so that the histogram of DCT coefficient would remain unchanged
- Since the Histogram is not altered the steganalysis technique proposed for F5 will be useless

Outguess



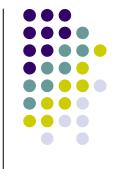
- Fridrich [6] proposes the "blockiness" attack
- Noise is introduced in DCT coefficients after embedding
- Spatial discontinuities along 8x8 jpeg blocks is increases
- Embedding a second time does not introduce as much noise, since there are cancellations
- Increase or lack of increase indicates if the image is clean or stego

Universal Steganalysis Techniques



- Techniques which are independent of the embedding technique
- One approach identify certain image features that reflect hidden message presence.
- Two problems
 - Calculate features which are sensitive to the embedding process
 - Finding strong classification algorithms which are able to classify the images using the calculated features

What makes a Feature "good"



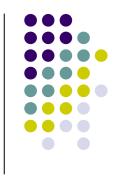
- A good feature should be:
 - Accurate
 - Detect stego images with high accuracy and low error
 - Consistent
 - The accuracy results should be consistent for a set of large images, i.e. features should be independent of image type or texture
 - Monotonic
 - Features should be monotonic in their relationship with respect to the message size

IQM



- Avcibas et al.[24,26] use Image Quality
 Metrics as a set of features
- IQM's are objective measures
- From a set of 26 IQM measures a subset with most discriminative power was chosen
- ANOVA is used to select those metrics that respond best to image distortions due to embedding

Choice of IQMs

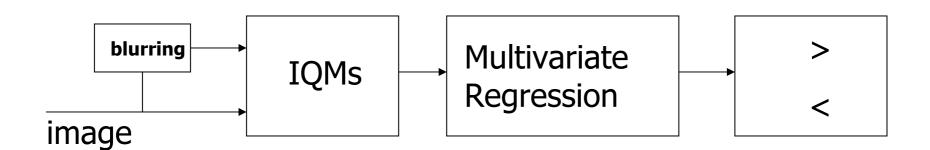


- Different metrics respond differently to different distortions.
 For example:
 - mean square error responds more to additive noise
 - spectral phase or mean square HVS-weighted error are more sensitive to blur
 - gradient measure reacts more to distortions concentrated around edges and textures.
- Steganalyzer must work with a variety of steganography algorithms
- Several quality metrics needed to probe all aspects of an image impacted by the embedding

IQM

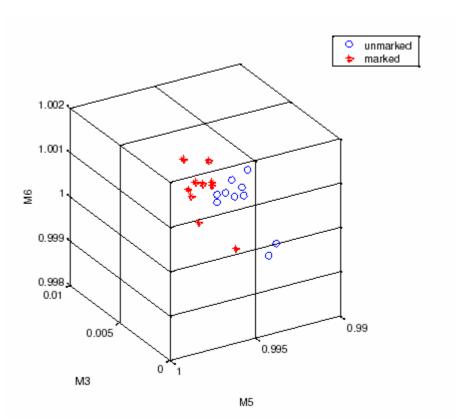


- The images are first blurred
- The IQM are then calculated from the difference of the original and blurred image



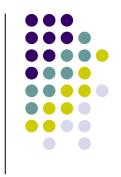
IQM





Scatter plot of 3 image quality measures showing separation of marked and unmarked images.

Farid



- Farid et. al.[27] argues that most steganalysis attacks look at only first order statistics
- But new techniques try to keep the first order statistics intact
- So Farid builds a model for natural images and then classifies images which deviate from this model as stego images

Farid



- Quadratic mirror filters are used to decompose the image, after which higher order statistics are collected
- These include mean, variance, kurtosis, skewness
- Another set of features used are error obtained from an optimal linear predictor of coefficient magnitudes of each sub band

Classifiers



- Different types of classifier used by different authors.
 - Avcibas et. al. use a MMSE linear predictor
 - Farid et. al. use Fisher linear discriminates as well as a SVM classifier
- SVM classifiers seem to do much better in classification
- All the authors show good results in their experiments, but direct comparison is hard since the setups are very much different.

So What Can Alice (Bob) Do?



- Limit message length so that detector does not trigger
- Use model based embedding.
 - Stochastic Modulation (Fridrich 02)
 - This conference Phil Sallee
- Adaptive embedding
 - Embed in locations where it is hard to detect.
- Active embedding
 - Add noise after embedding to mask presence.
 - Outguess

Adaptive Embedding

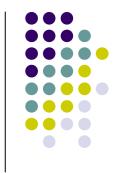
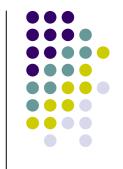


Image	Bits flipped	RS reported value
Baboon	4500	0.0207
Clock	5020	0.0249
Hats	1600	0.0216
Lena	5020	0.0204
New York	8080	0.0205
Peppers	200	0.0240
SAR	12760	0.0206
Teapot	2000	0.0246
Tolicon	22720	0.0209
Watch	200	0.0256

- LSB embedding in a location only if its 8neighborhood variance is high.
- Embedding locations still secret key dependent.
- Number of bits that can be embedded is significantly small.
- Would work against most steganalyzers?

Another Twist – Data Masking



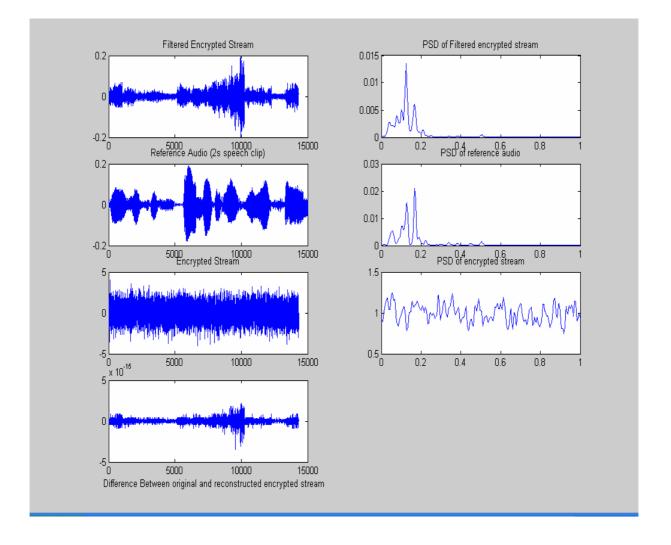
- Current model assumes Wendy also examines messages perceptually.
- However, in a large scale surveillance application this may not be feasible
- Wendy must solely rely on statistical tests and then only use perceptual tests on small set of "suspects".
- So as long it statistically seems to be an image it can have poor perceptual quality!!
- R. Radhakrishnan, K. Shanmugasundaram and N. Memon (2002).

Example Data Masked Stream

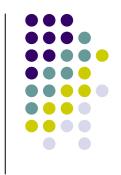


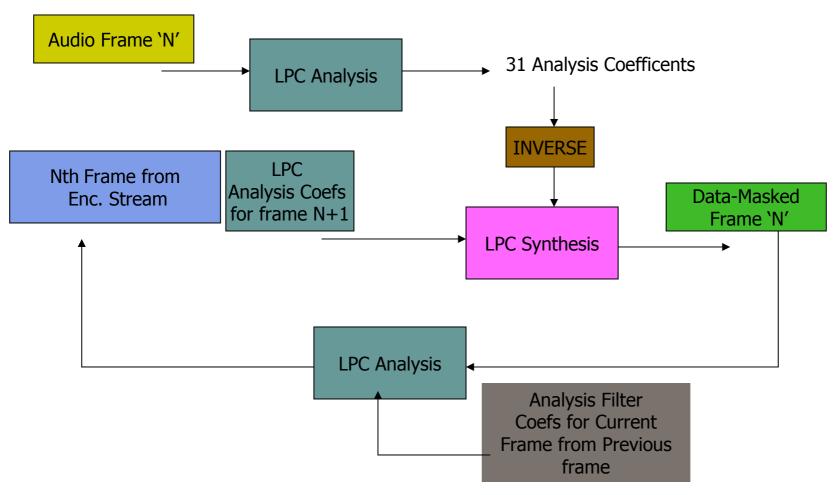






Data Masking by LPC Analysis/Synthesis

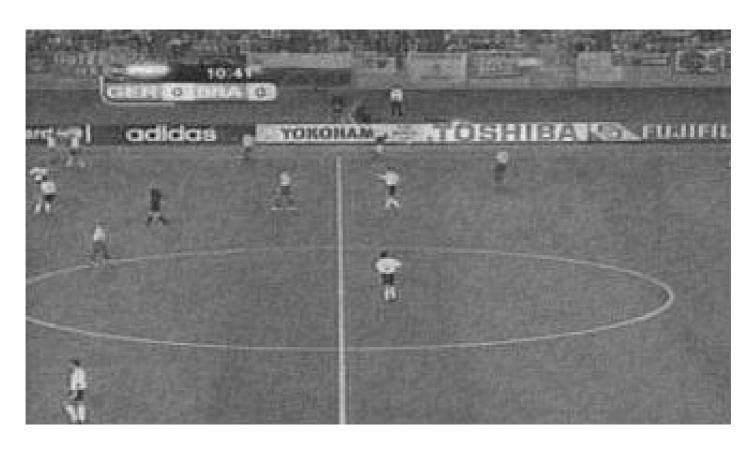




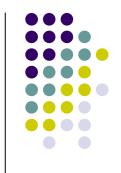
Data Masking with Images



 Take secret message and treat it as Huffman coded prediction errors.



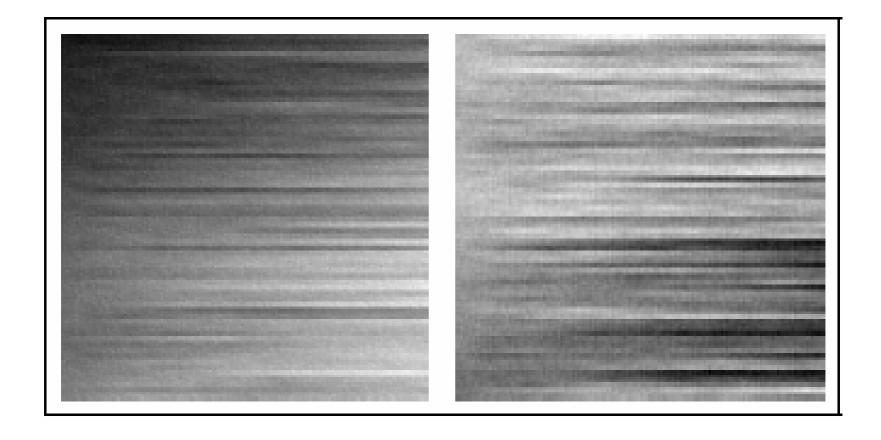
Stretching more ...



- In fact it need not look like an image or audio or video at all.
- Idea
 - Take encrypted secret message random stream.
 - Decompress it using some codec like JPEG, JPEG200 etc.
 - Compress the resulting stream losslessly and transmit.

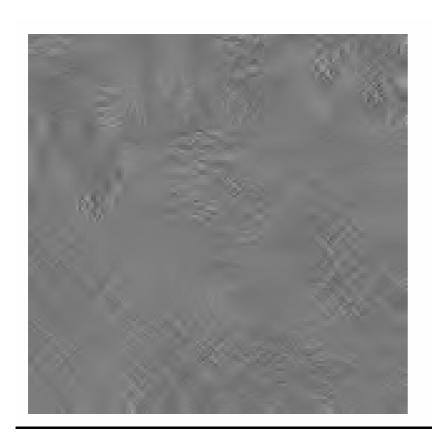
Images From DCT-based Image Decoders

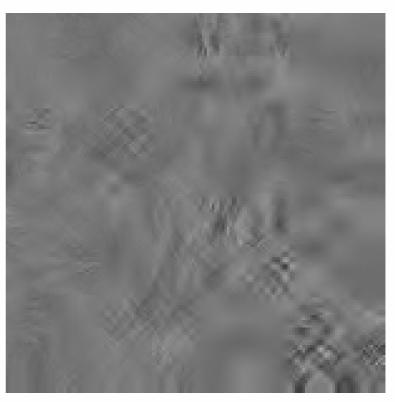




From Wavelet-based Image Decoders

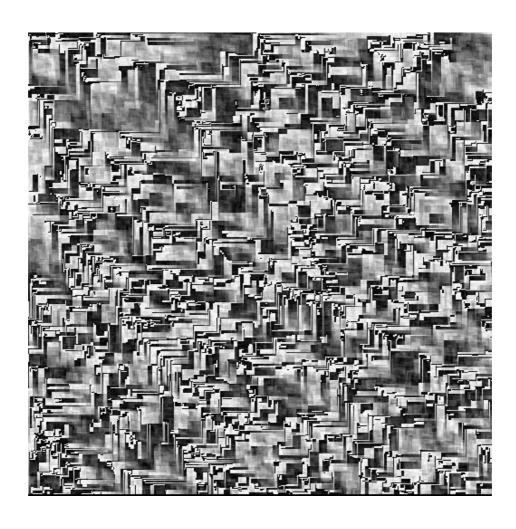




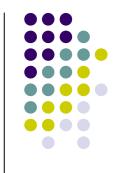


From JPEG-LS Lossless Image Decoder





Ton Kalker's Algorithm



- Fix positions in the image that will carry massage.
- Examine pictures until you find one in which bits in these positions are exactly what you want to embed.
- Clearly secure, but very low capacity. Much more than 10 bits or so will be impractical.
 - Capacity can be increased by blocking strategy.
 - But security becomes unclear.

Conclusion



- Steganography and steganalysis are still at an early stage of research
- In general, the covert channel detection problem is known to be undecidable!!
- Although in principle secure schemes exist, practical ones with reasonable capacity are not known.
- Notion of security and capacity for steganography needs to be investigated
- Steganography and corresponding steganalysis using image models needs to be further investigated

Other thoughts



- Unlike cryptography, Steganography allows you to choose the cover object.
- How do you choose good cover object for a given stego message
- What kind of images are good for using as cover objects?