Image Steganalysis based on SVD and Noise Estimation: Improve Sensitivity to Spatial LSB Embedding Families

Abolfazl Diyanat
Department of Electrical Engineering
Sharif University of Technology
Tehran, Iran
Email: diyanat@ee.sharif.edu

Farshid Farhat
Department of Electrical Engineering
Sharif University of Technology
Tehran, Iran
Email: farhat@ee.sharif.edu

Shahrokh Ghaemmaghami
Department of Electrical Engineering
and Electronics Research Institute
Sharif University of Technology
Tehran, Iran
Email: ghaemmagh@sharif.edu

***Abstract*—We propose a novel image steganalysis method, based on singular value decomposition and noise estimation, for the spatial domain LSB embedding families. We first define a content independence parameter, DS, that is calculated for each LSB embedding rate. Next, we estimate the DS curve and use noise estimation to improve the curve approximation accuracy. It is shown that the proposed approach gives an estimate of the LSB embedding rate, as well as information about the existence of the embedded message (if any). The proposed method can effectively be applied to a wide range of the image LSB steganography families in spatial domain. To evaluate the proposed scheme, we applied the method to a large image database. Using a large image database, simulation results of our steganalysis scheme indicate significant improvement to both true detection and false alarm rates.**

***Keywords*-Singular value Decomposition (SVD), steganalysis, noise estimation, LSB steganography.**

I. Introduction

In spatial domain schemes, a steganographer modifies the cover medium in the spatial domain, such as encoding at the level of pixel LSBs (least significant bits) [1]. Spatial domain steganography methods, as compared to transform domain methods, are simpler and of lower computational complexity, where could have a larger impact on the cover signal structure and statistics [2].

Pevny in [3] proposed a steganalysis scheme based on subtractive pixel adjacency matrix. In this scheme, the transition probability of Markov modeling is used as features for a Support Vector Machine (SVM) based classification. In [4], *Zhang* used statistical modeling of pixel difference distribution for detection of stego images. He estimated the number of the zero difference values from stego images to calculate the error between the estimated and actual values.

Chiew in [5] proposed a novel blind steganalysis method by defining a set of matrices. The mean, variance, skewness, and kurtosis of these matrices are selected to construct the feature set. *Fridrich* in [6], defines three groups: Regular, Singular and unusable. Making changes to the LSB plane pushes the difference between the numbers of regular and singular groups towards zero, when the length of message increases. This is the basic idea behind the *Fridrich* method.

Most researchers have used statistical features of the cover signal for steganalysis [7], [4]. In this paper, we propose a scheme based on Singular Value Decomposition (SVD) that is a powerful matrix decomposition procedure shown to grab graphical and numerical characteristics of the image [9]. So far, not much work in this area has been reported, where examples are methods introduced in [10], [11]. *Gul* obtains SVs from 25 overlapping block in an image, and then applies log function to inverse SVs to get SvB_j. Finally, an average of SVB_j is used as feature [10]. In [12], *Gul* describes a method for JPEG compression based Perturbed Quantization (PQ) and shows that JPEG-based PQ steganography distorts linear dependencies of rows/columns of pixel values.

Smith estimates and models the noise present in an image [13]. Using this estimation, he shows how steganography introduces detectable changes made to this natural noise. *Goyal* models LSB embedding by additive noise in color images [14]. The difference between the close color pairs and unique color pairs is used as the detection factor.

Fridrich [7] used histogram characteristic function (HCF) to develop an additive noise model based steganalysis of the LSB embedding in color images, but their algorithm almost failed in the case of grayscale images. Ker [15] extended the detection of LSB matching; a skilled variant of the LSB alteration that was undetectable by standard LSB steganalysis methods. An empirical matrix was used by Ker to boost the detection probability of the HCF technique [7]. The empirical matrix, equal to the adjacency histogram, improved the Ker's detection results.

In our SVD based analysis, we define a DS (a difference value defined in section III) curve and trace changes to the curve introduced by the LSB embedding. To get a better estimation of the DS curve, we use seventh bit plane of image noise variance estimation that makes it possible to classify our database into several classes based on signal details. For each class, We use a rational, linear function to estimate points of the DS curve.

The rest of the paper is organized as follows. A background on the SVD and the noise estimation is given in section II. The proposed steganalysis scheme is introduced

in section IIIand simulation results are represented and discussed in section IV.Finally, the paper is concluded in section V.

II. Background

A. SVD Analysis

Singular value decomposition (SVD) is considered as one of powerful matrix analysis tools for both real and complex matrices. The SVD comes with many applications in the field of pseudo-inverse matrix calculation, the least mean square estimation, matrix estimation, etc. Let A denote a $W \times H$ matrix, of rank R ,with real elements. It is shown that matrix A can be decomposed as:

$$A = U\Sigma V^* \tag{1}$$

In (1), U and V^* are $W \times W$ and $H \times H$ matrices, respectively and Σ is a diagonal matrix whose elements $(\sigma_1, \sigma_2, ..., \sigma_N)$ are non-negative real numbers, where $N = min(W, H)$.

B. LSB Steganography

It is assumed here that the inputs of the steganographic system are grayscale images. The cover signal (grayscale image) is typified by an W-row, H-column ($W \times H$) matrix whose elements (pixels) are integer numbers between 0 and 255 for an 8-bit digital image.

Secure steganography methods use a secret key shared between transmitter and receiver. This secret key (K) ordinarily addresses a pseudo-random number generator (PRNG) as seed. The output of the PRNG locates candidate pixels of the cover signal to embed the covert message bits. Specifically, a typical LSB steganography method randomly changes some LSBs of the cover image, based on a pseudo-randomized version of the secret message using a PRNG. To hide the secret message in the cover signal, PRNG selects some LSBs of cover signal, subject to the embedding capacity and the embedding strategy. The mathematical relation between stego (S), cover (C), message (M) of length L, and secret key (K) could be defined as:

$$S_{i,j} = LSB_K(C_{i,j}, M_l) \tag{2}$$

where $1 \leq i \leq W$, $1 \leq j \leq H$ and $1 \leq l \leq L$.

C. LSB Mathematical Modeling

We model the LSB embedding by additive independent noise [7], so:

$$S = LSB_K(C, M) \Longrightarrow S = C + noise \tag{3}$$

Assume S is the stego image, C is the cover image and Sv_i and Sv'_i is ith singular value of S and C. It can be proved that:

$$\begin{aligned} \Sigma_{i=1}^{W}\Sigma_{j=1}^{H} S(i,j)^2 &= \Sigma_{i=1}^{min(W,H)} Sv(i)^2 \\ \Sigma_{i=1}^{W}\Sigma_{j=1}^{H} C(i,j)^2 &= \Sigma_{i=1}^{min(W,H)} Sv'(i)^2 \end{aligned} \tag{4}$$

Table I: List of notations

Notation	Description
σ_i	SVs value
I	Stego gray sacle Image
I_0	It is generated by zeroing LSB bit plane
I'	It is generated by putting 7th bit plane (B_7) to LSB plane
$S(i)$	Average of some SVs at rate i
$S0(i)$	Average of some SVs for I_0 at rate i
$DS(i)$	Difference between $S(i)$ and $S0(i)$
B_7	7th bit plane
σ_{I_0}	I_0 noise variance
σ_{B_7}	B_7 noise variance
t	Number of part
d_1	Distance between Line 1 and 2 in zero
d_2	Distance between Line 1 and 2 in 50

Consider 3 and 4, and suppose noise has zero mean:

$$\Sigma_{i=1}^{min(W,H)} Sv'(i)^2 \leq \Sigma_{i=1}^{min(W,H)} Sv(i)^2 \tag{5}$$

As stated in (5), embedding secret data into the cover signal results in enlargement of the SVs energy.

D. Noise Estimation

As mentioned in section III, we need to estimate the additive noise using the signal at the receiver end. Several methods have been proposed in literature for noise estimation. Donoho [18] has proposed a method for noise estimation based on wavelet transform. According to the method, the noise variance is estimated from the HH component of the wavelet transform of the image, as:

$$\hat{\sigma} = \frac{Median(|W_i|)}{0.6745} \qquad W_i \in subbandHH \tag{6}$$

E. Definition of Notations

We assume that the aim is to hide message M into the image signal. We use a LSB steganography algorithm which randomly embeds message into image pixels. A list of notations is given in Table I.

III. Proposed Method

A. General Schema of Steganalysis Algorithm

General schema of the proposed algorithm for image steganalysis is shown in Figure1. The *SVD analysis* block shown in Figure1 performs the analysis by taking the following steps:

❶ Compute SVD from the image
❷ Choose some SVD values
❸ Make averaging

A grayscale image (I) of the size $W \times H$ has an actual capacity to place a $W \times H$-bit hidden message through a typical LSB embedding method. Now assume that just i-percent of the capacity of I is used. At the first step of the steganalysis, singular values of image I are calculated $SVD(I) = (\sigma_1, \sigma_2, \ldots, \sigma_N)$, where N is the number of singular values, $N = min(W, H)$.

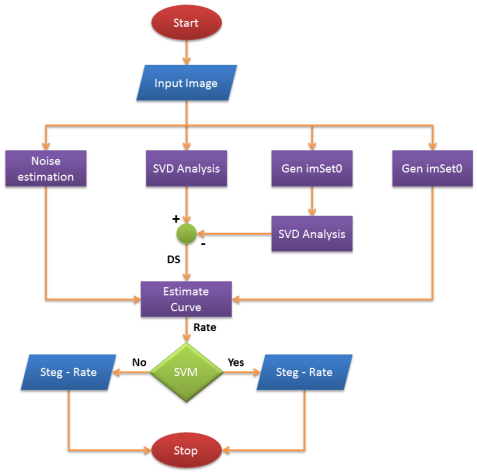

Figure 1: Block Diagram of Proposed method

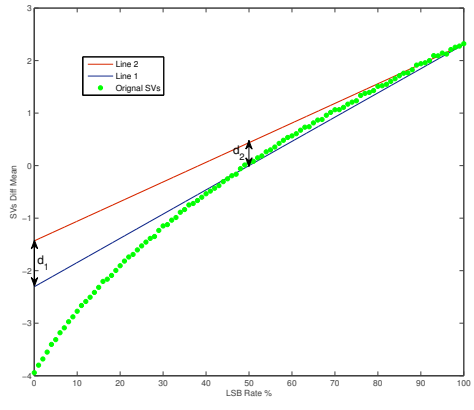

Figure 2: d_1 and d_2 for *Cameraman* image.

The LSB bit plane is the indicator of the details of an image, where the small variations of the matrix are found in lower SVD coefficients. Thus, we observe the lower half of the SVD coefficients $(\sigma_{\lceil \frac{N}{2} \rceil}, \ldots, \sigma_N)$ for steganalysis. Due to the assumptions made, the parameter $S(i)$ is defined as:

$$S(i) = \frac{1}{N'} \Sigma_{j=N'}^{N} \sigma_j \qquad N' = \lceil \frac{N}{2} \rceil \quad 0 \leq i \leq 100 \tag{7}$$

where i is LSB embedding rate. To impose the effect of the LSB embedding on the SVs, first, the image I_0 is generated by zeroing the LSB bit plane of the signal. Then, we compute the parameter $S(i)$ of the image, as described in the above. The result of the operation is indexed by $S0(i)$. The effect of LSB is hidden in a parameter called $DS(i)$, which is independent from the image content.

$$DS(i) = S(i) - S_0(i) \quad 0 \leq i \leq 100 \tag{8}$$

Figure2 shows DS values for different LSB embedding rates for the *Cameraman* image. As expected, DS changes almost linearly vs. the embedding rate. The DS-rate curve is estimated for a given suspicious image to discriminate an innocent image from a stego image. We try to fit DS curve by a quadratic curve. This procedure requires taking three points on the DS curve. We show these three points by $DS(0), DS(50), DS(100)$.

$$\begin{aligned} DS &\approx ax^2 + bx + c \\ a &= \frac{DS(100) + DS(0) - 2DS(50)}{5000} \\ b &= \frac{-DS(100) - 3DS(0) + 4DS(50)}{100} \\ c &= DS(0) \end{aligned} \tag{9}$$

B. Estimation Curve

The DS curve consists of values of DS at different rates between 0% to 100%. Finding the DS value for 100% $(100, DS(100))$ LSB embedding rate is a trivial task using *SVD Analysis* procedure. This is because the $DS(100)$, is generated by a message of the length of the number of image pixels, which is hidden in all LSBs of the signal.

Suppose $DS'(0)$ is the value of DS curve for I' at zero point. We generate I' by replacing the LSB plane with a replica of the 7th bit plane (B_7) (for an 8-bit grayscale image). This way, we make an approximation to the innocent image, which is supposed to be independent of the message at the expense of increasing the inherent correlation between image pixels, on one hand, and reducing the image resolution on the other hand. By making this approximation, we define a limit that is used in our steganalysis stage. We introduce two bounding lines over the DS curve, and then use σ_{B_7} as a parameter to specify the distance from these lines that are defined as:

f_1: Line passes through $(100, DS(100))$ and $(50, 0)$.
f_2: Line passes through $(100, DS(100))$ and $(0, DS'(0))$.

The distance of the DS curve from *Line1* will depend on the amount of randomness of the image LSBs, such that higher randomness is associated with a larger distance between this line and the curve, and vice versa. Henceforth, the image can be classified by evaluating the following parameters:

1) σ_{B_7}: Noise variance of the next upper plane (B_7).
2) $\sigma_{I'}$: Noise variance of the I'.

Considering the abovementioned parameters, we divide our database into several parts and subsequently estimate the DS curve for each part. We use the following relations to estimate $DS(0)$ and $DS(50)$.

$$DS(0) = x_0 + d_1.f_1(\sigma_{B_7}) \tag{10}$$

$$DS(50) = d_2.f_2(\sigma_{B_7}) \tag{11}$$

d_1 and d_2 are shown in Figure2 and f_1 and f_2 are given as:

$$f_1(\sigma_{B_7}) = \frac{\sigma_{B_7}}{a - \sigma_{B_7}} \tag{12}$$

$$f_2(\sigma_{B_7}) = \frac{b - \sigma_{B_7}}{c} \tag{13}$$

Table II: Databases properties

DB Name	ImageType	ImageNumber	size
NRCS [19]	Gray	2375	512×512
USC [20]	Gray , Color	44	512×512 , 256×256
Corel [21]	Gray	8185	512×512
USID	Color	1339	512×384 , 384×512
DB2	Gray	3165	512×512

To find f_1, we use the MMSE (minimum mean square error) based method to estimate a and b (Refer to (14)). The f_2 parameters (c, d) are estimated through linear MMSE (LMMSE) based method, as:

$$\frac{-1}{d} = \mu_{\sigma_{B_7}} - \frac{b}{c} \times \mu_{f_2(\sigma_{B_7})} \tag{15}$$

$$\frac{c}{d} = C_{\sigma_{B_7} f_2(\sigma_{B_7})} C^{-1}_{f_2(\sigma_{B_7})} \tag{16}$$

In (15), $C_{f_2(\sigma_{B_7})}$ is the covariance matrix of $f_2(\sigma_{B_7})$, and

$$C_{\sigma_{B_7} f_2(\sigma_{B_7})} = E\{(\sigma_{B_7} - \mu_{\sigma_{B_7}})(f_2(\sigma_{B_7}) - \mu_{f_2(\sigma_{B_7})})\}$$

x_0 depends on the class of image, which is selected as:

$$x_0 = \begin{cases} DS'(0) & \text{for} \quad \sigma_{B_7} \leq 0.35 \\ -DS(100) & \text{for} \quad \sigma_{B_7} \geq 0.35 \end{cases} \tag{17}$$

IV. Experiments

In this section, we refer to our MATLAB simulation results to evaluate the proposed steganalysis algorithm that is run over several standard image databases containing 15108 images of various types. Table II gives some details of the databases we have used.

We use a SVM (support vector machine) classifier for the detection process. Let $x_i, i = 1, 2, \ldots, N$, be the feature vectors of the training set, X, which are used to detect the class of image that could be either *clear* or *stego*. The goal is to design a hyperplane, as:

$$g(x) = w^T x + w_0 = 0 \tag{18}$$

that is expected to classify correctly the training vectors. To find the desired hyperplane, we use Quadratic Programming (QP) for the optimization (18). The polynomial function of order 5 to 8 for each class is used as the SVM kernel, based on an investigation we conducted on the type of the kernel. Selection of the *Kernel* influences greatly performance of the learning process [22]. Our training set includes the LSB embedding at rates 0 to 8%.

We use the conventional LSB replacement method for the steganography algorithm, so embed message bits in random places. First, the image is selected randomly from our database, and then is embedded at rate between 0 to 60 percentage. The stego images, as well as the clear images, undergo the detection procedure. Errors encountered at different rates within the given range is shown in Figure3 for 6000 images. Error at point zero indicates the false alarm error.

Table III: Comparison between steganalysis methods performance

Method	0.05%	0.10%	0.20%	0.40%
Our Method	48.5300	80.9100	95.3960	99.5344
WAM(72D)	0.6254	0.7749	0.8864	0.9103
BSM(18D)	0.6138	0.7394	0.8425	0.8675
FARID(72D)	0.5486	0.6005	0.6876	0.7673

Table III compares performance of the proposed method, applied to 6000 images, to that of other Steganalysis methods. The data shown in this table are extracted from [10], [11].

As Table III and Figure3 show, the proposed method, especially at low LSB embedding rates, outperforms the other methods in terms of detection accuracy. The *receiver operating characteristic (ROC)* for the proposed method and three well-knownschemes(RS [23], Sample Pair [24], and WS [25]) are displayed in Figure4, resulting from 1363 experiments conducted over both clear and stego images for embedding rates between 15 to 40 percent.

V. Conclusion

We have used singular value decomposition and noise estimation for grayscale image steganalysis. We have defined a content independence parameter (DS) and have shown that the method can be applied to any LSB embedding families. We obtain DS_i and estimate the DS curve for different embedding rate. We have used noise estimation of 7th bit-plane and two functions f_1 and f_2 to improve the curve approximation. It is one of the features of the present approach that the LSB embedding rate can be estimated in addition to the message existence detection. To evaluate the proposed scheme, we have applied our method to a large database. The simulation results show that the new method outperforms well-known steganalysis methods in terms of detection accuracy.

References

[1] A. Cheddad, J. Condell, K. Curran, and P. Mc Kevitt, "Digital image steganography: Survey and analysis of current methods," *Signal Processing*, vol. 90, no. 3, pp. 727–752, 2010.

[2] P. Alvarez, "Using Extended File Information (EXIF) File in Digital Evidence Analysis," *International Journal of Digital Evidence*, vol. 2, no. 3, 2004.

[3] T. Pevny, P. Bas, and J. Fridrich, "Steganalysis by Subtractive Pixel Adjacency Matrix," *IEEE Transactions on Information Forensics and Security*, vol. 5, pp. 215–224, June 2010.

[4] T. Zhang, W. Li, Y. Zhang, E. Zheng, and X. Ping, "Steganalysis of LSB matching based on statistical modeling of pixel," *Information Sciences*, vol. 180, pp. 4685–4694, 2010.

[5] K. L. Chiew and J. Pieprzyk, "Blind Steganalysis: A Countermeasure for Binary Image Steganography," in *International Conference on Availability, Reliability and Security*, pp. 653–658, Feb. 2010.

$$\begin{pmatrix} a \\ b \end{pmatrix} = \frac{\begin{pmatrix} -E(f_1(\sigma_{B_7})^2) & E(f_1(\sigma_{B_7})\sigma_{B_7}) \\ -E(\sigma_{B_7} f_1(\sigma_{B_7})) & E(\sigma_{B_7}^2) \end{pmatrix} \times \begin{pmatrix} -E(\sigma_{B_7}^2 f_1(\sigma_{B_7})) \\ -E(f_1(\sigma_{B_7})^2 \sigma_{B_7}) \end{pmatrix}}{-E(f_1(\sigma_{B_7})^2)E(\sigma_{B_7}^2) + E(f_1(\sigma_{B_7})x)E(\sigma_{B_7} f_1(\sigma_{B_7}))} \tag{14}$$

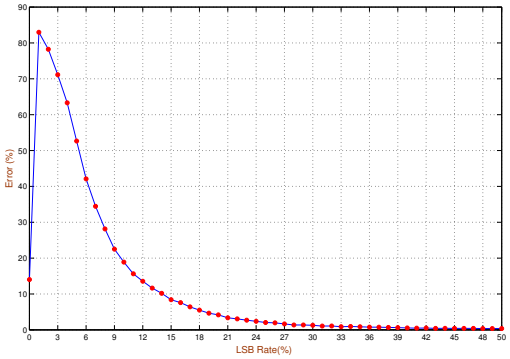

Figure 3: Error of our scheme for all rate between 0 to 60% LSB embedding rate.

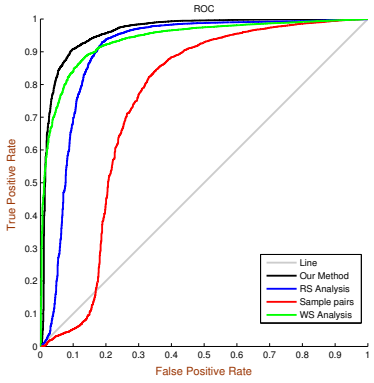

Figure 4: ROC for detection of LSB embedding

[6] J. Fridrich, G. Miroslav, and D. Rui, "Detecting LSB Steganography in Color and Gray-scale Images," *IEEE Trans. on Multimedia*, vol. 8, no. 4, pp. 22–28, 2001.

[7] J. Fridrich, M. Goijan, and D. Soukal, "Higher-order statistical steganalysis of palette images," *Proceeding of SPIE*, vol. 5020, pp. 178–190, 2003.

[8] J. Qin, X. Sun, X. Xiang, and Z. Xia, "Steganalysis based on difference statistics for LSB matching steganography," *Information Technology Journal*, vol. 8, no. 8, pp. 1281–1286, 2009.

[9] A. Shnayderman and A. Gusev, "A multidimensional image quality measure using singular value decomposition," *Proceedings of SPIE Image Quality and System Performance*, vol. 5294, pp. 82–92, 2004.

[10] G. Gul and F. Kurugollu, "SVD-based universal spatial domain image steganalysis," *Information Forensics and Security, IEEE Transactions on*, vol. 5, no. 2, pp. 349–353, 2010.

[11] G. Gul and F. Kurugollu, "A novel universal steganalyser design:LogSv," in *Image Processing (ICIP), 16th IEEE International Conference on*, pp. 4249–4252, IEEE, 2010.

[12] G. Gul and A. Dirik, "Steganalytic features for JPEG compression-based perturbed quantization," *Signal Processing Letters, IEEE*, vol. 14, no. 3, pp. 205–208, 2007.

[13] C. Smith, "Steganalysis using Noise Variance Estimation," in *Image Processing, ICIP. IEEE International Conference on*, vol. 1, pp. 417–420, IEEE, 2007.

[14] R. Goyal, S. Vijay, S. Agarwal, V. Laxmi, and M. Gaur, "Difference steganalysis for LSB embedding in images," in *Computer Design and Applications (ICCDA), International Conference on*, vol. 1, pp. V1–407 – V1–410, IEEE, 2010.

[15] A. D. Ker, "Steganalysis of LSB matching in grayscale images," *IEEE Signal Processing Letters*, vol. 12, no. 6, pp. 441–444, 2005.

[16] A. Ker, "A general framework for structural steganalysis of LSB replacement," in *Information Hiding*, pp. 296–311, Springer, 2005.

[17] M. Wall, A. Rechtsteiner, and L. Rocha, "Singular value decomposition and principal component analysis," *A practical approach to microarray data analysis*, pp. 91–109, 2003.

[18] D. Donoho and J. Johnstone, "Ideal Spatial Adaptation By Wavelet Shrinkage," *Biometrika*, vol. 81, no. 3, p. 425, 1994.

[19] [Online], "[Available]http://photogallery.nrcs.usda.gov/."

[20] [Online], "[Available]http://sipi.usc.edu/database/."

[21] "Corel stock photo library 3,Ontario, Canada. Available at: http://www.amazon.com/Corel-Stock-Photo-Library-3."

[22] R. Alhajj, H. Gao, X. Li, J. Li, O. Zaïane, J.-B. Li, S.-C. Chu, and J.-S. Pan, *Advanced Data Mining and Applications*, vol. 4632 of *Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007.

[23] J. Fridrich and M. Goljan, "Practical steganalysis of digital images-state of the art," in *Proc. SPIE Photonics West*, vol. 4675, pp. 1–13, Citeseer, 2002.

[24] S. Dumitrescu, "Detection of LSB steganography via sample pair analysis," *IEEE Transactions on Signal Processing*, vol. 51, pp. 1995–2007, July 2003.

[25] J. Fridrich, "On estimation of secret message length in LSB steganography in spatial domain," *Proceedings of SPIE*, vol. 5306, pp. 23–34, 2004.