

IMPROVED PERFORMANCE OF INVERSE HALFTONING ALGORITHMS VIA COUPLED DICTIONARIES

Pedro Garcia Freitas*, Mylène C.Q. Farias*⁺, and Aletéia P. F. de Araújo*

*Department of Computer Science, ⁺Department of Electrical Engineering
University of Brasília, Brasília, Brazil

ABSTRACT

Inverse halftoning techniques are known to introduce visible distortions (typically, blurring or noise) into the reconstructed image. To reduce the severity of these distortions, we propose a novel training approach for inverse halftoning algorithms. The proposed technique uses a coupled dictionary (CD) to match distorted and original images via a sparse representation. This technique enforces similarities of sparse representations between distorted and non-distorted images. Results show that the proposed technique can improve the performance of different inverse halftone approaches. Images reconstructed with the proposed approach have a higher quality, showing less blur, noise, and chromatic aberrations.

Index Terms— Inverse Halftone, Dictionary Training, Image Restoration, Sparse Coding, Sparse Modeling.

1. INTRODUCTION

Most printed materials are produced using halftoning techniques. Halftoning is the technique of converting continuous-tone images into images with a limited number of color levels. The technique generates images that, although having a limited number of levels, convey the illusion of having a higher number of levels. Inverse halftoning is the process of generating continuous-tone images from their halftoned versions. The reconstruction of scanned images is an application of inverse halftoning that is very important for the publishing industry [1]. Other applications include the protection of digital documents against piracy [2], authentication of video content [3], compression of images [4], and error concealment for images and videos [5]. In all these applications, the quality of the reconstructed image using the inverse halftoning algorithm is crucial.

Generally, algorithms proposed for inverse halftoning focus on the restoration of a specific kind of halftoning algorithm. In the case of dithering, there are several techniques to generate the patterns that create the illusion of a continuous-tone image. The corresponding inverse halftoning algorithms use the appropriate pattern to optimize image reconstruction. For example, Saika *et. al* [6] and Freitas *et*

al. [7] use stochastic models to restore continuous-tone image from Ordered Dithering (OD) halftones. But, when the halftoning technique is an error diffusion dithering technique, these approaches do not produce good results. This is a problem since error diffusion techniques [8] have a better performance than ordered dithering techniques [9, 10]

In this paper, we present a method to enhance the visual quality of images reconstructed using inverse halftoning techniques. We treat the inverse halftoned image as a distorted version of the original image and focus on recovering a non-distorted version. Although the technique presented here can be used to obtain the inverse halftone directly, the goal of the proposed technique is to detect and reduce distortions in reconstructed images generated by an inverse halftoning technique.

This paper is organized as follows. In Section 2, we describe the distortions associated with inverse halftoning algorithms. In Section 3, we discuss how to train a pair of dictionaries to match information between the original image and the reconstructed image. Section 4 details our strategy to improve inverse halftoning methods and Section 5 presents its results. Finally, in Section 6, we present our conclusions and discuss future works.

2. INVERSE HALFTONING

In this work, halftoning is the process of generating a binary (2 levels) image I_b from a continuous-tone (255-levels) gray-scale image I_g (or a channel of a colored image). Particularly,

$$I_b = \mathbf{H} \cdot I_g, \quad (1)$$

where \mathbf{H} is an operator representing the halftoning process that transforms I_g into the binary image (pixels values equal to ‘1’ or ‘0’). Although \mathbf{H} can be viewed as a simple operation, there are several ways to model it. As presented by Ulichney [11], different models correspond to different dithering patterns, leading to different ways to correlate the information between I_b and I_g .

Inverse halftoning is the process of reconstructing I_g from I_b . Since to generate I_b a considerable amount of information is discarded, the inverse halftoning algorithm can only generate an approximation of I_g . In other words, the reconstructed

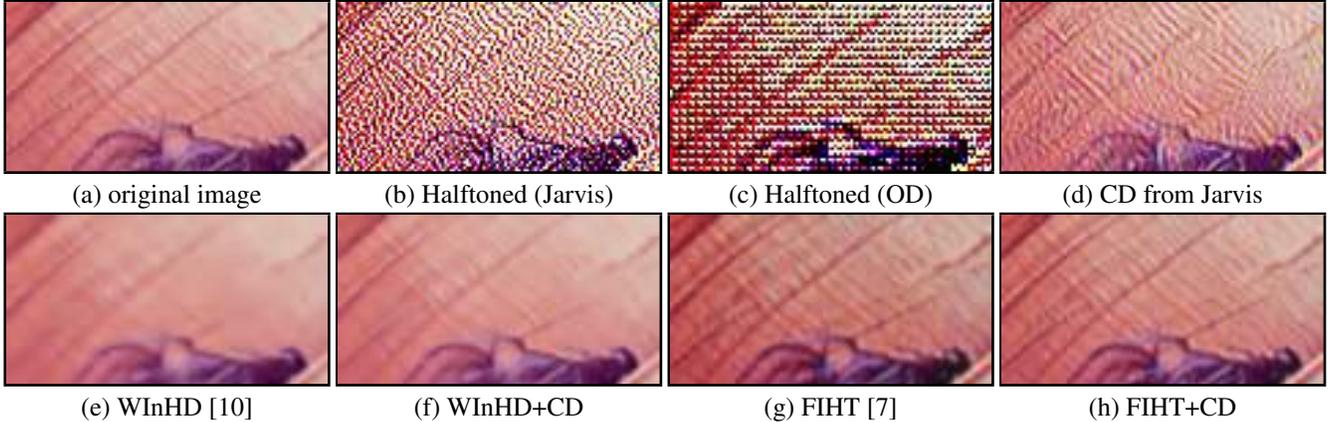


Fig. 1. Comparison of original, halftoned, and reconstructed images using different inverse halftoning methods (Lena).

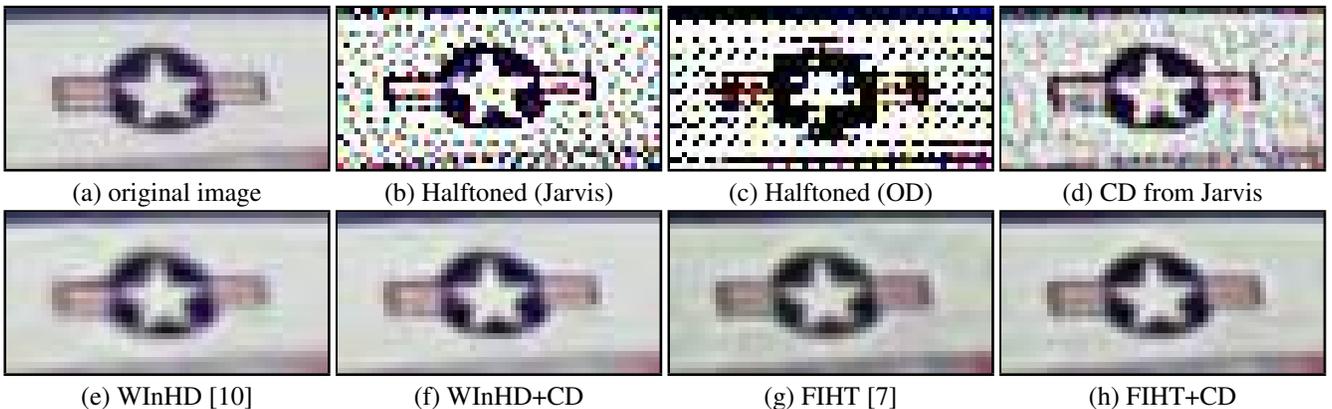


Fig. 2. Comparison of original, halftoned, and reconstructed images using different inverse halftoning methods (Airplane).

image \hat{I}_g contains an error, as modeled by following equation:

$$I_g = \hat{I}_g + \epsilon, \quad (2)$$

where ϵ is an approximation error and \hat{I}_g is the result of inverse halftoning process.

This approximation error ϵ is characterized by a degradation of the reconstructed image that we call *distortion*. The visual effect of this distortion varies according to the inverse halftoning technique, as shown in Fig. 1. The simplest inverse halftoning algorithm consists of low-pass filtering a halftoned image. This approach removes the noise introduced by halftoning patterns, but it also removes high-frequency information. Comparing the original Lena image (Fig. 1 (a)) and the reconstructed version generated using Wavelet-based inverse halftoning via deconvolution (WInHD) [10] (Fig. 1 (e)), we notice that the reconstructed image is blurred. Comparing the original (Fig. 1 (a)) and the image reconstructed using Fast Inverse Halftoning Algorithm for Ordered Dithered Images (FIHT) [7] (Fig. 1 (g)), we notice that the reconstructed image is noisier.

To tackle this problem, we observe the sparse representation of signals [12] and notice that there are linear relation-

ships among non-distorted and distorted signals. Therefore, it is possible to reconstruct an original image from its distorted version [13, 14]. To enhance the quality of reconstructed images, in this work we use a sparse representation of the original and halftoned images and an error concealment technique based on error correcting codes. The mathematical model used for the sparse representation is based on dictionary learning [15, 16], which is a matrix factorization problem that finds a dictionary (usually overcomplete) that sparsely encodes the fitted data. This work is inspired by the approach proposed by Yang *et al.* [17] that reduces blurring in interpolated images. Our work proposes a general model that reduces distortions in continuous-tone images reconstructed from halftoned images.

3. DICTIONARY LEARNING

Sparse representations make it possible to approximate a given signal, $x \in \mathbb{R}^d$, by a linear combination of elementary sets. These elementary sets are called *basic atoms* and can be stored in an over-complete dictionary $\mathbf{D} \in \mathbb{R}^{d \times k}$, where $d \ll k$. Sparse coding techniques consist of discover-

ing a good set of basic atoms to represent \mathbf{D} , what is usually done by learning from a set of training samples composed by patches of the signal x , i.e. $x = \{x_1, x_2, \dots, x_n\}$.

In the literature, there are many approaches to find dictionaries that guarantee the recovery of signals using sparse representations [16, 15, 18, 19]. In this paper, we use a learning-dictionary strategy that minimizes the differences between the signal and its sparse representation. In other words, we use the following optimization problem:

$$\begin{aligned} \arg \min_{\mathbf{D}, \alpha} \quad & \|x_i - \mathbf{D}\alpha_i\|_2^2 + \lambda \|\alpha_i\|_1 \\ \text{subject to} \quad & \|D_i\|_2^2 \leq 1 \quad i = 1, 2, \dots, m, \end{aligned} \quad (3)$$

where D_i is the i -th column of \mathbf{D} , λ is a constant that multiplies the ℓ_1 term (regularization parameter), and α_i is the sparse representation of x_i . This problem is convex when \mathbf{D} or $\{\alpha_i\}$ are fixed. When \mathbf{D} is fixed, $\{\alpha_i\}$ can be solved efficiently by linear programming, like in the Lasso problem [20, 21]. Otherwise, by fixing $\{\alpha_i\}$, \mathbf{D} can be solved as a constrained quadratic problem [22]. This joint optimization problem converges to a local minimum [23].

3.1. Coupled Dictionary Learning

Although the strategy described in the previous section is efficient to represent a signal x , we need a strategy to find the relationship between distorted and non-distorted signals. Our approach consists of learning two dictionaries, D_x and D_y , and find the relationship between these two spaces. For this, the distorted signal x in terms of D_x is the same as the non-distorted signal y in terms of D_y . A formulation to this problem is made by generalizing the single dictionary sparse coding as follows [17]:

$$\begin{aligned} \arg \min_{D_x, D_y, \alpha} \quad & \|x_i - D_x \alpha_i\|_2^2 + \|y_i - D_y \alpha_i\|_2^2 + \lambda \|\alpha_i\|_1 \\ \text{subject to} \quad & \|D_x\|_2^2 \leq 1 \\ & \|D_y\|_2^2 \leq 1 \end{aligned} \quad (4)$$

where

$$D_k = \arg \min_{D_k, \alpha} \|k_i - D_k \alpha_i^k\|_2^2 + \lambda \|\alpha_i^k\|_1, \quad (5)$$

for $k \in \{x, y\}$. We chose $\alpha_i^x = \alpha_i^y$ to guarantee that the distorted and non-distorted signals share the same sparse representation α_i .

Combining Equations 4 and 5, we can rewrite the optimization problem as:

$$\begin{aligned} \min_{D_c, \alpha} \quad & \|z_i - D_c \alpha_i\|_2^2 + \lambda \|\alpha_i\|_1 \\ \text{subject to} \quad & \|D_c\|_2^2 \leq 1, \end{aligned} \quad (6)$$

where

$$z_i = \begin{bmatrix} x_i \\ y_i \end{bmatrix} \quad (7)$$

and D_c is the coupled dictionary, given by:

$$D_c = \begin{bmatrix} D_x \\ D_y \end{bmatrix}. \quad (8)$$

4. PROPOSED METHOD

In this section, we discuss how to reduce distortions of images reconstructed with inverse-half-toning algorithms using the Coupled Dictionaries approach described in the previous section. For this approach, the distorted (y) and original (x) images are decomposed into sets of patches, i.e. $x = \{x_1, x_2, \dots, x_n\}$ and $y = \{y_1, y_2, \dots, y_n\}$, respectively. The main goal is to map the two spaces using coupled sparse dictionaries and, then, use these dictionaries to recover a patch of x_i from a patch of y_i ($i = 1, 2, \dots, n$).

4.1. Learning Stage

The first stage of the proposed method consists of extracting distorted and non-distorted patch-pairs from the distorted and non-distorted images, respectively. We use a database of training images, which is composed of a set of original (gray-scale) images and a set of corresponding images reconstructed using inverse half-toning algorithms. To generate the database, we first pick N training *non-distorted* images $\{I_g^n\}_{n=1}^N$. Then, we generate their N halftone versions $\{I_b^n\}_{n=1}^N$. Finally, for each halftoned image, we use an inverse half-toning algorithm to obtain each *distorted* image $\{\hat{I}_g^n\}_{n=1}^N$.

With these pairs of original and reconstructed images, I_g^n and \hat{I}_g^n , we compute a random permutation of correspondent 8×8 patches blocks, generating a set of original and reconstructed pairs of blocks. The pixels in the i -th pair of blocks, (b_i^n, \hat{b}_i^n) , are composed by the original information and its corresponding reconstructed information, respectively. These blocks are resized to vectors and referred as *patch pairs*. Patch pairs are appended to a list \mathbf{p} , in which the i -th element p_i is the previously defined pair (x_i, y_i) , where $x = \{x_1, x_2, \dots, x_n\}$ is the set of patches from original images (non-distorted) and $y = \{y_1, y_2, \dots, y_n\}$ is the corresponding set from reconstructed images (distorted).

Fig. 3 depicts the steps used for the extraction of the patch pairs. In step (1), a single image is converted to a halftoned image. In step (2), this halftoned image is reconstructed using an inverse half-toning technique. Then, in steps (3) and (4), the distorted and original images are divided into blocks. In step (5), the original image block and its corresponding block in the reconstructed image are extracted, generating the pair (b_i, \hat{b}_i) . Finally, in step (6), the non-distorted and distorted blocks are reshaped into vectors to create the patch pair $p_i = (x_i, y_i)$.

To learn dictionaries from distorted and non-distorted patches, we have to keep the sparse representation of non-distorted patches similar to the sparse representation of distorted

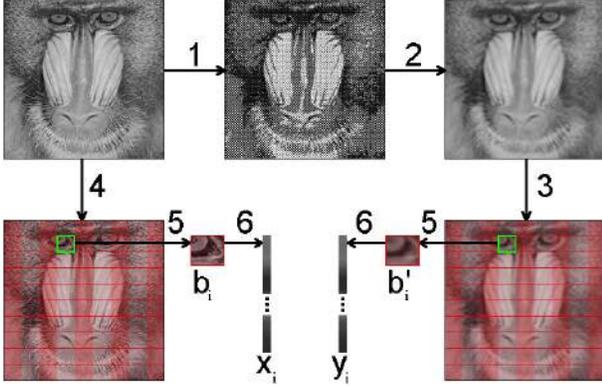


Fig. 3. First stage of proposed technique: extraction of patch pairs to create dictionaries.

ted patches. As mentioned earlier, non-distorted and distorted patches share the same $\{\alpha_i\}$, i.e. $\alpha_i^k = \alpha_i^x = \alpha_i^y$ in Equation 5. In other words, for representing distorted and original information, the dictionaries are coupled by sharing the same sparse representation.

4.2. Minimizing Distortions Stage

Once the training is done, we have two dictionaries, D_x and D_y , corresponding to non-distorted and distorted patches, respectively. To reduce the distortion of reconstructed images, we first extract the distorted patches $\{y_i\}$ from \hat{I}_g . Then, for each i -th distorted patch, we find the sparse representation corresponding to D_y , using the following equations:

$$\begin{aligned} \arg \min_{\alpha_i} \quad & \|\alpha_i\|_1 \\ \text{subject to} \quad & \|D_y \alpha_i - y_i\|_2^2 \leq \epsilon. \end{aligned} \quad (9)$$

This problem can be solved using one the several linear regression statistical techniques for ℓ_1 -norm [21, 23]. For a given optimal solution α_i , the corresponding non-distorted patch is given by:

$$x_i = D_x \alpha_i. \quad (10)$$

After non-distorted patches $\{x_i\}$ are obtained from the distorted ones $\{y_i\}$, we generate the reconstructed image \tilde{I}_g by substituting each non-distorted patch $\{x_i\}$ into the corresponding position in the binary image. In other words,

$$I_g = \tilde{I}_g + \sigma, \quad (11)$$

where $\sigma < \epsilon$ (see Equation 2) and, therefore, \tilde{I}_g is closer to I_g than \hat{I}_g .

5. EXPERIMENTAL RESULTS

We used a set of 64 images to train the dictionaries. These images were extracted from the ‘‘Texture’’ set of USC-SIPI

Image Database [24]. Fig. 4 shows a sample of these training images. The texture images varied from coarse to fine and from smooth to busy. Notice that the training images have a simple content, consistent with what is frequently found in blocks of larger and complex images.

To generate the distorted images, we used two different types of inverse halftoning techniques: FIHT [7] and WInHD [10]. These techniques were chosen because these algorithms use different approaches to reconstruct continuous-tone images. While WInHD reconstructs images from error diffusion halftones, FIHT restores images from ordered dithered halftones. As a consequence, the visual distortions produced by each algorithm look very different, as can be seen in Figs. 1 (e) and (g). Therefore, this choice of techniques makes it possible to test the proposed method for two very different halftoning algorithms. Both methods were implemented in Matlab® and the simulations were performed in a laptop with an Intel i7-4700MQ processor and 32GB of RAM.

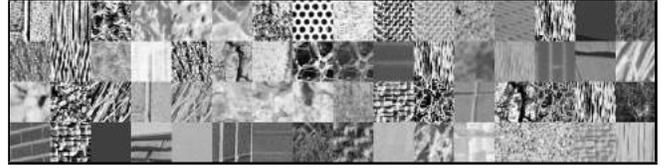


Fig. 4. Example of images used in the training stage [24].



Fig. 5. Images used in experimental tests. In order: Airplane, Baboon, Girl, House, Lena, Peppers, Sailboat, and Splash.

The proposed approach is tested using a set of 8 color natural images extracted from the ‘‘Miscellaneous’’ set of USC-SIPI Image Database [24]. These images are depicted in Fig. 5. For each image, to test the performance of the proposed technique for WInHD and FIHT, we generate an error diffusion halftoned image and an ordered dithered halftoned image, respectively. Since the images are colored, we generate two halftoned images for each RGB channel.

The results obtained for a detail of the image ‘Lena’ are shown in Fig. 1. Fig. 1 (a) corresponds to the original image, while Figs. 1 (b) and (c) correspond to its halftoned versions obtained using the Jarvis’s error diffusion method and

Table 1. SSIM values for images reconstructed with inverse halftoning algorithms.

Image	WInHD	WInHD+CD	FIHT	FIHT+CD	CD
Airplane	0.91833	0.92137	0.88987	0.91487	0.51259
Baboon	0.76191	0.76934	0.83559	0.85362	0.57378
Girl	0.82917	0.83345	0.88337	0.90417	0.59847
House	0.88919	0.89611	0.89713	0.92071	0.50451
Lena	0.88623	0.89091	0.86406	0.88107	0.41437
Peppers	0.85510	0.85742	0.84406	0.86044	0.47974
Sailboat	0.84204	0.84621	0.83499	0.85400	0.55574
Splash	0.89055	0.89378	0.85100	0.86376	0.48803

the ordered dithering method, respectively. Figs. 1 (e) and (g) show the restoration of the image in Fig. 1 (a) using Fig. 1(b) as input for the WInHD and FIHT algorithms, respectively.

As mentioned earlier, comparing Figs. 1 (a) and (e), we observe that the restoration using WInHD produces a blurred distortion. This distortion is concealed when we use the proposed technique (WInHD+CD), as shown in Fig. 1 (f). On the other hand, FIHT produces a sharper and noisier image, as shown in Fig. 1 (g). The noise is minimized using the proposed technique (FIHT+CD), as depicted in Fig. 1 (h).

Fig. 1 (d) shows the results of using the proposed technique directly on the image shown in Fig. 1 (b). This example shows that the training stage can be performed without using the inverse halftoning algorithm. Therefore, the proposed method can be used to restore graylevels directly from binary images, i.e. the proposed method can act as an inverse halftoning algorithm too.

Fig. 2 show the results for a detail of the ‘Airplane’ image. For better visualization of the distortions, a digital zoom was applied to the details of the image. When we compare the images in Figs. 2 (a), (e), and (f), we can notice that distortions in high-frequency content are minimized. Moreover, the noisy distortion inserted by FIHT method is also minimized by CD, as can be noticed when comparing Fig. 2 (a), (g), and (h).

Table 1 shows estimates of the quality of reconstructed images obtained using a popular full-reference image quality metric: the structural similarity index (SSIM) [25]. The second column of the table shows the SSIM values corresponding to the images obtained using the WInHD inverse halftone process, while the third column shows the SSIM values obtained using the proposed method for WInHD halftoned images. Similarly, the fourth and fifth column show the same results for the FIHT method. The last column shows the SSIM values corresponding to the images reconstructed using the CD directly on Jarvis’ halftoned images. We can notice that the use of CD increases the SSIM values of images reconstructed for WinHD and FIHT halftoning techniques.

The scalability of algorithm is analysed by varying the number of patches used in the training stage. As shown in Fig. 6, we can observe that the training time increases linearly with the number of samples. Moreover, the sampling

time is approximately constant (130 seconds average) and much smaller than the training time. This indicates that the main computational cost is, as expected, the optimization process. Therefore, an optimized implementation of the Lasso algorithm is required for implementation of the proposed algorithm on embedded hardware (e.g. scanning devices) or for real-time applications.

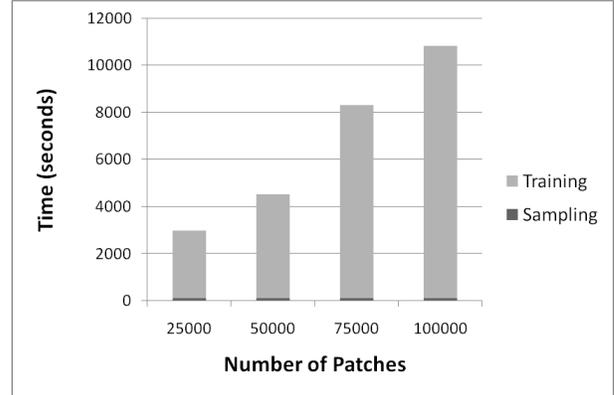


Fig. 6. Computation time of dictionaries as function of number of samples.

6. CONCLUSIONS

We have presented a novel approach to improve the quality of images reconstructed using inverse halftoning algorithms. The proposed approach uses coupled dictionaries, which are trained using original images and images reconstructed using inverse halftoning algorithms. Experimental results show that the proposed method is able to conceal distortions caused by inverse halftoning techniques, improving the quality of images generated by these techniques. Moreover, we have shown that coupled dictionaries can also be used as an inverse halftoning method. Future works include the investigation of the proposed method performance gain in inverse halftone applications, such as error concealment, image inpainting, digital content protection, and digital scanning software. Also, a study of the model’s optimal parameter values is needed. Finally, further investigation of the effect of the training image database on the quality of the restored images is necessary.

7. REFERENCES

- [1] Christopher M Miceli and Kevin J Parker, “Inverse halftoning,” *Journal of Electronic Imaging*, vol. 1, no. 2, pp. 143–151, 1992.
- [2] M Bharathi, R Charanya, PG Student, and T Vijayan, “Halftone visual cryptography & watermarking,” *International Journal of Engineering*, vol. 2, no. 4, 2013.

- [3] Chao-Yung Hsu, Chun-Shien Lu, and Soo-Chang Pei, "Content authentication of halftone video via flickering as sparse signal," in *Image Processing (ICIP), 2012 19th IEEE International Conference on*. IEEE, 2012, pp. 265–268.
- [4] Wei-Chih Chang and Li-Ming Chen, "Image compression system using halftoning and inverse halftoning to produce base and difference layers while minimizing mean square errors," Oct. 31 2000, US Patent 6,141,450.
- [5] Pedro Garcia Freitas, Ronaldo Rigoni, Mylene CQ Farias, and APFd Araujo, "Error concealment using a halftone watermarking technique," in *Graphics, Patterns and Images (SIBGRAPI), 2012 25th SIBGRAPI Conference on*. IEEE, 2012, pp. 308–315.
- [6] Y. Saika, K. Okamoto, and F. Matsubara, "Probabilistic modeling to inverse halftoning based on super resolution," in *Control Automation and Systems (ICCAS), 2010 International Conference on*, Oct 2010, pp. 162–167.
- [7] P.G. Freitas, M.C.Q. Farias, and A.P.F. de Araujo, "Fast inverse halftoning algorithm for ordered dithered images," in *Graphics, Patterns and Images (Sibgrapi), 2011 24th SIBGRAPI Conference on*, Aug 2011, pp. 250–257.
- [8] John F Jarvis, C Ni Judice, and WH Ninke, "A survey of techniques for the display of continuous tone pictures on bilevel displays," *Computer Graphics and Image Processing*, vol. 5, no. 1, pp. 13–40, 1976.
- [9] G.B. Unal and A.E. Cetin, "Restoration of error-diffused images using projection onto convex sets," *Image Processing, IEEE Transactions on*, vol. 10, no. 12, pp. 1836–1841, Dec 2001.
- [10] R.N. Neelamani, R. Nowak, and R. Baraniuk, "Model-based inverse halftoning with wavelet-vaguelette deconvolution," in *Image Processing, 2000. Proceedings. 2000 International Conference on*, 2000, vol. 3, pp. 973–976 vol.3.
- [11] Robert A Ulichney, "Review of halftoning techniques," in *Electronic Imaging*. International Society for Optics and Photonics, 1999, pp. 378–391.
- [12] John Wright and Yi Ma, "Dense error correction via minimization," *Information Theory, IEEE Transactions on*, vol. 56, no. 7, pp. 3540–3560, 2010.
- [13] John Wright, Yi Ma, Julien Mairal, Guillermo Sapiro, Thomas S Huang, and Shuicheng Yan, "Sparse representation for computer vision and pattern recognition," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1031–1044, 2010.
- [14] David L Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [15] Julien Mairal, Francis Bach, Jean Ponce, and Guillermo Sapiro, "Online dictionary learning for sparse coding," in *Proceedings of the 26th Annual International Conference on Machine Learning*. ACM, 2009, pp. 689–696.
- [16] Kenneth Kreutz-Delgado, Joseph F Murray, Bhaskar D Rao, Kjersti Engan, Te-Won Lee, and Terrence J Sejnowski, "Dictionary learning algorithms for sparse representation," *Neural computation*, vol. 15, no. 2, pp. 349–396, 2003.
- [17] Jianchao Yang, Zhaowen Wang, Zhe Lin, Scott Cohen, and Thomas Huang, "Coupled dictionary training for image super-resolution," *Image Processing, IEEE Transactions on*, vol. 21, no. 8, pp. 3467–3478, 2012.
- [18] Shenghua Gao, Ivor Wai-Hung Tsang, and Liang-Tien Chia, "Sparse representation with kernels," *Image Processing, IEEE Transactions on*, vol. 22, no. 2, pp. 423–434, 2013.
- [19] Yi Chen, Nasser M Nasrabadi, and Trac D Tran, "Hyperspectral image classification via kernel sparse representation," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 51, no. 1, pp. 217–231, 2013.
- [20] David L Donoho, "For most large underdetermined systems of linear equations the minimal 1-norm solution is also the sparsest solution," *Communications on pure and applied mathematics*, vol. 59, no. 6, pp. 797–829, 2006.
- [21] Robert Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.
- [22] Casper J Albers, Frank Critchley, and John C Gower, "Quadratic minimisation problems in statistics," *Journal of Multivariate Analysis*, vol. 102, no. 3, pp. 698–713, 2011.
- [23] Honglak Lee, Alexis Battle, Rajat Raina, and Andrew Y Ng, "Efficient sparse coding algorithms," *Advances in neural information processing systems*, vol. 19, pp. 801, 2007.
- [24] Allan G Weber, "The USC-SIPI image database version 5," *USC-SIPI Report*, vol. 315, pp. 1–24, 1997.
- [25] Zhou Wang and Alan C Bovik, "Mean squared error: love it or leave it? a new look at signal fidelity measures," *Signal Processing Magazine, IEEE*, vol. 26, no. 1, pp. 98–117, 2009.