Removal of Artifacts from JPEG Compressed Document Images

Basak Oztan[a*], Amal Malik[b], Zhigang Fan[b], Reiner Eschbach[b]

[a] University of Rochester, Rochester, NY, USA
[b]Xerox Corporation, Webster, NY, USA

ABSTRACT

We present a segmentation-based post-processing method to remove compression artifacts from JPEG compressed document images. JPEG compressed images typically exhibit ringing and blocking artifacts, which can be objectionable to the viewer above certain compression levels. The ringing is more dominant around textual regions while the blocking is more visible in natural images. Despite extensive research, reducing these artifacts in an effective manner still remains challenging. Document images are often segmented for various reasons. As a result, the segmentation information in many instances is available without requiring additional computation. We have developed a low computational cost method to reduce ringing and blocking artifacts for segmented document images. The method assumes the textual parts and pictorial regions in the document have been separated from each other by an automatic segmentation technique. It performs simple image processing techniques to clean out ringing and blocking artifacts from these regions.

Keywords: Ringing artifacts, Blocking artifacts, Document image, JPEG, Post-processing, Document Segmentation

1. INTRODUCTION

Digital image compression is commonly employed in many practical and commercial systems, where storage or transmission of the image is required over limited resources. Compression can be exercised by removing information-theoretic redundancies from the original image such that the original image can be recovered exactly from the compressed image as in *lossless* compression methods or by removing psycho-visual redundancies such that compressed image visually approximates the original image as in *lossy* compression methods. Over the long history of image compression, several standards have been established aiming to accomplish application specific requirements, satisfy computational complexity criterion and visual quality constraints.[1] Despite the increased choices of image compression standards, doubtlessly, *JPEG*,[2] a member of lossy compression standards family, is the most popular image compression standard among all others.

JPEG compression is based on quantization of the Discrete Cosine Transform (DCT) coefficients of each size 8×8 non-overlapping block that tile the image completely by a quantization table followed by a lossless entropy encoder.[3,4] Different compression ratios can be achieved by scaling the elements of the quantization table as a function of user-selectable *quality factor* (Q). As Q increases, more quantization noise is introduced on the DCT coefficients, which makes the quantized coefficients get closer values to each other. Consequently, the entropy coder can achieve higher compression ratios. However, at the same time, this harsh quantization may introduce visible artifacts on the image. These artifacts are typically sorted as:

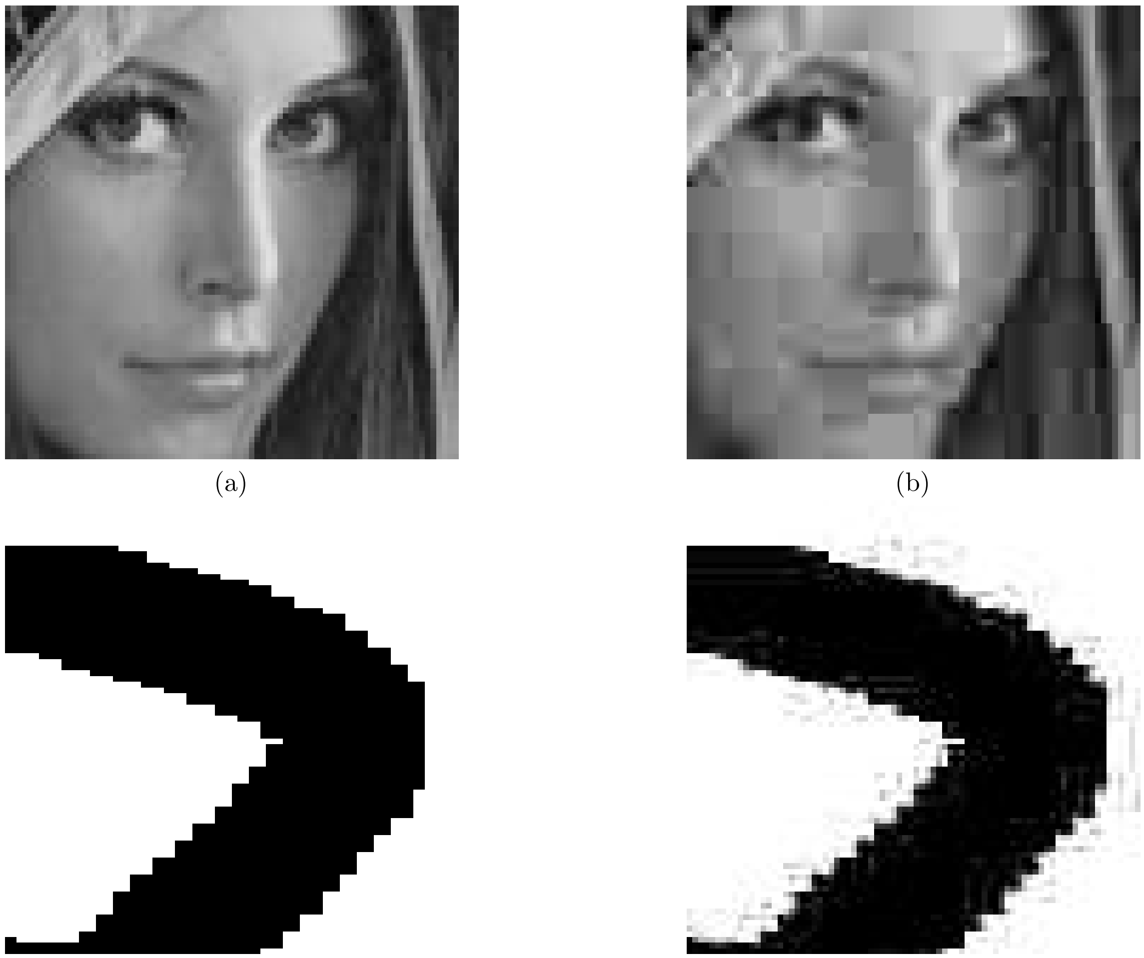
1. *Blocking artifacts*, which are mainly due to the coarse quantization of low-frequency DCT coefficients yielding decompressed image look like a mosaic at smooth regions, and

* B. Oztan was a summer intern at Xerox Corporation, Webster, NY during Summer 2005
Send correspondence to B. Oztan: E-mail: basak@ece.rochester.edu, Telephone: (585) 275-8122, Address: Electrical and Computer Engineering Department, University of Rochester, Rochester, NY, 14627-0126, USA, WWW: www.ece.rochester.edu/~basak

Color Imaging XII: Processing, Hardcopy, and Applications, edited by Reiner Eschbach, Gabriel G. Marcu,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 6493, 649306, © 2007 SPIE-IS&T · 0277-786X/07/$18

2. *Ringing artifacts*, which are mainly due to the coarse quantization of high-frequency DCT coefficients making the decompressed image exhibit noisy patterns known as ringing or mosquito noise near the edges.

Examples of such artifacts are shown in Fig. 1.



(a) (b) (c) (d)

Figure 1. Blocking and ringing artifacts in highly compressed JPEG images. (a) and (c) shows portions of original uncompressed images, (b) and (d) shows the same portions JPEG compressed with $Q = 10$. Blocking artifacts are visible in (b) and ringing artifacts are visible in (d).

Although the primary targets of JPEG compression are natural images, several other types of images such as document images can also be encountered in digital environments. Typical document images contain both textual and pictorial regions. Artifacts can also be observed on these regions. However, unlike natural images, which generally exhibit blocking artifacts at high compression ratios, in document images significant ringing artifacts are observed near the edges of textual components. Fig. 2 shows a magnified portion of an original document image including the word *document* and its JPEG decompressed form obtained by JPEG compressing the image with $Q = 20$. Ringing artifacts can be seen easily on the background region near the edges and visual appearance of the image is degraded. Depending on the application these artifacts can become intolerable and make the document even illegible. A trivial remedy to this problem can be using higher quality factors during the compression, however, high compression ratios can still be desirable for different applications. For instance, storage requirements for a book scanning project can be significantly reduced by utilization of high compression ratio. Similarly, the productivity of an image processing and printing system can be significantly increased by using higher compression ratios to better utilize the available network bandwidth. Thus, an artifact reduction scheme for JPEG decompressed document images is useful to improve visual quality of the image.

(a)

(b)

Figure 2. (a) shows a portion of an uncompressed document image, (b) shows the same portion after decompression.

In this work, we focus on reducing JPEG compression artifacts from the document images. Previously several other algorithms have been proposed to eliminate blocking[5–7] and ringing[8–10] artifacts from JPEG decompressed images. However, reducing these artifacts in an effective manner still remains challenging.

Segmentation is a procedure that partitions an image into multiple regions, each of which is uniform or similar in certain characteristics. Document images are often segmented into regions, such as text, synthetic graphics and natural pictures for various purposes. It could serve as a first step for applications like Optical Character Recognition (*OCR*), or enable object-optimized algorithms and/or parameters in processing for different regions. Segmentation-based compression, color conversion and halftoning techniques have been reported in the literature.[11–13] In this paper, we present a JPEG artifact reduction algorithm based on segmentation. We assume the image has already been segmented into text and picture regions for other purposes and the segmentation information is thus free in terms of computation. To reduce the JPEG artifacts, different strategies are applied to different regions. We propose simple post-processing methods using well-known image processing techniques. Our technique can significantly improve the visual quality of the document at a comparable level to prior techniques, but with a much less computation cost. Several successful document segmentation algorithms[14–17] have been proposed previously in the literature. However, choice of the segmentation technique and its integration to our method is outside the scope of this paper.

In the remainder of this paper, we explain our method in detail in Sec. 2 and show experimental results in Sec. 3. Finally, in Sec. 4, we present conclusions.

2. SEGMENTATION-BASED JPEG ARTIFACT REDUCTION IN DOCUMENT IMAGES

As mentioned in Sec. 1, segmentation of picture text/graphics parts is accomplished before we begin our process. Moreover, the text/graphics parts are also assumed to be segmented into regions, each of which contains a background of uniform color.

We divide our method into two parts. In the first part, removal of the ringing artifacts from text regions is handled. We explore two different scenarios for different levels of compression ratios namely low to moderate and high compression ratios. In the next part, we describe how blocking artifacts are reduced in pictorial regions.

2.1. Reduction of Ringing Artifacts in Text and Graphics Regions

The overview of our method is shown in Fig. 3. We first consider gray-level images. The results will be extended to color images at the end of this subsection. For each textual region, a gray value histogram is first built. Three pieces of information are derived from the histogram, namely, the gray value of the background, a threshold that separates the text and the background, and a Signal-to-Noise Ratio (*SNR*) level for the region.

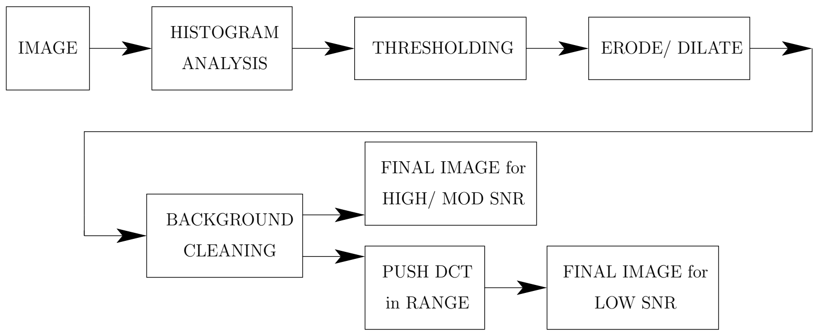


Figure 3. Overview of the method

Since in most text regions, the background pixels are dominant in number, it is easy to determine the background color of the image by either choosing the most frequent gray level or the weighted average of several frequent gray levels as the background color of the image region. From the histogram, we also determine a threshold value that can be used as a metric to assign each pixel as a member of the text or background. Several powerful thresholding algorithms can be found in the literature,[18] which generates a single (global) threshold for the whole image region or several (local) threshold values that vary according to the image content. Generally, local thresholding algorithms outperforms global thresholding techniques. However, since we have the same background formation throughout the image region, a global thresholding algorithm can outperform a local thresholding algorithm in consistency while offering the benefit of being more computationally efficient.

Although many global[19–21] and local[22,23] thresholding algorithms can be applied here, we choose Otsu's document binarization algorithm[19] to determine the optimal threshold value for the image in our experiments. The algorithm maximizes the ratio of between-class variance/within-class variance between the foreground(text) and background pixels. The threshold value can be computed as:

$$T = \arg\max_T R(T) = \arg\max_T \frac{P(T)(1-P(T))(\mu_f - \mu_b)^2}{P(T)\sigma_f^2 + (1-P(T))\sigma_b^2}, \tag{1}$$

where $P(T) = \sum_{i=0}^{T} p(i)$ is the cumulative distribution function of the histogram $p(T)$, $\mu_f = \sum_{i=0}^{T} ip(i)$ and $\mu_b = \sum_{i=T+1}^{255} ip(i)$ are the means, and $\sigma_f^2 = \sum_{i=0}^{T}(i-\mu_f)^2 p(i)$ and $\sigma_b^2 = \sum_{i=T+1}^{255}(i-\mu_b)^2 p(i)$ are the variances of the foreground and background regions, respectively.

The maximized $R(T)$ also indicates how well the two classes (text and background) are separated. A larger $R(T)$ implies a higher SNR level, which translates to a smaller classification error.

Once the threshold is determined, each pixel in the region is classified as a member of text or background. Although ringing occurs in both text and background, practically, it is rarely observable in text. This is because the text objects are typically much smaller in sizes and the ringing is often visually masked by the strong transitions between the text and the background. As a result, cleaning is only performed on the background pixels and not on the text ones. Specifically, the background value obtained in histogram analysis is assigned to all the background pixels in the region, with the exception of the edge pixels. The rationale for the exclusion is twofold. First, due to the masking effect, the ringing on the edge pixels is usually not perceptible. Second, misclassification rate increases at the edge pixels. To prevent cleansing on the edge pixels, simple morphological transforms are applied on the image [24, pp. 523-527] to preserve these pixels. If the foreground color is darker than the background, the image is eroded, otherwise it is dilated such that the textual parts on the thresholded image becomes thicker and after the masking they are still preserved. A 2×2 structuring kernel of 1's is used for erosion/ dilation. Size of this kernel can be increased to obtain stronger erosion/ dilation. On the next step, using the eroded/ dilated image as a mask, the background region of the decompressed image is cleansed accordingly. This particular operation mimics logical **OR** or **AND** operation using the original decompressed image and the mask. The final step is assigning the background value determined in the histogram analysis step to the background pixels of the image. For for low SNR scenarios, thresholding may not always correctly separate background from noise. To prevent major errors being introduced into the cleaning process such as classifying a significant portion of the text as background at high noise levels, we perform a sanity check to see if the DCT coefficients of the cleansed image fall within the original quantization levels. If not, the DCT coefficients are pushed back and restored to the original quantization range. This stage adopted from the method proposed by Eschbach.[25]

While the aforementioned method is described for grayscale images, it can be applied with equal efficacy to color images with a minor change to the method. Typical color images have 3 different color channels. The method can be applied on each of these 3 channels separately. However, this increases the computational requirements by a factor of 3. Instead of using 3 color channels, if the image is transformed into luminance-chrominance space, it can be seen that most of the visible noise is carried in the luminance channel. Essentially, the computational cost can be reduced by a factor of one third by applying the method only on the luminance channel of the image. Artifact removal only in the provides sufficiently good results compared to applying the method on each of the color channels separately.

2.2. Reduction of Blocking Artifacts in Pictorial Regions

In order to eliminate the blocking artifacts from a pictorial region, we first tag the pixels in the region as edge/ non-edge. This can be accomplished by any standard edge detection algorithms. In addition, we also identify the pixels that lie on the 8×8 tile boundary used in JPEG compression. If a priori knowledge about the tiling boundary is not available, it can be determined by a Maximum A Posteriori (MAP)-like estimator.[26] For each non-edge pixel on the tiling boundary, a *sigma filter* is applied to smooth out the blocking artifacts. The sigma filter is an edge preserving smoothing filter. Its output is an average over the pixels within a small window. In calculation of the average, the pixels whose absolute intensity differences with the current pixel exceed a threshold value are excluded. In the case of color images, the same algorithm is applied to the chrominance channels as well.

3. EXPERIMENTAL RESULTS

In this section, we first show effectiveness of our algorithm on eliminating the ringing artifacts of the image shown in Fig. 2.(b). Fig. 4 shows histogram of the image. The background gray level for this image is found as 253, which is the most frequent gray level in the histogram. Using (1), the optimal threshold value for the image is computed as 171. In Fig. 5, the intermediate stage images after thresholding and erosion are shown. The image in Fig. 5.(b) is used as a mask for cleaning out the noise from JPEG decompressed image. In the final step, we assign the background value 253 to the background pixels specified by the mask. Fig. 6.(a) shows the final image. We compare our method against the computationally intensive Projection Onto Convex Sets ($POCS$) based method described by Fan and Eschbach.[8] Fig. 6.(b) shows the decompressed image obtained by using this method. It can be seen that the visual quality of our method is close to that of POCS based method.

In order to evaluate our method's performance against a JPEG compressed color image, we use the image shown in Fig. 7.(a), which was compressed using $Q = 20$. We first transform the image from RGB to YC_bC_r domain and apply the method only on the luminance channel to reduce the noise. The image is then converted into RGB color space and the final image is shown in Fig. 7.(b). Once again the method is successful in reducing the noise level significantly.

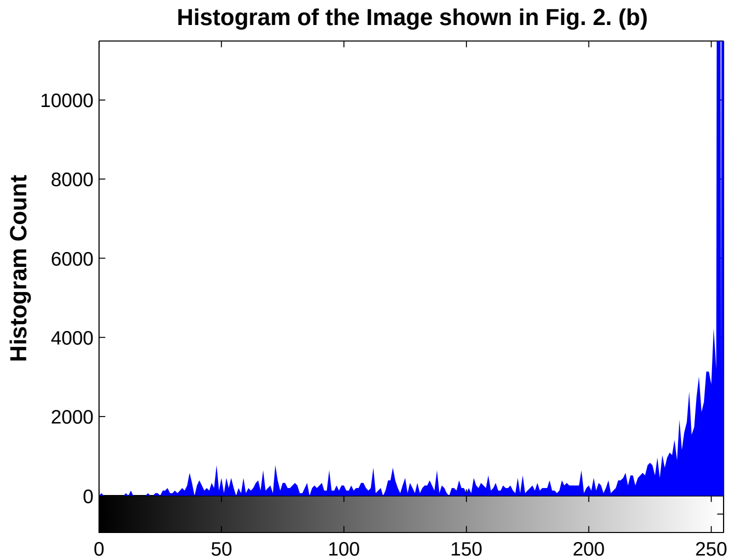


Figure 4. Overview of the method

(a)

(b)

Figure 5. (a) shows a portion of an uncompressed document image, (b) shows the same portion after compressing the image with JPEG $Q = 20$ and decompression.

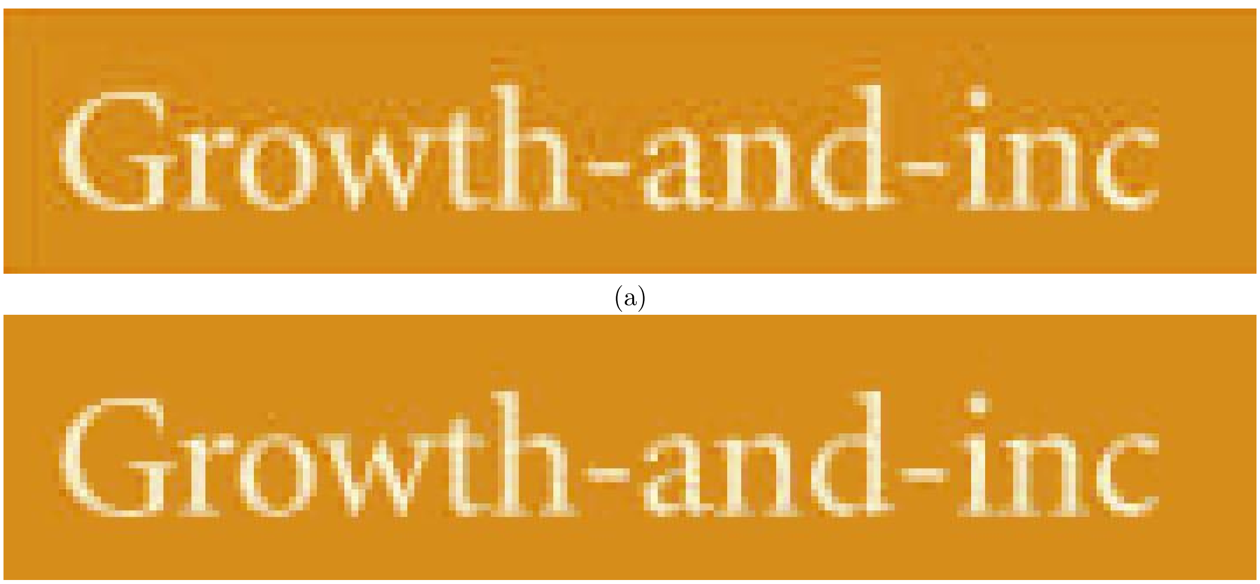
In the next experiment, we test our method's blocking artifact reduction performance for the image shown in Fig. 8.(a). Printed version of this image may not be exhibit the artifacts due to the halftoning, however, they can be observed on the computer screen (by zooming in) using the electronic version of this document. † Fig. 8.(b) shows the post-processed image using the method described in Sec. 2.2. The blocking artifacts are significantly reduced while the edge detail is still retained.

†Electronic version of this document can be accessed from http://www.spiedl.org

(a)

(b)

Figure 6. Final images obtained. (a) shows the result obtained by our method, (b) shows the result obtained by POCS based method.



(a)

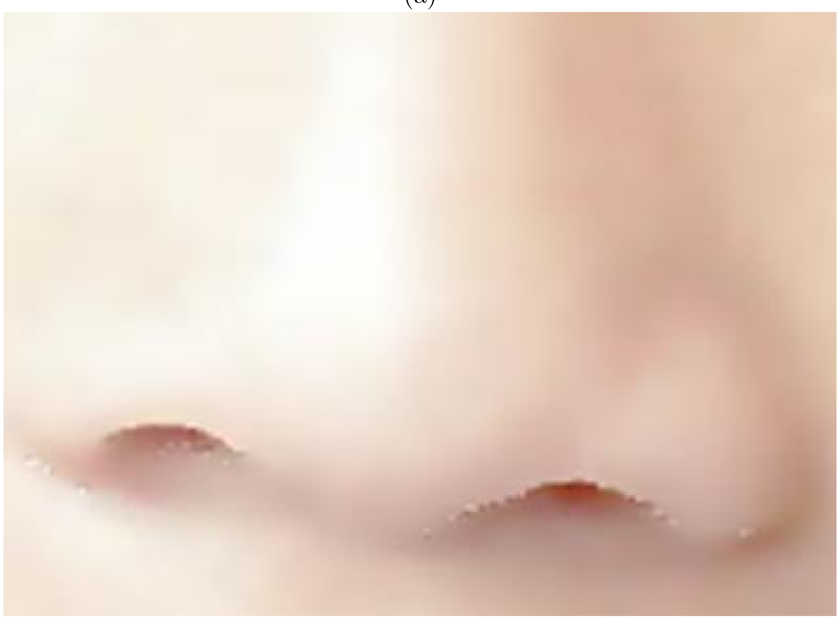
(b)

Figure 7. (a) shows a portion of the noisy color document image, (b) shows the same portion after post-processing the image using our method.

4. CONCLUSIONS

We present a simple post-processing method to reduce compression artifacts in JPEG compressed document images. The method is suitable for the cases where the document image has already been segmented into text and picture regions. The method can significantly reduce the artifacts with simple computation. The technique can be applied both to grayscale and color images.

(a)



(b)

Figure 8. (a) shows a portion of a JPEG decompressed image compressed with $Q = 20$ with visible blocking artifacts, (b) shows the same portion after blocking artifact reduction.

REFERENCES

1. M. Rabbani and P. W. Jones, *Digital Image Compression Techniques*, SPIE, Bellingham, WA, 1991.
2. W. B. Pennebaker and J. L. Mitchell, *JPEG Still Image data Compression Standard*, Van Nostrand Reinhold, New York, NY, 1993.
3. D. A. Huffman, "A method for the construction of minimum redundancy codes," *Proc. IRE* **40**, pp. 1098–1101, 1952.

4. P. G. Howard and J. S. Vitter, "Analysis of arithmetic coding for compression," *Information Processing and Management* **28**, pp. 749–763, Nov. 1991.
5. H. Reeve and J. Lim, "Reduction of blocking effects in image coding," *Optical Engineering* **23**, pp. 34–37, January/February 1984.
6. A. Zakhor, "Iterative procedures for reduction of blocking effects in transform image coding," *IEEE Trans. Circuits and Syst. Video Tech.* **2**, pp. 91–95, Mar. 1992.
7. T. Meier and K. N. Ngan, "Reduction of blocking artifacts in image and video coding," *IEEE Trans. Circuits and Syst. Video Tech.* **9**, pp. 490–500, Apr. 1999.
8. Z. Fan and R. Eschbach, "JPEG decompression with reduced artifacts," in *Proc. SPIE and IS& T Symposium on Electronic Imaging: Image and Video Compression*, **2186**, pp. 50–55, May 1994.
9. R. Eschbach, "Conditional post-processing of JPEG compressed images," in *IS&T's 50th Annual Conference*, **50**, pp. 443–445, May 1997.
10. Z. Fan and F. Li, "Reducing artifacts in JPEG decompression by segmentation and smoothing," in *Proc. IEEE Intl. Conf. Image Proc.*, **1**, pp. 17–20, Sept. 1996.
11. R. L. de Queiroz, R. Buckley, and M. Xu, "Mixed raster content (MRC) model for compound image compression," in *Proc. SPIE: Visual communications and image processing*, K. Aizawa, R. L. Stevenson, and Y. Zhang, eds., pp. 1106–1117, Jan. 1999.
12. H. Cheng and C. A. Bouman, "Document compression using rate-distortion optimized segmentation," *J. Electronic Imaging* **10**, pp. 460–474, Apr. 2001.
13. S. Prabhakar, H. Cheng, J. C. Handley, Z. Fan, and Y. Lin, "Picture-graphics color image classification," in *Proc. IEEE Intl. Conf. Image Proc.*, **2**, pp. 785–788.
14. R. L. de Querioz and R. Eschbach, "Fast segmentation of the JPEG compressed documents," *J. Electronic Imaging* **7**, pp. 367–377, Apr. 1998.
15. A. K. Jain and B. Yu, "Document representation and its application to page decomposition," *IEEE Trans. Pattern Anal. Mach. Intel.* **20**, pp. 294–308, Mar. 1998.
16. H. Cheng and Z. Fan, "Background identification based segmentation and multilayer tree based representation of document images," in *Proc. IEEE Intl. Conf. Image Proc.*, **3**, pp. 1005–1008.
17. Z. Fan and T. Jacobs, "Segmentation for mixed raster contents with multiple extracted constant color areas," in *Proc. SPIE: Color Imaging X: Processing, Hardcopy, and Applications*, R. Eschbach and G. G. Marcu, eds., pp. 251–262, Jan. 2005.
18. M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *J. Electronic Imaging* **13**, pp. 146–165, Jan. 2004.
19. N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Sys., Man, and Cyber.* **9**, pp. 225–236, Jan. 1979.
20. J. Kittler and J. Illingworth, "Minimum error thresholding," *Patt. Recogn.* **19**, pp. 41–47, Jan. 1986.
21. P. K. S. J. N. Kapur and A. K. C. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," *Comp. Vis., Graphics and Image Proc.* **29**(3), pp. 273–285, 1985.
22. W. Niblack, "An introduction to image processing." Prentice- Hall, Englewood Cliffs, NJ, 1986.
23. J. Sauvola and M. Pietikäinen, "Adaptive document image binarization," *Patt. Recogn.* **33**, pp. 225–236, 2000.
24. R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice-Hall, Upper Saddle River, NJ, second ed., 2002.
25. R. Eschbach, "Improved decompression of standard adct compressed images." United States Patent No. 5 379 122, 1995.
26. Z. Fan and R. L. Queiroz, "Identification of bitmap compression history: JPEG detection and quantizer estimation," *IEEE Trans. Image Proc.* **12**, pp. 230–235, Feb. 2003.