



Proof of Concept Watermarking Implementations for the Protection and Tracking of Restricted Data

***Tod Reinhart
Air Force Research Laboratories
AFRL/IFTA
2241 Avionics Circle
WPAFB, Dayton, OH 45433-7334
Tod.Reinhart@wpafb.af.mil***

***Robert J. Moore, Roberta L. Gotfried
Raytheon
Sensors and Electronics Systems Center
2000 E. Imperial Highway
RE/R01/A521
El Segundo, CA 90245-3571
{rjmoore,rlgotfried}@raytheon.com***

***This work was funded by AFRL Information Directorate under contract
number F33615-97-D-1153 D0005***

Prepared for STC 2002 for presentation on 1 May 2002



Presentation Roadmap

Raytheon

- ***Introductory Quiz***
- ***Brief Introduction to Digital Watermarking Technology and its Applications***
- ***A Data Protection Problem We Addressed with Digital Watermarking Technology***
- ***Proof of Concept Digital Watermarking Algorithms, Implementations, and Results***
- ***Summary of Our Preliminary Work***
- ***Our Ongoing Research***



Introductory Quiz: Question

Raytheon

- Which of the following images (if any) is watermarked?

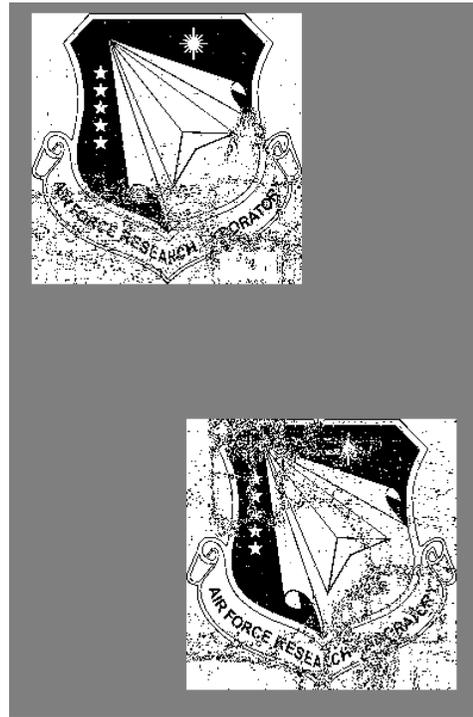




Introductory Quiz: Answer

Raytheon

- The image on the left in the previous chart is a watermarked version of the one on the right
- The watermark below was extracted from the green color channel (spatial domain) of the first image without using the original image and requires knowledge of a secret key





Digital Watermarking Technology (1) **Raytheon**

- *Algorithms have been proposed for digital media such as*
 - Audio
 - Still Images
 - Video
 - Graphics
 - Text
 - Multimedia
- *Some digital media problems which can be addressed*
 - Copyright Protection
 - Copy Control
 - Fingerprinting
 - Authentication
 - Tamper Detection
 - Broadcast Monitoring



Digital Watermarking Technology (2)

Raytheon

- *Digital Watermarking is a very broad, technically challenging, and rapidly advancing field with myriad possible applications*
- *We start by restricting our attention to the digital watermarking of still images which:*
 - (1) May either be perceptually invisible or visible
 - (2) May either be detectable by anyone or require knowledge of a secret key
 - (3) May or may not require the original image for detection, and
 - (4) Fragile, robust, or semi-robust to image modifications
- *Furthermore, there are many ways to implement image watermarks. For example they may be:*
 - Directly embedded into the image (spatial domain), or
 - Embedded after first performing some discrete image transform such as Fourier (DFT), Cosine (DCT), or Wavelet (DWT) (frequency domain)

Thus, it is very important to define exactly what problem you want to address before proceeding with watermark design and implementation



Problem 1: Ownership Protection of Digital Content

Raytheon

- **Consider a system where the owners of digital images wish to mark them in order to enforce policy restrictions on their use and distribution**
 - **If a restricted image is found in the possession of an unauthorized party, provide a method capable of proving that the image is owned by the content owner or custodian**
 - **Do not require the original image or original watermarked image**
 - ◆ **Even if modifications that do not degrade the usefulness of the image have been made to the watermarked image**



Watermark System Requirements

Raytheon

- *After analysis, we determined the appropriate type of watermark required, summarized in this case by*
 - **Invisible**
 - It is unacceptable to degrade image quality
 - **Private**
 - Watermark detection will require knowledge of a secret key (even though the underlying algorithms are assumed known to all)
 - **Blind Extraction**
 - It is impossible or impractical to require reference to the original image (see next chart for algorithm)
 - **Robust**
 - The watermark must still be detectable even in the face of specific image attacks or modifications
 - ◆ The 14 specific attacks enumerated below in this case



Blind Extraction Algorithm Summary

● Use standard statistical hypothesis testing techniques to decide whether a watermark has been detected or not

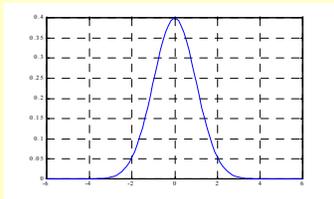
- H0 : A watermark W is **not** present (null hypothesis)
- H1: A watermark W is present

$$s(W, W_e) = \frac{W \bullet W_e}{\sqrt{W_e \bullet W_e}}$$

● Choose a very small probability of a false positive

- The probability that a watermark is declared present when in fact a watermark is not present
 - Tradeoff: decreasing probability of a false positive increases probability of a false negative
- If $|s(W, W_e)| = T$, then the probability of a false positive is equal to $1 - P_n(Z = T)$ where P_n is cumulative standard normal distribution, and W (W_e) the original (extracted) watermark.
- Referring to the table of values below:

$T=6$ is seen to be a good threshold



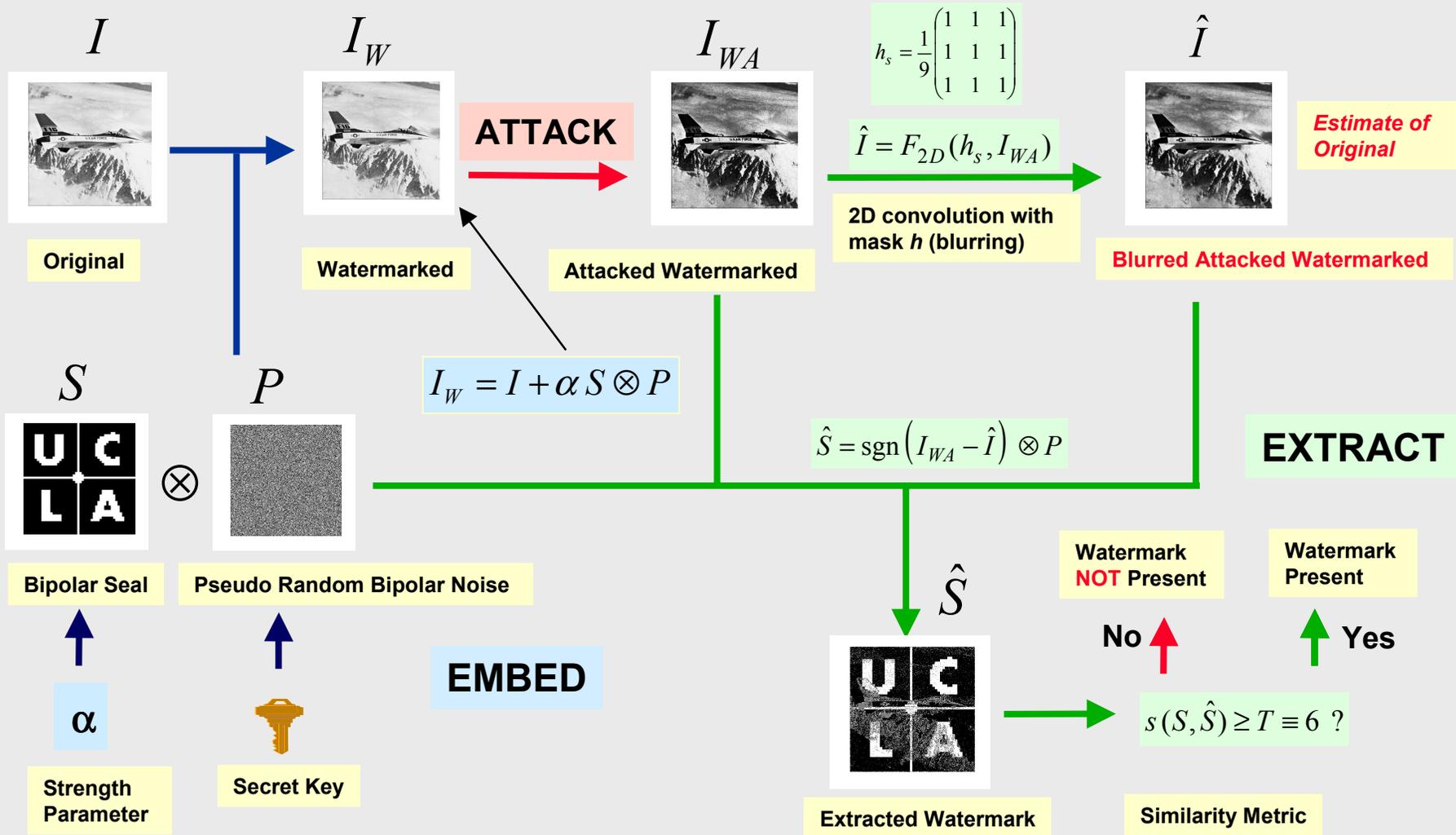
T	$1 - P_n(T)$
1.0	3.173E-01
2.0	4.550E-02
3.0	2.670E-03
6.0	1.973E-09

$$W \bullet W_e \sim N(0, W_e \bullet W_e)$$

Reference: I. Cox et. al. , *Secure Spread Spectrum Watermarking for Multimedia*, IEEE Trans. Image Processing 6,12, 1673-1687 (1997)



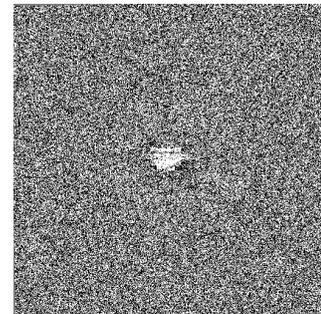
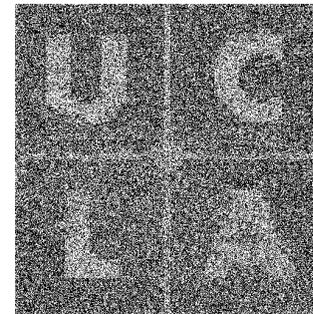
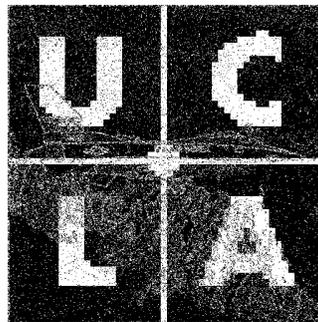
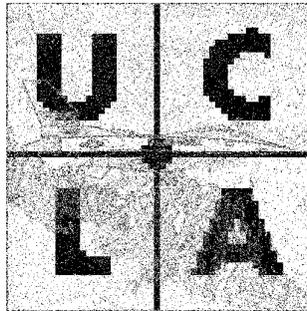
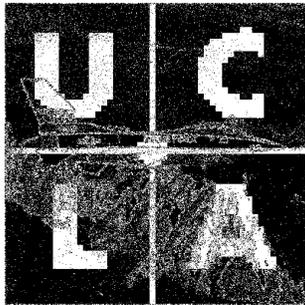
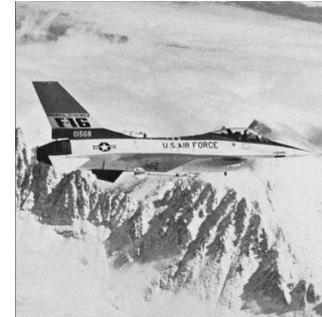
Watermark Embedding and Blind Detection (Spatial, Image Independent)





Blind Extraction of Spatial Watermark (selected cases)

Raytheon



No Attack
 $c=0.64$ $s=368.63$

Laplacian
 $c= -0.69$ $s=-389.85$

Unsharp
 $c= 0.70$ $s=394.62$

JPEG (qf=50)
 $c= 0.15$ $s=92.77$

Rotate ccw 1°
 $c=5.2E-3$ $s=0.81$



Spatial Watermark Blind Extraction Metrics (all cases)

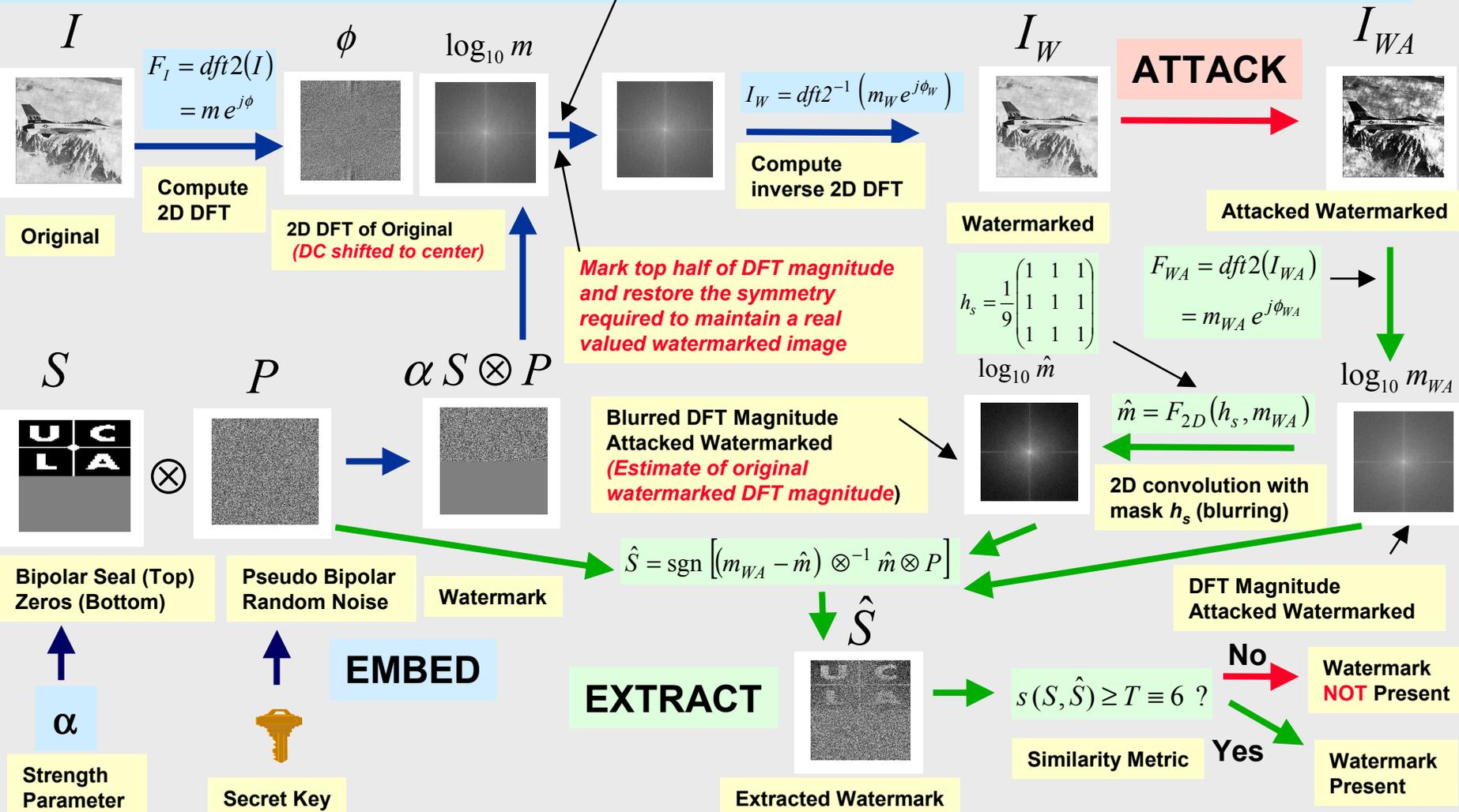
Raytheon

Watermark Implementation Type	Image Embedding Method	Attack	Correlation Coefficient	Similarity Metric	
Spatial	Independent	00: None	0.69	392.05	
Spatial	Independent	01: Gaussian Low Pass Filter	0.64	368.63	
Spatial	Independent	02: Laplacian High Pass Filter	0.69	389.85	
Spatial	Independent	03: Unsharp Contrast Enhancement	0.70	394.62	
Spatial	Independent	04: Sobel Edge Detection	8.47E-4	0.55	FAIL
Spatial	Independent	05: JPEG lossy compression (qf=50)	0.15	92.97	
Spatial	Independent	06: Increase brightness +14 gray levels	0.69	392.30	
Spatial	Independent	07: Add N(0,36) Gaussian Noise	0.64	370.33	
Spatial	Independent	08. Resize from 512x512 to 530x520 with nearest pixel interpolation. Then resize from 530x520 back to the original size using nearest pixel interpolation	0.69	392.05	
Spatial	Independent	09. Rotate image 1 degree counterclockwise and crop to original image size	5.20E-03	0.82	FAIL
Spatial	Independent	10. Adjust contrast. Relative range 0.3 to 0.7 adjusted to range 0.0 to 1.0	0.12	74.71	
Spatial	Independent	11. Median Filtering	7.41E-02	48.00	
Spatial	Independent	12. Histogram Equalization	0.68	385.43	
Spatial	Independent	13. Crop (set pixels in outer perimeter 10 pixels thick to zero).	0.66	374.85	
Spatial	Independent	14. Circular shift of image to right by 10 pixels	-7.85E-4	-1.04	FAIL



Watermark Embedding and Blind Detection (Frequency, Image Dependent)

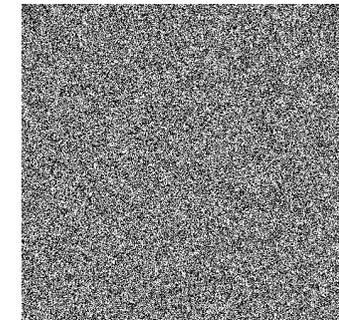
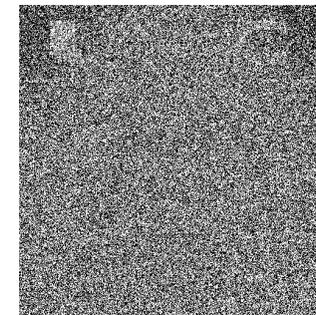
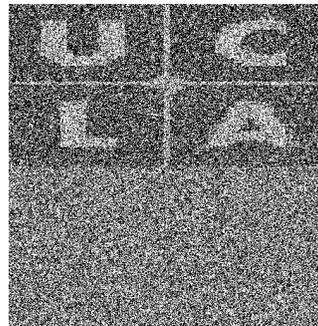
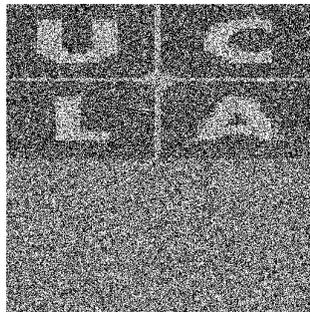
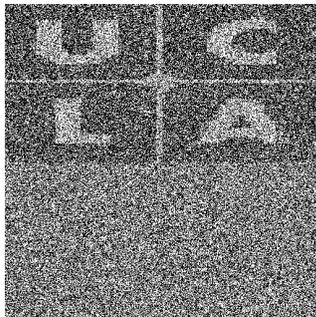
$$m_W = m(1 + \alpha S \otimes P); \phi_W = \phi; F_W(N_1 - k_1, N_2 - k_2) = \bar{F}_W(k_1, k_2) \text{ for } k_1 = 0 \text{ to } \frac{1}{2}N_1, k_2 = 0 \text{ to } N_2 - 1 \text{ excluding } F_W(0,0) \text{ and } F_W(\frac{1}{2}N_1, \frac{1}{2}N_2)$$





Frequency Watermark Blind Extraction (selected cases)

Raytheon



No Attack
 $c=0.18$ $s=46.92$

Laplacian
 $c=0.18$ $s=100.67$

Unsharp
 $c=0.18$ $s=101.00$

JPEG (qf=50)
 $c=3.95E-02$ $s=22.90$

Rotate ccw 1°
 $c=4.63E-04$ $s=0.34$



Frequency Watermark Blind Extraction Metrics (all cases)

Raytheon

Watermark Implementation Type	Image Embedding Method	Attack	Correlation Coefficient	Similarity Metric
DFT Magnitude	Dependent	00: None	0.18	71.01
DFT Magnitude	Dependent	01: Gaussian Low Pass Filter	0.18	70.61
DFT Magnitude	Dependent	02: Laplacian High Pass Filter	0.18	71.04
DFT Magnitude	Dependent	03: Unsharp Contrast Enhancement	0.18	71.28
DFT Magnitude	Dependent	04: Sobel Edge Detection	0.15	56.38
DFT Magnitude	Dependent	05: JPEG lossy compression (qf=50)	3.95E-02	16.16
DFT Magnitude	Dependent	06: Increase brightness +14 gray levels	0.18	71.01
DFT Magnitude	Dependent	07: Add N(0,36) Gaussian Noise	7.98E-02	31.09
DFT Magnitude	Dependent	08. Resize from 512x512 to 530x520 with nearest pixel interpolation. Then resize from 530x520 back to the original size using nearest pixel interpolation	0.18	71.01
DFT Magnitude	Dependent	09. Rotate image 1 degree counterclockwise and crop to original image size	8.64E-04	0.24
DFT Magnitude	Dependent	10. Adjust contrast. Relative range 0.3 to 0.7 adjusted to range 0.0 to 1.0	3.75E-02	15.10
DFT Magnitude	Dependent	11. Median Filtering	4.40E-02	17.91
DFT Magnitude	Dependent	12. Histogram Equalization	0.13	51.56
DFT Magnitude	Dependent	13. Crop (set pixels in outer perimeter 10 pixels thick to zero).	0.16	62.20
DFT Magnitude	Dependent	14. Circular shift of image to right by 10 pixels	0.18	71.01

FAIL



Analysis of Results (1)

Raytheon

- **A watermark is declared to be present if it passes the threshold test:** $|s(W, \hat{W})| > T$ where W [\hat{W}] is the original [extracted] watermark seal and $T = 6$
- **Spatial, Image Independent embedding *FAILS* the following attacks**
 - **04 Sobel**
 - **09 Rotation counterclockwise by 1°**
 - **14 Circular Shift to right by 20 pixels**
- **Frequency (DFT magnitude), Image Dependent embedding *FAILS* the following attacks:**
 - **09 Rotation counterclockwise by 1°**



Analysis of Results (2)

Raytheon

● **Observations**

- In general, no one implementation watermark type will be robust against every specified attack (both fail the rotation attack in this example)
- Particular watermark implementation types generally will have different degrees of robustness against the same attacks
- Several design and implementation iterations will generally be required to provide the required level of robustness
 - In general, some requirements may have to be relaxed



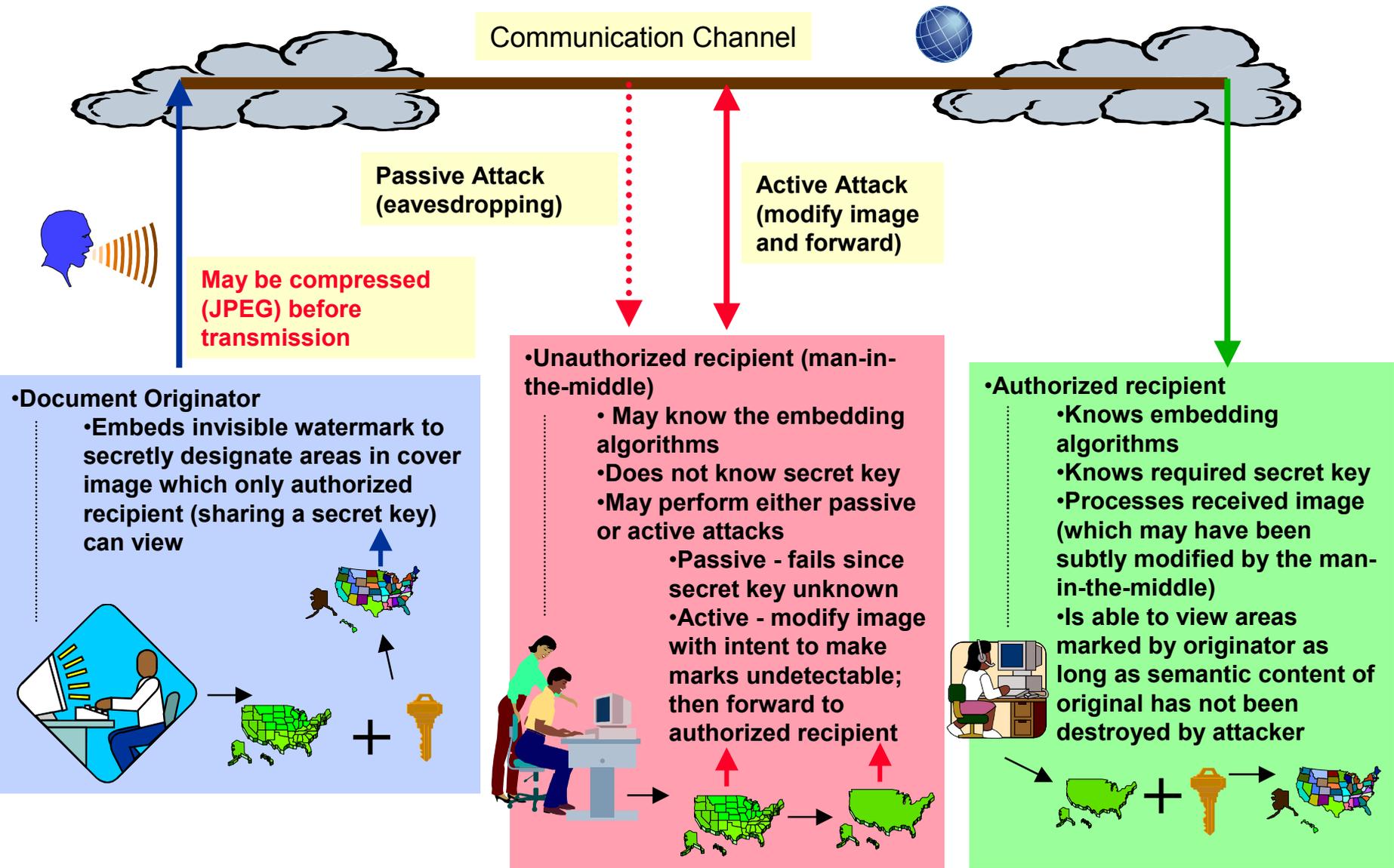
Summary

Raytheon

- *The previous charts have illustrated the general principles involved in a typical application of digital watermarking technology*
- *The algorithms shown were preliminary and implementations were “proof of concept” only*
- *This work formed the basis for solving more challenging problems which we are currently addressing*



Problem 2: Operational Picture



•Document Originator

- Embeds invisible watermark to secretly designate areas in cover image which only authorized recipient (sharing a secret key) can view

•Unauthorized recipient (man-in-the-middle)

- May know the embedding algorithms
- Does not know secret key
- May perform either passive or active attacks
 - Passive - fails since secret key unknown
 - Active - modify image with intent to make marks undetectable; then forward to authorized recipient

•Authorized recipient

- Knows embedding algorithms
- Knows required secret key
- Processes received image (which may have been subtly modified by the man-in-the-middle)
- Is able to view areas marked by originator as long as semantic content of original has not been destroyed by attacker



Problem 2: Requirements (1)

Raytheon

- *R1: Embedded watermark location(s) shall be invisible to the human visual system*
- *R2: Embedded watermark location(s) shall **not** be detectable using standard statistical tests*
- *R3: Image region(s) marked with watermark elements shall only be revealed with knowledge of the secret key used to create them*
- *R4: Knowledge of the watermark embedding algorithms shall **not** reveal the location of the marked region(s)*
- *R5: The original image shall not be required for an authorized recipient to view the marked regions*



Problem 2: Requirements (2)

Raytheon

- ***R6: The embedded watermarks shall be robust against the following attacks limited in extent such that they do not degrade the semantics of the image contents***
 - **R6a: Lossy compression**
 - **R6b: Common spatial domain image processing operations**
 - **R6c: Minor geometric distortions**
 - **Rotations, translation, scaling, shearing**
- ***R7: The system resources needed to embed and extract the watermark(s) shall not exceed the maximum limits established for this purpose***



Comparison of Our Solutions to Problems 1 and 2

Raytheon

Problem 1

- *Entire image was used to store only 1-bit of information:*
 - Is the watermark present? (yes or no)
- *The watermark seal (UCLA logo) was known a priori*
- *The blind extraction process was tolerant to moderate levels of image degradation*
- *Multiple types of embedding and extraction algorithms were successfully used*
 - Both spatial and frequency domain
 - Spatial
 - Frequency (DFT magnitude)

Problem 2

- *Many bytes of information must be encoded in the form of a watermark “seal”*
 - Information about the image
- *The watermark “seal” (the embedded information) was **NOT** known a priori*
- *The blind extraction process was very sensitive to image degradation*
- *More sophisticated embedding and extraction algorithms were required*
 - Some form of hybrid spatial-frequency technique will probably be required
 - For example, blocked DCT



Digital Watermarks: Selected References

Raytheon

- [01] I. Cox, J. Kilian, T. Leighton, T. Shamoon, *Secure Spread Spectrum Watermarking for Multimedia*, IEEE Trans. On Image Processing, Vol. 6, No. 12, 1673-1687, (1997).
- [02] G. Langelaar, I. Setyawan, R. Legendijk, *Watermarking Digital Image and Video Data (A State-of-the-Art Overview)*, IEEE Signal Processing Magazine, pp.20-45, IEEE 1053-5888/00 (September 2000).
- [03] W.Kim, J.C. Lee, W.D. Lee, *A Watermarking Scheme for both Spatial and Frequency Domain to Extract the Seal Image without the Original Image*, Fifth International Symposium on Signal Processing and its Applications, ISSPA '99, Brisbane, Australia (22-25 August 1999).
- [04] J.J.K. O Ruanaidh, W.J. Dowling, F.M. Boland, *Phase Watermarking of Digital Images*, IEEE 0-7803-3258-X, (1999)
- [05] M. Wu, B. Liu, *Attacks on Digital Watermarks*, IEEE 0-7803-5700-0/99 (1999)
- [06] J. Fridrich, *Secure Encryption and Hiding of Intelligence Data*, Air Force Research Laboratory Information Directorate, Rome, New York, *AFRL-IF-RS-TR-1998-192* (September 1998)
- [07] H. Maître, *IMAGE WATERMARKING, Why is watermarking a hard problem*, Korea-France Workshop on Multimedia, Seoul, Korea (July 6-9, 1998)
- [08] S. Pereira, T. Pun, *Robust Template Matching for Affine Resistant Image Watermarks*, IEEE Transactions on Image Processing, Vol. 9, No. 6 (June 2000)
- [09] A. Piva, M. Barni, F. Bartolini, V. Cappellini, *DCT-based Watermarking Recovering without Resorting to the Uncorrupted Original Image*, Dipartimento di Ingegneria Elettronica, Università di Firenze via di S. Marta, 3, 50139 Firenze, Italy
- [09] A. Piva, M. Barni, F. Bartolini, *Copyright Protection of Digital Images by Means of Frequency Domain Watermarking*, Dipartimento di Ingegneria Elettronica, Università di Firenze via di S. Marta, 3, 50139 Firenze, Italy
- [10] F. Alturki, R. Nersereau, *An Oblivious Robust Digital Watermark Technique for Still Images using DCT Phase Modulation*, IEEE 0-7803-6293-4/00 (2000)
- [11] F. Hartung, M. Kutter, *Multimedia Watermarking Techniques*, Proceedings of the IEEE, Vol, 87, No.7 (July 1999)