Medical Image Watermarking: A Study on Image Degradation

B. Planitz and A. Maeder
e-Health Research Centre, ICT CSIRO
Brisbane, QLD 4000
Birgit.Planitz@csiro.au

Abstract

Digital watermarking has been proposed to increase medical image security, confidentiality and integrity. Medical image watermarking is a special subcategory of image watermarking in the sense that the images have special requirements. Particularly, watermarked medical images should not differ perceptually from their original counterparts, because the clinical reading of the images (e.g. for diagnosis) must not be affected. This paper presents a preliminary study on the degradation of medical images when embedded with different watermarks, using a variety of popular systems. Image quality is measured with a number of widely used metrics, which have been applied elsewhere in image processing. The general conclusion that arises from the results is that typical watermark embedding can cause numerical and perceptual errors in an image. The greater the robustness of a watermark, the greater the errors are likely to be. Consequently medical image watermarking remains an open area for research, and it appears that a selection of different watermarks for different medical image types is the most appropriate solution to the generic problem.

1 Introduction

Digital image watermarking is a particular subset of steganography, which is the art of hiding a covert message in a carrier message. Examples of messages are other images, or ASCII code such as text files, or numbers. Three elements are required to hide a message within a digital image. These are [5]:

Carrier message: the original, unmarked image I;

Payload message: the hidden message or watermark W; and

Steganography key: K, which is used to encrypt the watermark and/or for randomisation in the watermarking scheme.

The result is a stego image $\tilde{I}$. Mathematically, the embedding process can be described as a mapping $I \times W \times K \rightarrow \tilde{I}$ [7].

This paper considers the particular case of *medical image watermarking*. Watermarking has become an important issue in medical image security, confidentiality and integrity [1]. Medical image watermarks are used to authenticate (trace the origin of an image) and/or investigate the integrity (detect whether changes have been made) of medical images. One of the key problems with medical image watermarking, is that medical images have special requirements. A hard requirement is that the image may not undergo any degradation that will affect the reading of images. Generally, images are required to remain intact to achieve this, with no visible alteration to their original form [2]. This paper presents a preliminary investigation on medical image watermarking, by applying three popular watermarking systems to medical images, and examining the level of degradation that occurs. First, aspects of recent medical image watermarking systems are reviewed in Section 2. Section 3 then outlines three popular, general-purpose, watermarking systems, which are used in the study on medical image degradation. The quality metrics that are used to determine image degradation when applying a watermark, are presented in Section 4. These metrics are applied to investigate the three aforementioned systems, and their appropriateness for medical images, in Section 5. Finally, Section 6 summarises the paper and discusses future work.

2 Review

Medical image watermarking systems can be broken into three broad categories: robust, fragile, and semi-fragile. This section explains these terms and provides a brief review of existing systems in each category.

Robust watermarks are designed to resist attempts to remove or destroy the watermark [9]. They are used primarily for copyright protection and content tracking. Many traditional robust methods are spread-spectrum, whereby the watermark is spread over a wide range of image frequencies [5]. More recent work includes the creation of image-adaptive watermarks, where parameters change depending on local image characteristics [9].

A number of robust medical image watermarking systems have been developed. For example one system uses a spread spectrum technique to encode copyright and patient information in images [17]. Another embeds a watermark in a spiral fashion around the Region Of Interest (ROI) of an image [19]. Any image tampering that occurs will severely degrade the image quality. The Gabor transform has also been applied to hide information in medical images [6]. One observation that is generally applicable to robust systems is the greater the robustness of the watermark, the lower the image quality [7].

Fragile watermarks are used to determine whether an image has been tampered with or modified [9]. The watermark is destroyed if the image is manipulated in the slightest manner. Fragile watermarks are often capable of localisation, and are used to determine where modifications were made to an image. Traditional methods embed checksums or pseudo-random sequences in the Least Significant Bit (LSB) plane [5]. More recent work has employed increasingly sophisticated embedding techniques such as cryptographic hash functions [9].

Fragile invertible authentication schemes have been proposed for medical images, whereby a watermark can be removed from a stego image, and the exact original image results [2, 10]. Another medical image watermarking system embeds information in bit planes, which results in stego images with very low normalised root mean square errors (NRMSE), indicating that the watermark is practically invisible [4]. A watermark that is embedded in the high frequency regions of an image has also been proposed, which also resulted in low NRMSEs [4].

Semi-fragile watermarks combine the properties of both robust and fragile watermarks [9]. Like robust methods, they can tolerate some degree of change to the watermarked image (for example, quantisation noise from lossy compression). Like fragile methods, they are capable of localising regions of an image that are authentic and those that have been altered. Recent work in the area includes embedding a heavily quantised version of the original image in the image, embedding key-dependent random patterns in blocks of the image, wavelet embedding, and embedding multiple watermarks [9].

Recently, much emphasis has been placed on semi-fragile medical image watermarking. Jagadish *et al.* investigated interleaving hidden information in the Discrete Cosine Transform (DCT) and the Discrete Wavelet Transform (DWT) domains [4]. DCT and DWT domains are widely studied because they relate to the JPEG and JPEG2000 compression methods respectively. The NRMSEs of encoding in these domains are higher than in the spatial and DFT domains, but the image changes are still barely visible to the human eye. Another example of embedding watermarks using DCT coefficients is presented in [14]. Multiple watermark embedding has also been used by a number of researchers [3, 12, 13]. Multiple watermarking systems have the advantages that different watermarks can be applied for different purposes (e.g. copyright, authentication, data integrity) [3]. Also, image alterations can be detected by investigating the watermarks after the image has undergone degradation [12, 13].

A number of recent medical image watermarking systems have been proposed in this section. These were categorised into robust, fragile, and semi-fragile systems. The remainder of the paper will consider three systems of varying levels of robustness, in a preliminary study that investigates the degradation of medical images, when embedded with a watermark.

3 Watermarking Systems

This section briefly describes three widely used watermarking systems. These systems vary in robustness, and are applied to hide information in medical images later in the paper.

S-Tools is a popular package for image watermarking [16]. The system embeds one or more fragile watermarks in the LSBs of an image. Given a low insertion rate (i.e. the watermark is significantly smaller than the image), the watermark should be perceptually invisible in the stego image. Although widely used, LSB techniques such as this are sensitive to factors such as quantisation noise [5], which can easily destroy the watermark.

Hide4PGP is more robust than S-Tools [15]. This is due to the fact that information is generally embedded in the fourth LSB of an image, which increases the watermark's robustness against noise. However, this increase in robustness causes a decrease in image quality.

JPHide hides files in JPEG images [8], whereas the two aforementioned systems generally embed watermarks in BMP files. This system changes the statistics of JPEG coefficients, so that the embedded information can easily be retrieved when required. The system aims to provide high stego image quality, but maintains that low insertion rates ($< 5\%$) should be observed. Higher rates will cause the watermark to become visible in the stego image.

Figure 1 illustrates image embedding by applying the three aforementioned systems. A small text file (108 characters) is embedded in Figure 2(b). The difference image between the original and stego image is shown in Figures 1(a), (c), and (e), using S-Tools, Hide4PGP, and JPHide respectively. A significantly larger image file (40kb) is also embedded in the same image, resulting in the difference images shown in the right hand column of Figure 1. It can be seen that the more information is embedded in an image, the more visible the difference between the original and stego images. Image degradation increases when using Hide4PGP rather than S-Tools, and greatly increases when using JPHide. These results will be discussed further in Section 5. However, they were provided here as a means of comparing the robustness (and related image degradation caused) by the three systems discussed.

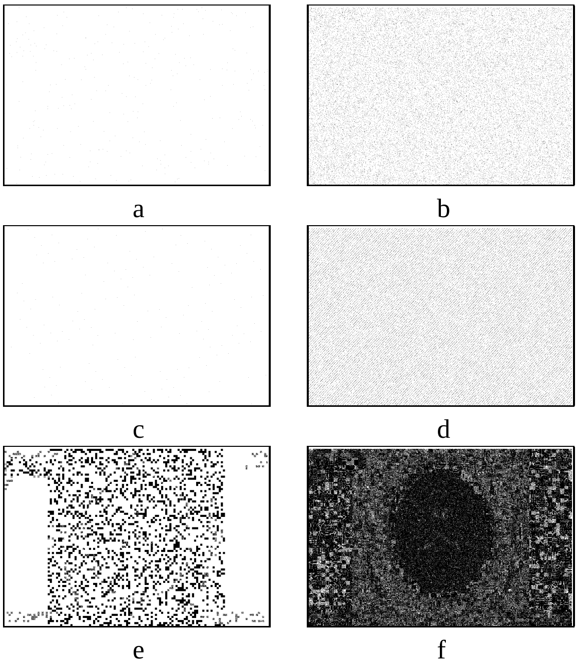


Figure 1. Difference images for implementing a 108 character text file using (a) S-Tools, (c) Hide4PGP and (e) JPHide, and implementing a 40kb image using (b) S-Tools, (d) Hide4PGP, and (f) JPHide.

4 Quality Metrics for Testing Image Degradation

As shown in Figure 1, watermarking causes image degradation. This section lists a number of metrics that quantify image degradation. These metrics have been applied widely in image quality assessment, including for medical imaging [11]. The metrics measure quality degradation using pixel-based comparisons, and the last one considers perceptual error in terms of the Human Vision System (HVS).

Entropy quantifies the amount of information that is present in an image. *Relative entropy*, or the Kullback-Leibler distance, normalises the entropy of an image $\tilde{I}$, with respect to a reference image I. Mathematically, it is expressed as:

$$m_e = \sum_k p_k log_2 \left(\frac{p_k}{q_k} \right), \qquad (1)$$

where p and q are the probability distributions of $\tilde{I}$ and I respectively, over all pixel intensities k. Given an image I, and a watermarked image $\tilde{I}$, m_e is expected to be low for similar images (0 if the images are equal) and high if the relative information differs significantly.

The Peak Signal-To-Noise Ratio (PSNR) is another commonly used image quality metric. *PSNR* is given by:

$$m_p = 10log_{10} \frac{B}{RMS}, \qquad (2)$$

where B is the largest possible value of the signal and RMS is the Root Mean Square difference between the two images. PSNR penalises the visibility of noise in an image [18]. Thus, two images that are exactly the same will produce an infinite PSNR value.

The Mean Square Error (MSE) compares two images on a pixel-by-pixel basis. Mathematically, *MSE* is expressed as:

$$m_s = \frac{1}{MN} \sum_i \sum_j \left(I_{ij} - \tilde{I}_{ij} \right)^2, \qquad (3)$$

where both images contain $M \times N$ pixels. This measure gives an indication of how much degradation was introduced at a pixel based level. The higher the MSE, the greater the level of degradation.

An alternative metric is the Mean Absolute Error (MAE). *MAE* is given by:

$$m_a = \frac{1}{MN} \sum_i \sum_j \left| I_{ij} - \tilde{I}_{ij} \right|. \qquad (4)$$

This equation quantifies the mean of all the absolute pixel-by-pixel differences in I and $\tilde{I}$.

Each of the four aforementioned metrics give an understanding of the actual differences in I and $\tilde{I}$, however these metrics do not focus on image differences in terms of the HVS. The *Watson* model has been designed to provide a measure that reflects image degradation as perceived by the HVS [20]. The basic aim of the model is to weight the DCT coefficients in an image block by its corresponding sensitivity threshold. The threshold is a compound function of sensitivity, luminance masking and contrast masking [18, 20]. The objective is to minimise the perceptual error between two images. Two images that are exactly the same will have an error of zero.

The metrics that have been presented here are used to measure image degradation in the following section, where medical images are watermarked using the tools discussed in Section 3. The metrics are used to compare system performance, and provide a general indicator of the appropriateness of each tool for embedding hidden data in medical images.

5 Results

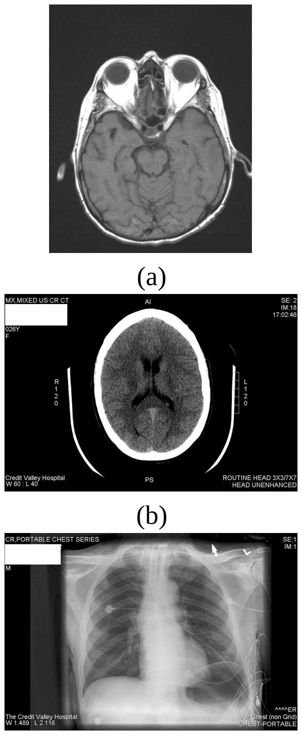
This section compares the image quality of three medical images that were embedded with a variety of watermarks. First, the three test images are presented. This is followed by a discussion on the watermarks that have been hidden in the images. The quantitative image quality results of each experiment are shown next. Finally, the appropriateness of each watermarking system for medical image data is discussed.

Three medical images were used in the watermarking experiment. The first image is from a Magnetic Resonance Imaging (MRI) modality. The second image is from a Computed Tomography (CT) modality. The third image was captured using a specialised CXR system. Figure 2 illustrates all three test data sets. Note that the images vary in size: 470×579 for the MRI image and 1022×689 for the CT and CXR images.

Four different watermarks were embedded in the medical images: text files with 108 and 1080 characters each, and JPEG images of size 4kb and 40kb. The text files were hidden in the images to test the image quality difference between embedding a text file, and another that is ten times larger, in an image. The same type of experiment was replicated with the image watermarks (based on the logo shown in Figure 3).

Table 1 presents the results of embedding the four watermarks in each of the medical images. Before analysing the results, some notes must be made about the outcomes. Firstly, JPHide was not able to produce results for the MRI data. Secondly, the JPHide program informs the user if a watermark is too large to embed in an image (in the sense that the watermark will cause significant visible distortions in the image). This was the case when embedding the 40kb logo in the CT and CXR images, and hence the results are shown in parentheses. The results are included for completeness, and to compare JPHide with the other two systems. Note also that in many cases, both MSE and MAE provide the same quantitative values. This is due to the binary nature of the images. Both sets of results are shown to emphasise the weakness of JPHide when embedding the logo within the medical images.

Some general observations can be made about the outcomes in Table 1. Firstly, image quality degrades as more data is embedded in an image. Secondly, increased watermark robustness is related to a decrease in image quality, as expected. Some specific results are now presented, by considering each quality metric separately.



(a)

(b)

(c)

Figure 2. Test data: (a) MRI head, (b) CT head and (c) CXR chest images. Images supplied by Queensland Health.

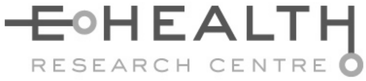


Figure 3. Image watermark: e-Health logo.

The *relative entropy* outcomes show that S-tools performs better than Hide4PGP for the MRI image. S-Tools and Hide4PGP entropies are approximately the same order of magnitude for the CT and CXR images. JPHide produced much higher values than other systems for the CT and CXR images, due to the ’heaviness’ of the embedding, which greatly increased the amount of information in the stego images.

An interesting anomaly occurred in the *PSNR* results for the S-Tools watermarked MRI image. The PSNR values were very low, although all four other metrics indicated that S-Tools embedding provided minimal image degradation. The reason for the result is unknown. In other results, S-Tools and Hide4PGP again provided similar values for the

Image	System	Watermark	Rel. Ent.	PSNR (dB)	MSE	MAE	Watson
MRI	S-tools	text 108 char	4.9678e-6	9.9224	9.7747e-4	9.7747e-4	0.0541
		text 1080 char	6.0984e-6	9.9224	0.0010	0.0010	0.0560
		4kb logo	0.0010	9.9228	0.0169	0.0169	0.2421
		40kb logo	0.0297	9.9251	0.1156	0.1156	0.4399
	Hide4PGP	text 108 char	1.4808e-5	37.2769	0.0017	0.0017	0.2188
		text 1080 char	8.0463e-4	32.3498	0.0167	0.0167	0.3035
		4kb logo	0.0083	29.7059	0.0564	0.0564	0.4387
		40kb logo	0.3899	23.0330	0.7826	0.6091	1.6202
CT	S-tools	text 108 char	3.0532e-6	42.8411	3.7492e-4	3.7492e-4	0.3158
		text 1080 char	2.6226e-7	42.7358	4.0190e-4	4.0190e-4	0.2396
		4kb logo	4.3526e-5	36.7336	0.0065	0.0065	0.7185
		40kb logo	0.0017	32.4548	0.0445	0.0445	1.5455
	Hide4PGP	text 108 char	2.1947e-7	43.9802	2.3574e-4	2.3574e-4	1.2268
		text 1080 char	7.5416e-6	39.1583	0.0021	0.0021	1.0032
		4kb logo	5.2060e-5	36.5181	0.0073	0.0073	1.1681
		40kb logo	0.0045	31.4453	0.0747	0.0747	2.1601
	JPHide	text 108 char	0.3092	27.9301	0.1687	0.1687	13.1836
		text 1080 char	0.3211	27.8969	0.1713	0.1738	13.2734
		4kb logo	0.7275	24.9565	0.8264	0.5877	17.9248
		40kb logo	(1.1291)	(18.1810)	(21.6782)	(3.0818)	51.4543
CXR	S-tools	text 108 char	5.2204e-7	42.8291	3.9338e-4	3.9338e-4	0.3008
		text 1080 char	5.0935e-7	42.7473	3.7634e-4	3.7634e-4	0.3025
		4kb logo	5.8960e-5	36.7648	0.0063	0.0063	0.5646
		40kb logo	0.0022	32.4558	0.0446	0.0446	1.2507
	Hide4PGP	text 108 char	2.2899e-7	44.0885	2.4426e-4	2.4426e-4	1.2267
		text 1080 char	7.4450e-6	39.0936	0.0022	0.0022	1.0012
		4kb logo	7.2196e-5	36.5025	0.0073	0.0073	1.1571
		40kb logo	0.0060	31.4491	0.0750	0.0750	1.8501
	JPHide	text 108 char	0.0110	27.5046	0.2052	0.2052	6.0806
		text 1080 char	0.0114	27.5261	0.2032	0.2032	6.2000
		4kb logo	0.0475	24.4695	1.0285	0.7060	9.6438
		40kb logo	(0.3343)	(17.2415)	(34.2228)	(4.3160)	43.2141

Table 1. Differences between original and stego images, with four different watermarks.

CT and CXR images, and the poor performance of JPHide was clear.

MSE computation resulted in much lower values for S-Tools than Hide4PGP for the MRI image, due to the fact that S-Tools embeds much less information in an image. This result is reflected in the MSE values for the CT and CXR images. The results of JPHide were again significantly poorer than the other two systems, due to the greater image alterations that is causes.

The *MAE* results generally followed the same pattern as the MSE ones. Some MAE values were lower however, because the errors were not squared.

The *Watson* metric showed that S-tools was the best overall performer visually, providing lower perceptual errors than Hide4PGP and JPHide. The poor performance of JPHide was again evident. Embedding data using this system can cause great visual disturbances, as shown in Figures 1(e) and (f).

From the discussions above, some general conclusions have been reached about medical image watermarking, using these approaches. Firstly, S-Tools generally provides less image degradation than Hide4PGP or JPHide. Secondly, more research is required before systems such as S-Tools, which provide minimal image degradation, are used to embed watermarks in the images. This is because even high quality stego images may have small changes in image pixel values, which can change the interpretation of the image. Note that image interpretation is used by radiologists for diagnosis and in imaging applications such as automatic image segmentation.

It may be more appropriate to embed an invertible watermark, such as [2], which can be removed completely to attain the original image. Alternatively, if more robustness is required, embedding watermarks in non-ROI image sections, such as [19], is another possibility. For images such as Figure 2(c) however, this may not be possible, because if cropped, the ROI takes up the whole image. Given these issues, it is appropriate to conclude that different watermarks should be applied to different medical image types, and therefore systematic ways to achieve this should be investigated. An example where the same watermark will produce different effects on two different image types is using LSB embedding for (1) X-ray and (2) Ultrasound images. Image enhancement, a common operation on the X-ray images, will destroy patches of the watermark, where the image is brightened. On the other hand, denoising, which is commonly used to smooth Ultrasound images, will destroy the watermark on edges where the image has been smoothed. As stated, a systematic approach will be required to select the most appropriate watermarks for different medical image types.

6 Conclusion and Future Work

This preliminary study has shown that medical image watermarking is still an open field of research. This is primarily due to the special nature of the images, which should not be perceptually altered. The study compared three watermarking systems, applying their techniques to hide data in medical images. As expected, watermark robustness is related to a decrease in image quality. Also, even stego images from the most fragile system, S-Tools, resulted in perceptual image degradation. Thus, future work in the area should include considering invertible techniques, or ROI techniques if increased robustness is required, and that different watermarks should be applied to different medical image types.

Acknowledgments

The assistance of Queensland Health in providing the images for this work, Justin Boyle for providing the image quality metric code, and Craig Kennedy for his insightful comments on the subject, are noted with gratitude.

References

[1] G. Coatrieux, H. Main, B. Sankur, Y. Rolland, and R. Collorec. Relevance of watermarking in medical imaging. In *IEEE-embs Information Technology Applications in Biomedicine*, pages 250–255, Arlington, USA, Nov. 2000.

[2] J. Fridrich, M. Goljan, and R. Du. Invertible authentication. In *Proc. SPIE, Security and Watermarking of Multimedia Contents III*, volume 3971, pages 197–208, San Jose, USA, Jan. 2001.

[3] A. Giakoumaki, S. Pavlopoulos, and D. Koutsouris. A medical image watermarking scheme based on wavelet transform. In *Proc. of the 25th Annual Int. Conf. of the IEEE-EMBS*, pages 856–859, Cancun, Mexico, Sept. 2003.

[4] N. Jagadish, P. S. Bhat, R. Acharya, and U. C. Niranjan. Simultaneous storage of medical images in the spatial and frequency domain: a comparative study. *Biomedical Engineering Online*, 3(1):record 17, June 2004.

[5] N. F. Johnson, Z. Duric, and S. Jajodia. *Information Hiding: Steganograph and Watermarking - Attacks and Countermeasures*. Kluwer Academic Press, Dordrecht, the Netherlands, 2001.

[6] X. Kong and R. Feng. Watermarking medical signals for telemedicine. *IEEE Trans on. Information Technology in Biomedicine*, 5(3):195–201, Sept. 2001.

[7] M. Kutter and F. A. P. Petitcolas. A fair benchmark for image watermarking systems. In *Proc. SPIE Security and Watermarking of Multimedia Contents*, volume 3657, pages 226–239, San Jose, CA, USA, Jan. 1999.

[8] A. Latham. Steganography. Website: http://linux01.gwdg.de/ alatham/stego.html, 1999. accessed 21 January 2005.

[9] E. T. Lin, C. I. Podilchuk, and E. J. Delp. Detection of image alterations using semi-fragile watermarks. In *Proc. of the SPIE Int. Conf. on Security and Watermarking of Multimedia Contents II*, volume 3971, pages 152–163, San Jose, CA, USA, Jan. 2000.

[10] B. Macq and F. Dewey. Trusted headers for medical images. In *DFG VIII-D II Watermarking Workshop*, Erlangen, Germany, Oct. 1999.

[11] A. Maeder and M. Eckert. Medical image compression: Quality and performance issues. *SPIE: New Approaches in Medical Image Analysis*, 3747:93–101, 1999.

[12] M. Nishio, Y. Kawashima, S. Nakamuar, and N. Tsukamoto. Development of a digital watermark method suitable for medical images with error correction. RSNA 2002 Archive Site: http://archive.rsna.org/index.cfm, 2002. accessed 18 January 2005.

[13] D. Osborne, D. Abbott, M. Sorell, and D. Rogers. Multiple embedding using robust watermarks for wireless medical images. In *IEEE Symposium on Electronics and Telecommunications*, page section 13(34), Timisoara, Romania, Oct. 2004.

[14] W. Puech and J. M. Rodrigues. A new crypto-watermarking method for medical images safe transfer. In *Proc. of the 12th European Signal Processing Conference*, pages 1481–1484, Vienna, Austria, Sept. 2004.

[15] H. Repp. Hide4PGP info & demo page. Website: http://www.heinz-repp.onlinehome.de/Hide4PGP.htm, Year unspecified. accessed 17 January 2005.

[16] Spychecker. S-tools 4.0 steganography tool. website: http://www.spychecker.com/program/stools.html, Nov. 2000. accessed 18 January 2005.

[17] H. Tachibana, H. Harauchi, T. Ikeda, Y. Iwata, A. Takemura, and T. Umeda. Practical use of new watermarking and vpn techniques for medical image communication and archive. RSNA 2002 Archive Site: http://archive.rsna.org/index.cfm, 2002. accessed 4 January 2005.

[18] S. Voloshynovskiy, S. Pereira, V. Iquise, and T. Pun. Attack modelling: Towards a second generation watermarking benchmark. *Signal Processing, Special Issue on Information Theoretic Issues in Digital Watermarking*, 81(6), June 2001.

[19] A. Wakatani. Digital watermarking for ROI medical images by using compressed signature image. In *Annual Hawaii Int. Conf. on System Sciences*, pages 2043– 2048, Hawaii, USA, Jan. 2002.

[20] A. Watson. DCT quantization matrices visually optimized for individual images. In *Proc. SPIE: Human vision, visual processing and digital display IV*, volume 1913, pages 202–216, 1993.