Detecting Hidden Messages Using Image Power Spectrum

Palak Amin and K. P. Subbalakshmi
Department of Electrical and Computer Engineering
Stevens Institute of Technology
Hoboken, New Jersey 07030
Email: {pamin,ksubbala}@stevens.edu

***Abstract*— In this paper we present a study of the effects of data hiding on the power spectra of digital images. Several imperceptible data hiding techniques have been proposed that provide strong visual security and robustness. Although imperceptible to the human visual system, the hidden data affects the natural qualities of the image, such as the image power spectrum. In this study, we classify a large image database into a number of categories. For each category, we calculate the slope of the power spectra for the marked and unmarked images. We note that in the case of spatial data hiding the average slope of the power spectra of marked images is 54.93% higher compared to that of the unmarked images. Also in the cases of transform domain data hiding we note that the average slope of the power spectra of the images marked using a discrete cosine (wavelet) transform (DC(W)T) based technique is higher by 9.12% (38.39%). We also test a commercially available data hiding software namely Digimarc corp.'s MyPictureMarc 2005 V1.0. In this case the average power spectra of the marked images is 35.99% higher. Hence we see that the proposed scheme is a tool for universal steganalysis with varying degrees of success depending on the type of embedding.**

Keywords: Steganalysis, power spectra, data-hiding, watermarking

I. Introduction

Steganography addresses the issue of hiding information in some host data for various applications including covert communication, copy protection, copyright protection etc [1], [2], [3], [4], [5], [6], [7], [8], while steganalysis is concerned with discovering the presence of hidden messages and/or extracting the hidden message. Some definitions and several methods of steganalysis were proposed in the literature [9]. Voloshynovskiy et. al. [10], present an overview of some key characteristics that could lead to detecting the existence of hidden information. Johnson et al. [9] gives a good description of steganalysis of several popular steganographic software. Current steganalytic methods can be classified into two categories: one category aims at a specific embedding method and the other aims at universal stego detection. A good amount of research has been done on detecting specific steganographic methods. For instance Fridrich et. al.[11][12][13] have proposed methods that detect data hidden using popular algorithms such as Outguess[14], F5 [15] etc.. Farid [16] proposed a universal steganalysis method based on the Fisher linear classifier[17]. Chandramouli et. al [18] propose a general framework for a steganalysis system. In [19], a Gaussian distributed hidden message was shown to be detectable due to the non-Gaussian nature of the cover data. Harmsen [20] modelled message embedding as noise addition which can amount to low pass filtering of the histogram of the image.

In this paper, we investigate the effects of data hiding on the power spectra of images and devise a steganalysis algorithm that can detect stego data hidden within images using several classes of data-hiding techniques. Section-II gives an overview of the power spectrum of digital images. In Section-III we provide the classification of the image database and present the properties of the power spectra within these classes. In Section-IV we provide experimental results of power spectra calculation after processing these images using different data hiding algorithms. Finally, Section-V summarizes our work.

II. Power Spectrum

The power spectrum of an image is a crucial image statistic and is defined [21] as,

$$\Gamma(k_x, k_y) = \frac{1}{N^2}|I(k_x, k_y)|^2 \tag{1}$$

where,

$$I(k_x, k_y) = \frac{1}{N^2}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} i(x,y)e^{(-\frac{j2\pi}{N}(xk_x+yk_y))} \ ; \tag{2}$$

where, N is the image dimension in either direction (power spectrum of a non-square image is calculated by cropping the image to a largest concentric square sub-image) and k_x/N and k_y/N are the discrete spatial frequencies. Torralba et. al. [22] prove that the average power spectrum of natural images assumes the form $1/f^{\alpha}$, where $1 \leq \alpha \leq 2$. The power spectra of images is usually represented as the slope of the log-log graph of the spatial frequency versus the amplitude of the image as seen in Figure-1. Different environment categories of images exhibit different orientations and frequency distributions, which is captured by the images' power spectra [23][24][22]. In Figure-2 we show spectral signatures for different image categories. Each spectral signature is an average of 25 test images within the category. In Figure-2 we show that the differentiation between various image categories

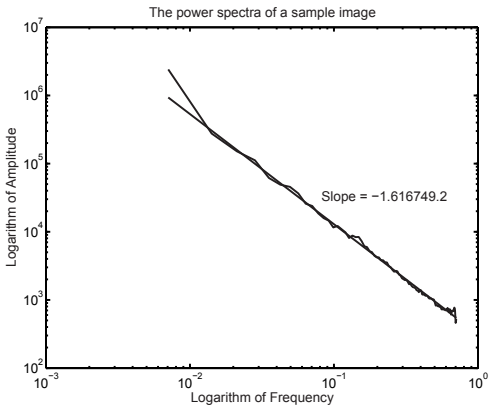

Fig. 1. Power spectra of a test image of an airplane in flight

Image Category	Avg. Slope of Power Spectra
Airplanes	-1.4140
Buildings	-1.2610
Cars	-1.2898
Cities	-1.3069
Crowds	-1.2889
Faces	-1.3123
Underwater Fishes	-1.3168
Forests	-1.2772
Integrated Circuits	-1.1179
Microscope Slides	-1.2979
MRI Scans	-1.2691
People	-1.2916
Satellite Images	-1.2970
Ships	-1.2547
Space Images	-1.2727
Storms	-1.3294
Artificial Textures	-1.0513
Tornadoes	-1.3618
Underwater Images	-1.3349
X-Ray Images	-1.2682

TABLE I
AVERAGE SLOPE OF THE POWER SPECTRA FOR EACH IMAGE CATEGORY

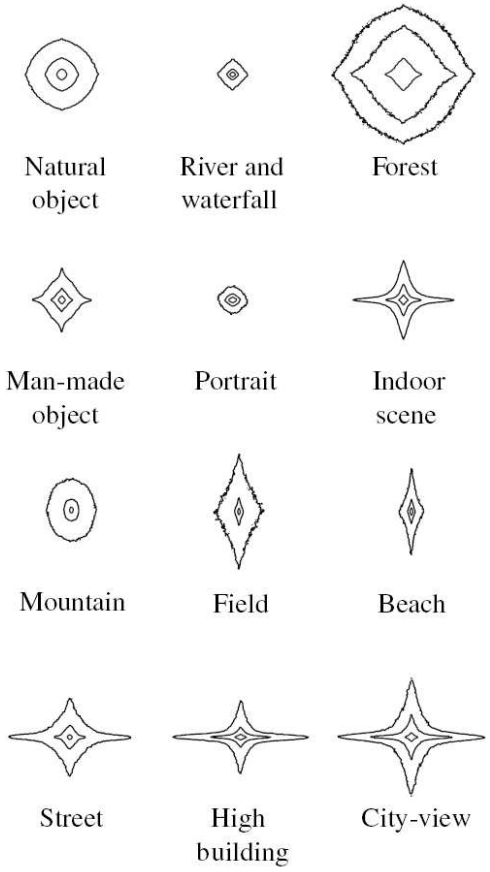

Fig. 2. Spectral Signatures of Different Image Categories

resides mainly in the relationship between the horizontal and vertical contours at different scales. These studies show that the shape of the spectral signature of an image is correlated with the scale at which major objects in the image are found (e.g. finer texture and detail in a forest scene, coarser texture and noise in a waterfall scene, etc.)

III. POWER SPECTRA OF THE IMAGES IN THE DATABASE

The image database used in this study comprised of a collection of 500 images of scenes, objects and people, taken from the NASA stock photo library [25], the coral NOAA photo library [26], and the national geodetic survey [27]. These images are classified into one of 20 categories, based on characteristics such as backgrounds, depths, resolutions, etc. Figure-3 shows a collection of representation images created by averaging several images from the same category in the database. In Table-I, we provide an average value for the slope of the power spectra for each of the categories in the database.

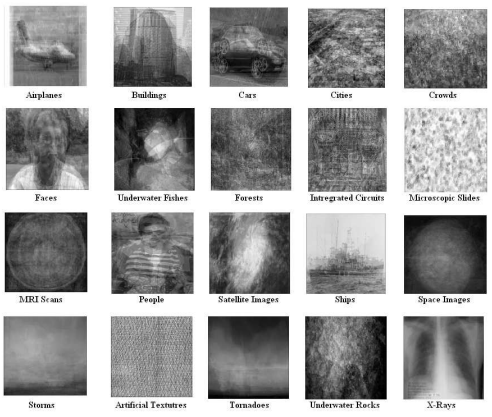

Fig. 3. Average Representation Images Generated for Each Semantic Category in the Database

IV. Experimental Results

We marked the test image dataset using four data hiding techniques, namely: correlation based spatial technique [28], discrete cosine transform (DCT) based technique [29], discrete wavelet transform (DWT) based technique [30], and Digimarc's MyPictureMarc 2005 V1.0 adobe photoshop plugin [31]. Equations-3, 4, and 5 give the embedding function used in the spatial, DCT, and DWT data hiding schemes respectively.

$$\tilde{I}_m(x,y) = I_m(x,y) + kW_b, \tag{3}$$

where, k is the gain factor used to specify the strength of the embedded data, W_b is the appropriate pseudo random noise sequence, based on the hidden bit being a 0 or a 1. $I_m(x,y)$ represents the m^{th} 8x8 spatial block of the host image and $\tilde{I}_m(x,y)$ represents the corresponding marked spatial block. Each block, $\tilde{I}_m(x,y)$, is then spatially rearranged to give the final stego image $\tilde{I}(x,y)$.

$$\tilde{I}_m(u,v) = \begin{cases} I_m(u,v) + kW_b, & \text{if } u,v \in F_M \\ I_m(u,v), & \text{otherwise} \end{cases} \tag{4}$$

where, k is the gain factor used to specify the strength of the embedded data, F_M represents the midband DCT coefficients, W_b is the appropriate pseudo random noise sequence, based on the hidden bit being a 0 or a 1. $I_m(u,v)$ represents the m^{th} 8x8 DCT block of the host image and $\tilde{I}_m(u,v)$ represents the corresponding marked DCT block. Each block, $\tilde{I}_m(u,v)$, is then inverse transformed to give the final stego image $\tilde{I}(x,y)$.

$$\tilde{I}_m(u,v) = \begin{cases} I_m(u,v) + k|I_m(u,v)|h_i, & \text{if } u,v \in HL, LH \\ I_m(u,v), & \text{otherwise} \end{cases} \tag{5}$$

where, k is the gain factor used to specify the strength of the embedded data, HL and LH represent the midband DWT coefficients, h_i is the hidden bit that is to be embedded. $I_m(u,v)$ represents the m^{th} DWT coefficient of the host image and $\tilde{I}_m(u,v)$ represents the corresponding marked DWT coefficient. Each coefficient, $\tilde{I}_m(u,v)$, is then inverse transformed to give the final stego image $\tilde{I}(x,y)$.

In Table-II, we summarize the average slope of the power spectra for each of the category of images marked with each of the data hiding algorithm above. In Table-III, we summarize the percent difference values between average slope of the power spectra for each of the category of images marked with each of the data hiding algorithm and that of unmarked images. From Tables-II and -III, we note that the average power spectra of the marked images significantly changes compared to the unmarked host images. Although each of the four algorithms are designed as imperceptible means of data hiding, we note that they significantly increase the slope of the power spectra of the images by marking the perceptibly significant areas of the image.

Image Category	Spatial	DCT	DWT	Digimarc
Airplanes	-0.6616	-1.2466	-0.8892	-0.8938
Buildings	-0.6337	-1.1758	-0.8438	-0.8369
Cars	-0.5266	-1.1633	-0.7322	-0.8076
Cities	-0.6913	-1.2181	-0.9155	-0.8858
Crowds	-0.6013	-1.2011	-0.8216	-0.8505
Faces	-0.6382	-1.2063	-0.8491	-0.8618
Underwater Fishes	-0.5878	-1.1883	-0.8044	-0.8491
Forests	-0.6049	-1.1765	-0.8154	-0.8340
Integrated Circuits	-0.6067	-1.0798	-0.8246	-0.7671
Microscope Slides	-0.5923	-1.1930	-0.8169	-0.8424
MRI Scans	-0.6607	-1.1936	-0.8929	-0.8797
People	-0.5860	-1.1794	-0.7970	-0.8223
Satellite Images	-0.4915	-1.1486	-0.6876	-0.7879
Ships	-0.5347	-1.1410	-0.7459	-0.7799
Space Images	-0.4591	-1.1156	-0.6624	-0.7808
Storms	-0.5293	-1.1411	-0.6799	-0.7707
Artificial Textures	-0.5293	-1.0066	-0.7537	-0.7050
Tornadoes	-0.4645	-1.1472	-0.6655	-0.7873
Underwater Images	-0.5914	-1.2048	-0.8088	-0.8557
X-Ray Images	-0.5178	-1.1195	-0.7195	-0.7746

TABLE II
AVERAGE SLOPE OF THE POWER SPECTRA FOR MARKED IMAGES

Image Category	Spatial	DCT	DWT	Digimarc
Airplanes	-53.21%	-11.84%	-37.11%	-36.79%
Buildings	-49.75%	-6.76%	-33.08%	-33.63%
Cars	-59.17%	-9.81%	-43.23%	-37.39%
Cities	-47.10%	-6.79%	-29.95%	-32.22%
Crowds	-53.35%	-6.81%	-36.26%	-34.01%
Faces	-51.37%	-8.08%	-35.30%	-34.33%
Underwater Fishes	-55.36%	-9.76%	-38.91%	-35.52%
Forests	-52.64%	-7.88%	-36.16%	-34.70%
Integrated Circuits	-45.73%	-3.41%	-26.24%	-31.38%
Microscope Slides	-54.36%	-8.08%	-37.06%	-35.10%
MRI Scans	-47.94%	-5.95%	-29.64%	-30.68%
People	-54.63%	-8.69%	-38.29%	-36.33%
Satellite Images	-62.10%	-11.44%	-46.99%	-39.25%
Ships	-57.38%	-9.06%	-40.55%	-37.84%
Space Images	-63.93%	-12.34%	-47.95%	-38.65%
Storms	-60.19%	-14.16%	-48.86%	-42.03%
Artificial Textures	-49.65%	-4.25%	-28.31%	-32.94%
Tornadoes	-65.89%	-15.76%	-51.13%	-42.19%
Underwater Images	-55.70%	-9.75%	-39.41%	-35.90%
X-Ray Images	-59.17%	-11.73%	-43.27%	-38.92%

TABLE III
PERCENTAGE DIFFERENCE OF THE AVERAGE SLOPE OF THE POWER SPECTRA BETWEEN MARKED AND UNMARKED IMAGES

V. Conclusion

We presented an in depth analysis of the effects of data hiding using several techniques on the image power spectra. Although, current data-hiding techniques are imperceptible to the human eye, we note from the experimental results that the image power spectra is severely affected by data hiding. From previous studies on natural image qualities and the experimental results obtained by our research we also conclude that the slope of the power spectra of images which are under the same category are very close to each other. From experimental results, we note that in the case of spatial data hiding the average power spectra of marked images is 54.93%

higher as compared to that of the unmarked images. In the cases of transform domain data hiding we note that the average power spectra of the images marked using a DCT (DWT) based technique is higher by 9.12% (38.39%). Upon testing a commercially available data hiding software namely Digimarc corp.'s picture marc 2005 V1.0, we note that the average power spectra of the marked images is 35.99% higher.

ACKNOWLEDGMENT

This work was funded by National Science Foundation (NSF) award number NSF-DAS 0242417.

REFERENCES

[1] I. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE transactions on image processing*, vol. 6, pp. 1673–1687, 1997.

[2] C. Lin, M. Wu, J. Bloom, I. Cox, M. Miller, and Y. Lui, "Rotation, scale, and translation resilient watermarking for images," *IEEE Transactions on Image Processing*, vol. 10, pp. 767–782, 2001.

[3] W. Fung and A. Kunisa, "Rotation, scaling, and translation-invariant multi-bit watermarking based on log-polar mapping and discrete Fourier transform," in *IEEE International Conference on Multimedia and Expo, 2005. ICME 2005.*, Japan, 2005.

[4] P. Amin and K. Subbalakshmi, "Rotation and cropping resilient data hiding with Zernike moments," in *International conference on image processing ICIP'04*, Singapore, Oct 2004.

[5] I. Cox and M. Miller, "The first 50 years of electronic watermarking," vol. 2002, no. 2, pp. 126–132, Feb. 2002.

[6] P. Amin, N. Liu, and K. Subbalakshmi, "Statistically secure digital image data hiding," in *IEEE Multimedia Signal Processing MMSP05*, China, Oct 2005.

[7] P. Moulin and A. Briassouli, "A stochastic QIM algorithm for robust, undetectable image watermarking," in *International conference on image processing*, Oct. 2004, pp. 1173–1176.

[8] N. Liu, P. Amin, and K. Subbalakshmi, "Secure quantizer based data embedding," in *IEEE Multimedia Signal Processing MMSP05*, China, Oct. 2005.

[9] N. Johnson and S. Jajodia, "Steganalysis of images created using current steganography software," *Lecture notes in computer science*, vol. 1525, pp. 273–289, 1998.

[10] S. Voloshynovskiy, A. Herrigel, Y. Rytsar, and T. Pun, "Stegowall:blind statistical detection of hidden data," *Proceedings of SPIE*, vol. 4675, pp. 57–68, Aug. 2002.

[11] J. Fridrich, M. Goljan, and R. Du, "Steganalysis based on JPEG compatibility," in *SPIE multimedia systems and applications IV*, Denver CO, 2001.

[12] J. Fridrich and M. Goljan, "Practical steganalysis of digital images - state of the art," *SPIE photonics west*, vol. 4675, pp. 01–13, 2002.

[13] J. Fridrich, M. Goljan, and D. Hogea, "Steganalysis of JPEG images: Breaking the f5 algorithm," in *5th information hiding workshop*, Netherlands, 2002.

[14] N. Provos, "Defending against statistical steganalysis," in *10th USENIX security symposium,*, Washington DC, 2001.

[15] A. Westfeld, "High capacity despite better steganalysis (F5A steganographic algorithm)," *Lecture notes in computer science*, vol. 2137, pp. 289–302, 2001.

[16] H. Farid, "Detecting steganographic message in digital images. report tr2001-412, dartmouth college, hanover, NH," 2001.

[17] R. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. VII, pp. 179–188, 1936.

[18] R. Chandramouli and N. Memon, "A distributed detection framework for steganalysis," in *ACM Multimedia Workshop*, Marina, USA, 2000, pp. 123–126.

[19] R. Chandramouli, "A mathematical framework for active steganalysis," in *ACM Multimedia Systems Journal*, 2003.

[20] J. Harmsen and W. Pearlman, "Steganalysis of additive noise modelable information hiding," in *Proceedings of SPIE Electronic Imaging*, vol. 5020, Santa Clara CA, 2003.

[21] R. Balboa and N. Grzywack, "Power spectra and distribution of contrasts of natural images from different habitats," *Pergamon Science Direct: Vision Research*, vol. 43, pp. 2527–2537, Mar. 2003.

[22] A. Torralba and A. Oliva, "Statistics of natural image categories," *IOP Network: Computation in neural systems*, vol. 14, pp. 391–412, Mar 2003.

[23] R. Baddeley, "The correlational structure of natural images and the calibrations of the spatial representations," *Cogn. Science*, vol. 21, pp. 351–372, 1997.

[24] A. Oliva, A. Torralba, A. Guerin-Dugue, and J. Herault, "Global semantic classification using power spectrum templates," *Proc. Challenge of image retrieval*, vol. 41, pp. 176–210, 1999.

[25] M. Hahn, "Great images in nasa," 2002. [Online]. Available: http://grin.hq.nasa.gov/

[26] O. of the CIO/HPCC, "NOAA photo library," 2003. [Online]. Available: http://www.photolib.noaa.gov/

[27] N. O. . A. Administration, "National geodetic survey," 2005. [Online]. Available: http://www.ngs.noaa.gov/

[28] G. Langelaar, I. Setyawan, and R. Lagendijk, "Watermarking digital image and video data," *IEEE Signal Processing Magazine*, vol. 17, pp. 20–43, 2000.

[29] J. Hernandez, M. Amado, and F. Perez-Gonzalez, "DCT-domain watermarking techniques for still images: Detector performance analysis and a new structure," *IEEE transactions on image processing*, vol. 9, no. 1, pp. 55–68, 2000.

[30] D. Wei, Y. Weiqi, and Q. Dongxu, "Digital image watermarking based on discrete wavelet transform," *Journal of Computer Science and Technology*, vol. 17, no. 2, pp. 129–139, 2002.

[31] D. Corp., "My picture marc 2005 v1.0 adobe photoshop plugin," 2005. [Online]. Available: http://www.digimarc.com