Further Progress in Watermark Evaluation Testbed (WET)

Hyung Cook Kim, Eugene T. Lin, Oriol Guitart and Edward J. Delp

Video and Image Processing Laboratory (VIPER)
School of Electrical and Computer Engineering
Purdue University
West Lafayette, Indiana USA

ABSTRACT

While Digital Watermarking has received much attention in recent years, it is still a relatively young technology. There are few accepted tools/metrics that can be used to evaluate the suitability of a watermarking technique for a specific application. This lack of a universally adopted set of metrics/methods has motivated us to develop a web-based digital watermark evaluation system called the *Watermark Evaluation Testbed* or *WET*. There have been more improvements over the first version of *WET*. We implemented batch mode with a queue that allows for user submitted jobs. In addition to StirMark 3.1 as an attack module, we added attack modules based on StirMark 4.0. For a new image fidelity measure, we evaluate conditional entropy as an image fidelity measure for different watermarking algorithms and different attacks. Also, we show the results of curve fitting the Receiver Operating Characteristic (ROC) analysis data using the Parzen window density estimation. The curve fits the data closely while having only two parameters to estimate.

Keywords: digital watermarking, watermark evaluation, watermark benchmarking, image database, Taguchi loss function, mutual information

1. INTRODUCTION

Digital Watermarking is the practice of hiding a message in an image, audio, video or other digital media element. Since the late 1990s, there has been an explosion in the number of digital watermarking algorithms published [1–6]. The sudden increase is mostly due to the increase in concern over copyright protection of content [7]. Because digital devices store content in digital form, there is no degradation in quality of data after copy is made [8,9]. Currently, cryptographic techniques are the most popular method used to prevent piracy [10]. The problem with cryptographic techniques is that after the content is decrypted to be consumed, the content is no longer protected against piracy. Because digital watermarking can embed information related to the content in the digital media element and can be designed to survive many changes, it is seen as a technique to complement cryptography in preventing piracy. Applications of watermarking [6, 11–13] include broadcast monitoring, owner identification, proof of ownership, transaction tracking, authentication, copy control, and device control.

An important and often neglected issue in the design of digital watermarking methods is proper evaluation and benchmarking [14–16]. This lack of a proper evaluation in designing watermarking methods causes confusion among researchers and hinders the adoption of digital watermarking in various applications. With a well-defined benchmark, researchers and watermarking software developers would just need to provide a table of results, which would give a good and reliable summary of the proposed scheme performances for end users. To address this issue, a few still image digital watermarking benchmarks have been proposed. These include StirMark [16–18], Certimark [19], Checkmark [20], Optimark [21, 22] and the *Watermark Evaluation Testbed (WET)* [23]. In [23], different watermarking algorithms were compared using the Taguchi loss function [24, 25].

We are still developing the *WET* system and in this paper we describe the progress in *WET*. Also, we describe a new image quality comparison measure and a ROC analysis interpolation technique that are based on image registration methods using mutual information [26–29]. This paper is organized as follows: In section 2, we describe a brief overview of the architecture, implementation and new features of *WET*. In section 3, we

This work was supported by the Air Force Research Laboratory, Information Directorate, Rome Research Site, under research grant number F30602-02-2-0199. Address all correspondence to E. J. Delp, ace@ecn.purdue.edu.

evaluate watermarks using mutual information as an image fidelity measure and show the results of curve fitting the ROC analysis data using the Parzen window density estimation. The conclusion and future work is given in section 4.

2. WATERMARK EVALUATION TESTBED UPDATE

In the following, we give you a brief overview and new features of *WET* [23]. For detailed description of *WET* refer to [23]*. The testbed consists of three major components: the Front End, the Algorithm Modules and the Image Database. Each component will be described below. Currently, *WET* runs on a 2.4 GHz Pentium 4 computer using the Fedora Core 1.0 operating system†.

2.1. Front End

The Front End is the end users' main interface into *WET*. The Front End provides a web interface whereby a user can select various tasks to be performed. Two versions of *WET*, the *initial version* and *advanced version*, are available. The *initial version* is intended for either first time users to familiarize themselves with the system or for users who want to get a feel for how watermarking works. The *advanced version* provides an extensive image database. All metrics reported in the *initial version* are also reported in the *advanced version.* The *advanced version* is available in two modes: interactive and batch modes. The interaction between all components that implements *advanced version* is illustrated in Figure 1. The interactive mode allows the user to manually step through the process illustrated in Figure 2. The interactive mode is similar to that of the *initial version* with more freedom in selecting images for evaluation.

Batch mode is a part of *WET* that allows users to make a more exhaustive evaluation of the chosen watermarking techniques. Unlike the interactive mode, batch mode does not run the processes at the time that the user submits the job. Instead, batch mode queues the user requests into a job and sends e-mail to the user when the job has completed. The e-mail contains a URL where the user can retrieve the evaluation results. Thereby, jobs can be lengthy and process much more data than an interactive display, where a user expects results to be available quickly. This job queue was not included in the previous version of *WET*. To submit a job using the batch mode, first the user chooses the images to watermark. Second, the user selects one or more watermarking techniques to be evaluated and specifies their watermark embedding parameters. Third, the user (optionally) chooses the attacks to be performed on the watermarked images. In the fourth step, the user selects the watermark detection parameters. Finally, the user submits the e-mail address to complete the procedure and submit the job.

2.2. Algorithm Modules

It is desirable to develop tools that users can use standalone in their own test environments, allowing them to validate their tests locally before submitting them to a watermark benchmark site. To achieve this, the GNU Image Processing Program (GIMP) [30] was selected for use in *WET*. GIMP is designed to be augmented with plug-ins and extensions. Additionally, it provides full scripting support in various languages (Scheme, Perl, Java, Python, etc.). The combination of extensibility and scripting support make GIMP a powerful environment for *WET*.

We have been implementing several plug-ins for GIMP. Our plug-ins can be used in two different modes: Interactive Mode and Non-Interactive Mode. The former has a user interface where the user can choose the input parameters and the output parameters are displayed after using the plug-in. The Non-Interactive Mode performs the same function as the Interactive Mode but allows the plug-in to be called from another GIMP plug-in or scripts. The Non-Interactive Mode is used in *WET*.

*A copy of this paper is located at http://www.datahiding.org/about.html.

†The system is located at http://www.datahiding.org. To obtain access to use *WET*, contact wetbug@ecn.purdue.edu.

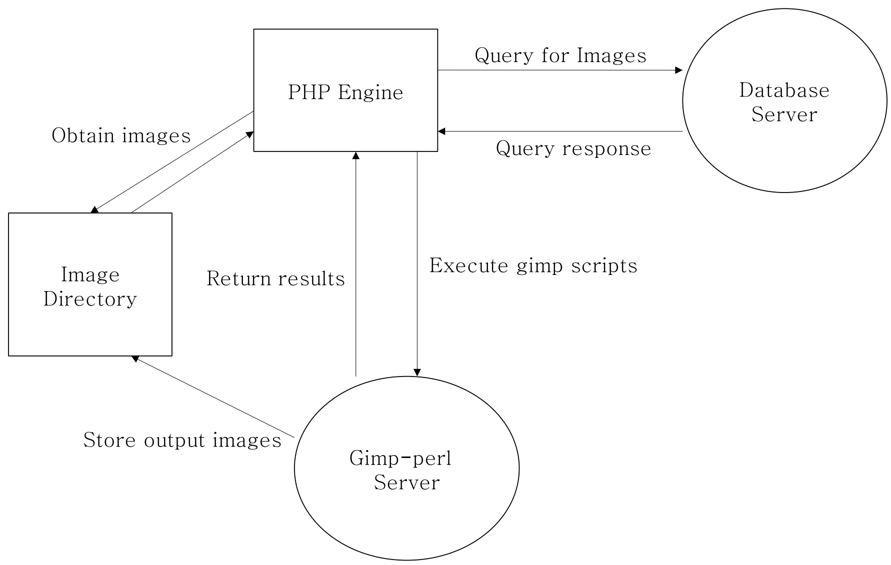


Figure 1. Interaction of major components in advanced version

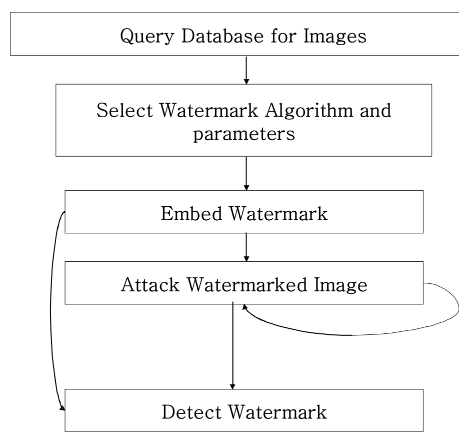


Figure 2. Watermarking steps for the advanced version

2.2.1. New Watermarking Algorithms

To extend watermarking algorithms to color images, we use the reversible color transform (RCT) [31] developed for JPEG2000. The Red, Green and Blue components of an image are transformed by RCT and we embed the watermark into the luminance component. We chose RCT among different color transforms because it preserves the image when a watermark is not embedded. We watermarked the luminance component because a watermark should be placed in the perceptually most significant components of an image [3].

Three algorithms new blind watermarking algorithms are implemented as GIMP plug-ins. These are: Additive Spread Spectrum Watermarking (ASSW) [3], Multiplicative Spread Spectrum Watermarking (MSSW) [32] and Improved Spread Spectrum Watermarking (ISSW) [33]. Implementation of each algorithm is described in [23]

2.2.2. StirMark 4.0

In addition to StirMark [17,18,34] 3.1 attacks we implemented the StirMark 4.0 attacks as separate GIMP plug-ins. The attacks are classified into geometric transforms, signal processing operations and special transforms. Geometric transforms include affine transform, rescale, rotation and small random distortions. Signal processing attacks operations include adding dither noise, adding uniformly distributed noise, filtering by convolution, median cut and self-similarities attack. Special Transforms include flipping(horizontal, diagonal, and vertical), cropping, remove lines, and special rotations (90, 180, and 270 degrees).

2.3. Image Database

The image database maintains the attributes of all images available in the test-bed. We currently have approximately 1301 images which are copyright free. The images are from a variety of cameras and sensors including scanned photographs, x-ray images, ultrasound images, astrometrical images, line drawings, digital cameras, maps, and computer generated images. Each image in the database is stored with its attributes including chrominance, resolution, height, width, and category.

3. WATERMARK EVALUATION: A STATISTICAL FRAMEWORK

Our goal is to develop a "theory" for watermark evaluation and use *WET* to deploy the developed techniques. In [23], we developed a new watermark evaluation framework. To improve on this technique, we introduce two new techniques for watermark evaluation. First, we introduce conditional entropy as an image fidelity measure to compare different watermarking techniques. In [23], we used Mean Square Error (MSE) as an image fidelity measure. One of the problems with MSE as an image fidelity measure [6] is that it does not compensate against attacks such as amplitude changes and contrast changes such as gamma correction. In medical imaging community, mutual information has been successfully used in image registration as a similarity measure between images obtained from different modalities [26, 28, 29]. This motivated us to use conditional entropy which is related to mutual information as an image fidelity measure to evaluate watermarks and attacks. Second, we extrapolate the Receiver Operating Characteristics (ROC) using Parzen window estimation used in [26, 27]. ROC analysis is a valuable tool to compare different watermarks [35]. When there are no attacks, the value of the false negative rate is so small for specified false positive rates that we need to extrapolate the distribution of the detection values for watermarked images [6, 36]. In [23], we extrapolated the ROC data assuming that the watermark detection statistics for true positive and true negative are normally distributed and independent [6,21]. For better estimation is to use chernoff bounds [6, 36, 37]. Also using the spherical method to estimate the false positive probabilities has been proposed [6, 38].

3.1. Watermarking Model

We assume the following watermarking model

$$\begin{aligned} X &= S + W \\ Y &= Attack(X) \end{aligned}$$

where S is a random variable that models host images, W is a random variable that models the watermark, X is a random variable that models the watermarked images, and Y is a random variable that models the attacked watermarked images [5].

3.2. Conditional Entropy as an Image Fidelity Measure

If the watermarked image undergoes attacks such as amplitude scaling and gamma correction, mean square error does not compensate for such effects. In image registration, mutual information has been successfully used to measure the alignment of images obtained through different modalities [26, 28, 29]. This motivated us to use conditional entropy as a fidelity measure to evaluate still image watermarking algorithms and attacks.

3.2.1. Definitions and Background

The entropy $H(X)$ of a discrete random variable X is defined by [39]

$$H(X) = -\sum_x p(x) \log p(x)$$

where $p(x)$ is the probability mass function of the random variable X. The joint entropy $H(X, Y)$ of a pair of discrete random variables (X, Y) with a joint distribution $p(x, y)$ is defined as

$$H(X, Y) = -\sum_x \sum_y p(x, y) \log p(x, y).$$

The conditional entropy $H(X|Y)$ is defined as

$$H(X|Y) = -\sum_x \sum_y p(x, y) \log p(x|y).$$

The properties of H include

$$\begin{aligned} H(X) &\geq 0 \\ H(X|Y) &\leq H(X) \\ H(X, Y) &= H(Y) + H(X|Y). \end{aligned}$$

The mutual information $I(X; Y)$ between two random variables X and Y is defined as

$$I(X; Y) = \sum_x \sum_y p(x, y) \log \frac{p(x, y)}{p(x)p(y)}$$

Properties of $I(X; Y)$ are

$$\begin{aligned} I(X; Y) &= \sum_x \sum_y p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \\ &= \sum_x \sum_y p(x, y) \log \frac{p(x|y)}{p(x)} \\ &= H(X) - H(X|Y) \\ &= H(X) + H(Y) - H(X, Y). \end{aligned}$$

$I(S; Y)$ is a measure of the amount of information that Y contains about S [39]. $I(S; Y)$ is a "larger is better" measure because it is the reduction in the uncertainty of S by knowing the distribution of Y. To use $I(X; Y)$ in a Taguchi loss function, it has to be modified such that the function is a "smaller is better" function [23–25]. In this sense, we use the conditional entropy $H(S|Y)$ as a similarity measure. $H(S|Y)$ is the average bits needed to describe S given Y. The relation between $H(S|Y)$ and $I(S; Y)$ is given as follows:

$$\begin{aligned} H(S|Y) &= H(Y, S) - H(Y) \\ &= H(S) - I(Y; S). \end{aligned}$$

We estimate entropies using histograms [28]. Let $h(s, y)$ be the joint histogram of the pixel values. We approximate the probability as follows:

$$\begin{aligned} p(x, y) &\approx \frac{h(x, y)}{\sum_x \sum_y h(x, y)} \\ p(x) &= \sum_y p(x, y) \\ p(y) &= \sum_x p(x, y). \end{aligned}$$

3.3. Parzen Window Density Estimation

In this paper, we use the Parzen Window Density Estimation to estimate the underlying detection statistic densities to interpolate and extrapolate ROC. Parzen Window Density Estimation was used in image registration to estimate the underlying probability densities to evaluate mutual information between two images [26, 27].

For a random variable Z, we divide the samples z_j into 2 groups A and B and approximate the underlying probability density as

$$p(z) \approx \frac{1}{N_A} \sum_{z_j \in A} G_\sigma(z - z_j)$$

where

$$G_\sigma(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{z^2}{2\sigma^2}).$$

We choose the value of σ that maximizes the likelihood for sample data set B, that is

$$\begin{aligned} \sigma_{ML} &= \arg\max_\sigma \prod_{z_i \in B} p(z_i) \\ &= \arg\min_\sigma \frac{-1}{N_B} \sum_{z_i \in B} \log \frac{1}{N_A} \sum_{z_j \in A} G_\sigma(z_i - z_j). \end{aligned}$$

3.4. Experimental Results

3.5. Watermark Evaluation Parameters

The watermark evaluation parameters we use to evaluate watermarks are described in [23] and are similar to the parameters used in Optimark [21, 22]. Also, the watermark evaluation procedure is described in [23]. For fair benchmarking, one has to ensure that the methods under investigation are tested under comparable conditions [35]. We use the Taguchi loss function [23–25] to compare different watermark algorithms for fair evaluation. We define a new Taguchi loss function for the conditional entropy $H(S|Y)$ as follows:

$$\mathrm{H(S|Y)}' = (\frac{1}{N_I N_K} \sum_k (H(S|Y)_k)^2)^{\frac{1}{2}}.$$

where $H(S|Y)_k$ is the conditional entropy of the kth image. We chose PSNR′, H(S|Y)′ as a measure of fidelity and use BER′ and ROC analysis as a robustness measure for a watermarking technique. We estimate the watermark detection statistics for true positive and true negative using the Parzen Window density estimation defined in subsection 3.3. We need to only estimate σ for true positive and true negative.

3.6. Experiments

We selected ASSW, ISSW, and MSSW as our test algorithms. We set the lower specification limit for PSNR′ as 45dB. We chose the data payload to be 16 bits. This specification can be used as a specification for "fingerprinting" applications [40]. We tested N_K=2 key for each image. For the image test set **I**, we used our image database. It has $N_I = 1301$ images. We test $N_K \times N_I$=2602 keys. To simplify the evaluation procedure, we converted the color images to gray scale images.

3.6.1. Attacks

We use StirMark [16–18] 4.0 software and histogram equalization [41] for our attacks. We evaluate each watermarking algorithms for the following attacks:

- Gaussian filtering (blur): $\begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}/16$

Name	No Attack	Gaussian Filtering	Sharpening	JPEG	Amplitude Scaling	Hist. Equal.
ASSW	(45.0, 45.0)	(31.6, 33.5)	(21.9, 22.8)	(39.0, 39.8)	(12.4, 13.3)	(13.3, 16.1)
MSSW	(45.0, ∞)	(31.5, 33.4)	(22.5, 23.8)	(38.1, 38.9)	(12.4, 13.3)	(13.0, 16.3)
ISSW	(45.0, 45.0)	(31.6, 33.4)	(21.9, 22.8)	(39.0, 39.8)	(12.4, 13.3)	(13.3, 16.1)

Table 1. $(\text{PSNR}', \text{PSNR}_{average})$ pair for different attacks

Name	No Attack	Gaussian Filtering	Sharpening
ASSW	(1.9e-2, 2.0e-3)	(6.7e-2, 2.1e-2)	(9.3e-3, 3.6e-4)
MSSW	(5.2e-2, 1.3e-2)	(9.4e-2, 4.0e-2)	(3.4e-2, 4.5e-3)
ISSW	(6.6e-3, 2.6e-4)	(5.4e-2, 1.3e-2)	(9.3e-3, 4.1e-4)

Name	JPEG	Amplitude Scaling	Hist. Equal.
ASSW	(9.3e-2, 3.4e-2)	(2.2e-2, 2.4e-3)	(1.7e-3, 4.8e-5)
MSSW	(7.4e-2, 2.5e-2)	(5.2e-2, 1.3e-2)	(4.3e-2, 7.8e-3)
ISSW	(7.6e-2, 2.3e-2)	(8.6e-3, 4.1e-4)	(1.3e-2, 7.2e-4)

Table 2. $(BER', \text{BER}_{average})$ pair for different attacks

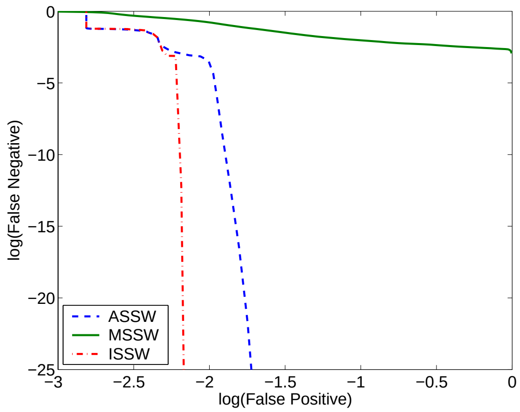
- Simple sharpening: $\begin{pmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{pmatrix}$
- JPEG compression with quality factor 70
- Amplitude scaling with scaling factor 9/16
- Histogram equalization [41]

3.6.2. Results

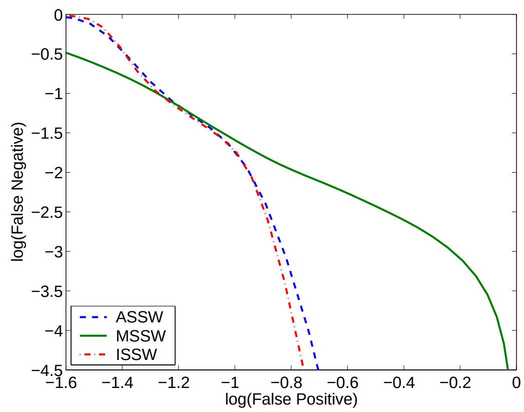
As mentioned in [23], Table 1 shows that we can compare different watermarking algorithms in terms of PSNR using the Taguchi loss function, even in cases where a watermark algorithm fails to embed a watermark as shown in the MSSW $\text{PSNR}_{average}$ result. PSNR results are nearly identical to those of [23]. Table 2 shows the results for bit error rates. We can see sharpening and histogram equalization reduces the bit error rate in some cases. Table 3 shows the new results for conditional entropy. For no attack, conditional entropy is smaller for MSSW compared to ASSW and ISSW. This is because MSSW embeds a watermark near edges.

Table 1 and Table 3 shows that there is a correlation between PSNR and conditional entropy except for amplitude scaling and histogram equalization attack. Compared to PSNR, conditional entropy gives a better score to amplitude scaling and histogram equalization. Conditional entropy also gives a better score to amplitude scaling compared to other attacks. Table 3 shows that MSSW has a better score than ASSW and ISSW in terms of conditional entropy except for JPEG attack. For JPEG attack, we speculate that MSSW has a worse score than other algorithms because the watermark is embedded in frequency components that is significant in the original image and JPEG is designed to preserve the frequency characteristics of typical images.

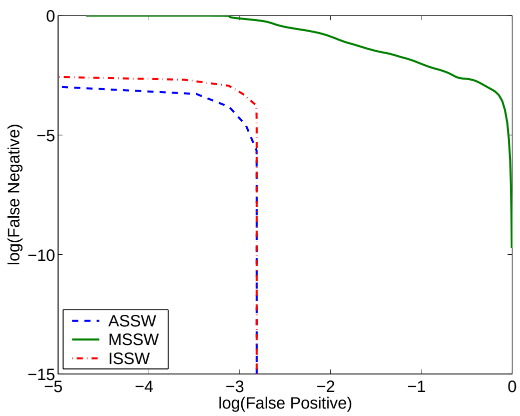
Figure 3 shows that ISSW has better detection error characteristic than ASSW for no attack. This is due to the host signal cancelation properties of ISSW. For Gaussian filtering attack, we can see there is a difference in the slope of the curves and MSSW curve crosses the other curves. For sharpening attack, there seems to be no difference for MSSW compared to no attack. For ASSW and ISSW, we can see that sharpening attack improved the detection error characteristics compared to no attack. This is different from the results obtained from [23]. For JPEG attack with quality factor 70, MSSW is better than other algorithms in terms of detection probability. For amplitude scaling, it seems that MSSW is affected less than the other two algorithms. For histogram attack, there seems to be no difference for MSSW while there are improvements for ASSW and ISSW compared to no attack.



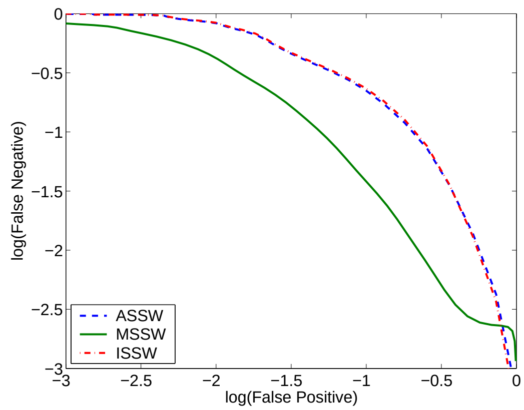
(1) ROC with no attack



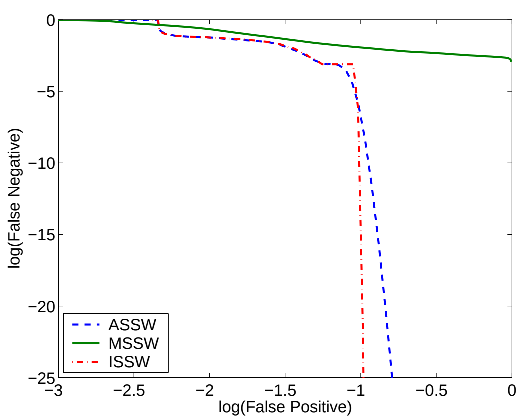
(2) ROC with Gaussian filtering attack



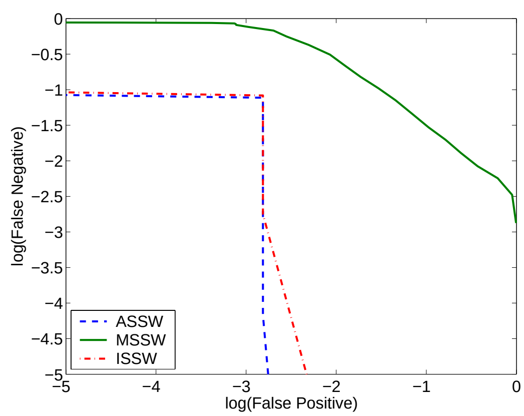
(3) ROC with simple sharpening attack



(4) ROC with JPEG attack with quality 70



(5) ROC with amplitude scaling attack with scale factor 9/16



(6) ROC with histogram equalization attack

Figure 3. ROC for various attacks

Name	No Attack	Gaussian Filtering	Sharpening	JPEG	Amplitude Scaling	Hist. Equal.
ASSW	(2.41, 2.36)	(3.34, 3.19)	(5.19, 5.05)	(3.09, 2.95)	(2.48, 2.43)	(2.45, 2.40)
MSSW	(1.76, 1.61)	(3.28, 3.11)	(4.75, 4.58)	(3.14, 3.00)	(2.04, 1.95)	(1.81, 1.67)
ISSW	(2.41, 2.36)	(3.34, 3.19)	(5.19, 5.05)	(3.09, 2.95)	(2.48, 2.43)	(2.45, 2.40)

Table 3. $(\mathrm{H(S|Y)}', \mathrm{H(S|Y)}_{average})$ pair for different attacks

4. CONCLUSION AND FUTURE WORK

We described our watermark evaluation architecture in this paper. New features compared to the previous version include more watermarking algorithms, StirMark 4.0 attack modules, and a job queue added to the batch mode. We also evaluated three watermark algorithms in terms of fidelity, bit error rate, false positive probability and false negative probability. We used the Taguchi loss function to define the fidelity and robustness requirements. We used conditional entropy to evaluate the image quality of the image. We used Parzen window density estimation for interpolation and extrapolation of ROC analysis.

For future work, we need to add features to *WET* to produce the results shown in the results. Furthermore, we need to improve the Front End of *WET* to support evaluation procedures similar to StirMark 4.0 [16]. In StirMark 4.0, the watermarking algorithm is evaluated using a evaluation profile that describes which images, fidelities and attacks are used in the evaluation procedure. An evaluation profile is necessary to evaluate watermarks in a fair manner. For ROC analysis, we need to further evaluate how accurate the interpolated and extrapolated detection error rates are using more samples [6]. We need to further investigate different image fidelity metrics, including minimum mean square error (MMSE) and perceptual metrics [42].

REFERENCES

1. R. B. Wolfgang, C. I. Podilchuk, and E. J. Delp, "Perceptual watermarks for digital images and video," *Proceedings of the IEEE*, Vol. 87, No. 7, pp. 1108–1126, Jul. 1999.
2. E. T. Lin, A. M. Eskicioglu, R. L. Lagendijk, and E. J. Delp, "Advances in digital video content protection," *Proceedings of the IEEE*, Vol. 93, No. 1, pp. 171–183, January 2005.
3. I. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Transactions on Image Processing*, Vol. 6, No. 12, pp. 1673–1687, 1997.
4. J. Fridrich and M. Goljan, "Comparing robustness of watermarking techniques," *Proceedings of the SPIE/IS&T Conference on Security and Watermarking of Multimedia Contents*, Vol. 3657, San Jose, CA, Jan. 1999, pp. 214–225.
5. C. I. Podilchuk and E. J. Delp, "Digital watermarking: Algorithms and applications," *IEEE Signal Processing Magazine*, Vol. 18, No. 4, pp. 33–46, Jul. 2001.
6. I. Cox, M. Miller, and J. Bloom, *Digital Watermarking*. Morgan Kaufmann, 2001.
7. E. J. Delp, "Is your document safe: An overview of document and print security," *NIP18: International Conference on Digital Printing Technologies*, San Diego, CA, Sep. 29–Oct. 4 2002, presented at.
8. A. M. Eskicioglu and E. J. Delp, "An overview of multimedia content protection in consumer electronics devices," *Signal Processing: Image Communication*, Vol. 16, pp. 681–699, 2001.
9. A. M. Eskicioglu, J. Town, and E. J. Delp, "Security of digital entertainment content from creation to consumption," *Signal Processing: Image Communication*, Vol. 18, pp. 237–262, 2003.
10. B. Schneier, *Applied Cryptography*, 2nd ed. New York: John Wiley and Sons, 1996.
11. I. J. Cox and M. L. Miller, "The first 50 years of electronic watermarking," *EURASIP Journal of Applied Signal Processing*, Vol. 2002, No. 2, pp. 126–132, 2002.
12. I. J. Cox, M. L. Miller, and J. A. Bloom, "Watermarking applications and their properties," *International Conference on Information Technology: Coding and Computing*, 2000, pp. 6–10.
13. F. Mintzer, G. Braudaway, and M. Yeung, "Effective and ineffective digital watermarks," *Proceedings of the IEEE International Conference on Image Processing*, Santa Barbara, CA, Oct. 1997, pp. 9–12.
14. B. Macq, J. Dittmann, and E. J. Delp, "Benchmarking of image watermarking algorithms for digital rights management," *Proceedings of the IEEE*, Vol. 92, No. 6, pp. 971–984, June 2004.

15. M. Kutter and F. A. P. Petitcolas, "A fair benchmark for image watermarking systems," *Proceedings of the SPIE/IS&T Conference on Security and Watermarking of Multimedia Contents*, P. W. Wong and E. J. Delp, Eds., Vol. 3657, San Jose, CA, Jan. 1999, pp. 226–239.
16. F. A. P. Petitcolas, M. Steinebach, F. Raynal, J. Dittman, C. Fontaine, and N. Fates, "A public automated web-based evaluation service for watermarking schemes: StirMark benchmark," *Proceedings of the SPIE/IS&T Conference on Security and Watermarking of Multimedia Contents*, P. W. Wong and E. J. Delp, Eds., Vol. 4314, San Jose, CA, Jan. 2001.
17. F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Attacks on copyright marking systems," *Information Hiding, Second International Workshop*, D. Aucsmith, Ed. Portland, OR: Springer-Verlag, Apr. 1998, pp. 219–239.
18. F. A. P. Petitcolas, "Watermarking schemes evaluation," *IEEE Signal Processing Magazine*, Vol. 17, No. 5, pp. 58–64, Sep. 2000.
19. J. C. Vorbruggen and F. Cayre, "The Certimark benchmark: architecture and future perspectives," *IEEE International Conference on Multimedia and Expo*, Vol. 2, Lausanne, Switzerland, Aug. 2002, pp. 485–488.
20. S. Pereira, S. Voloshynovskiy, M. Madueño, S. Marchand-Maillet, and T. Pun, "Second generation benchmarking and application oriented evaluation," *Information Hiding Workshop*, Pittsburgh, PA, Apr. 2001.
21. V. Solachidis, A. Tefas, N. Nikolaidis, S. Tsekeridou, A. Nikolaidis, and P. Pitas, "A benchmarking protocol for watermarking methods," *Proceedings of the IEEE International Conference on Image Processing*, Vol. 3, Thessaloniki, Greece, Oct. 2001, pp. 1023–1026.
22. "http://poseidon.csd.auth.gr/optimark/."
23. H. C. Kim, H. Ogunleye, O. Guitart, and E. J. Delp, "The Watermark Evaluation Testbed (WET)," *Proceedings of the SPIE/IS&T Conference on Security and Watermarking of Multimedia Contents*, ser. Proceedings of SPIE Electronic Imaging, San Jose, CA, Jan. 2004, pp. 236–247.
24. T. F. Rodriguez and D. A. Cushman, "Optimized selection of benchmark test parameters for image watermark algorithms based on Taguchi methods and corresponding influence on design decisions for real-world applications," *Proceedings of the SPIE/IS&T Conference on Security and Watermarking of Multimedia Contents*, P. W. Wong and E. J. Delp, Eds., Vol. 5020, San Jose, CA, Jan. 2003.
25. E. E. Lewis, *Introduction to Reliability Engineering.* John Wiley and Sons, Inc., 1996.
26. P. A. Viola, "Alignment by maximization of mutual information," Ph.D. dissertation, Massachusetts Inst. Technol., Boston, MA, 1995.
27. P. Viola and W. M. W. III, "Alignment by maximization of mutual information," *International Journal of Computer Vision*, Vol. 24, No. 2, pp. 137–154, 1997.
28. F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging*, Vol. 16, No. 2, pp. 187–198, Apr. 1997.
29. J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: A survey," *IEEE Transactions on Medical Imaging*, Vol. 22, No. 8, pp. 986–1004, Aug. 2003.
30. "http://www.gimp.org."
31. D. Taubman and M. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards, and Practice.* Kluwer Academic Publishers, 2002.
32. A. Piva, M. Barni, F. Bartolini, and V. Cappellini, "DCT-based watermark recovering without restoring to the uncorrupted original image," *Proceedings of the IEEE International Conference on Image Processing*, Vol. 3, 1997, pp. 520–523.
33. H. S. Malvar and D. A. F. Florencio, "Improved spread spectrum: A new modulation technique for robust watermarking," *IEEE Transactions on Signal Processing*, Vol. 51, No. 4, pp. 898–905, Apr. 2003.
34. "http://www.petitcolas.net/fabien/watermarking/stirmark/."
35. M. Kutter and F. A. P. Petitcolas, "Fair evaluation methods for image watermarking systems," *Journal of Electronic Imaging*, Vol. 9, No. 4, pp. 445–455, Oct. 2000.
36. J. R. Hernandez and F. Perez-Gonzalez, "Statistical analysis of watermarking schemes for copyright protection of images," *Proceedings of the IEEE*, Vol. 87, No. 7, pp. 1142–1166, 1999.

37. J. J. Hernandez, F. Perez-Gonzalez, J. M. Rodriguez, and G. Nieto, "Performance analysis of a 2-d multipulse amplitude modulation scheme for data hiding and watermarking still images," *IEEE Journal of Selected Areas in Communication*, Vol. 16, No. 4, pp. 510–524, 1998.
38. M. L. Miller and J. A. Bloom, "Computing the probability of false watermark detection," *Proceedings of the Third International Workshop on Information Hiding*, 1999, pp. 146–158.
39. T. M. Cover and J. A. Thomas, *Elements of Information Theory.* John Wiley & sons, 1991.
40. *Deliverables D21: Watermarking applications and requirements for benchmarking.* The Certimark Consortium, Oct. 2000, Available from http://www.certimark.org.
41. A. K. Jain, *Fundamentals of digital image processing.* Englewood Cliffs, NJ: Prentice-Hall, Inc., 1989.
42. A. B. Watson, "DCT quantization matrices visually optimized for individual images," *SPIE Proceedings, Human Vision, Visual Processing and Digital Display IV*, J. P. Allebach and B. E. Rogowitz, Eds., Vol. 1913, San Jose, CA, Feb. 1993, pp. 202–216.