Abstract—In this work, a new approach is proposed that deals with blocking effects in JPEG compressed images. High-frequency details of the coded images are mainly contaminated by quantization noise. Preserving the image details and reducing the effect of quantization noise as much as possible can improve the ability of any enhancing method. To achieve this goal along with the removal of the blocking effect, the high-frequency components of the image are first extracted by high pass filtering. The result is then scaled by a factor that depends on the compression ratio and subtracted from the observed image. This result is used to design an adaptive filter that depends on the statistical behavior of the preprocessed image. The adaptive filter is applied to the resultant image. The result shows high SNR, significant improvement in the separation between blocking noise and image features, and effective reduction of image blurring. Other steps are required to preserve the global and local edges of the processed image, remove blocking noise, and ensure smoothness without blurring. These steps are dedicated to remove blocking artifacts and to enhance feature regularities. The evaluation of this approach in comparison with other techniques is carried out both subjectively and qualitatively.

Index Terms—Adaptive filtering, deblocking, edge enhancement, image restoration, JPEG coding, statistical analysis.

I. INTRODUCTION

T HE discrete cosine transform (DCT) is the integrated part of the basic compression algorithm in JPEG. The importance of this compression technique can be utilized due to its performance that matches the Karhunen-Loeve Transform (KLT), which is known to be optimal in the mean square error sense. Although this is the most popular compression approach, its main drawback is what is usually referred to as the “blocking effect.” Dividing the image into small blocks prior to coding causes discontinuities between adjacent blocks and affects the strong edges in the image. In this approach, correlation between adjacent blocks is not exploited. Owing to the quantization errors generated in different blocks, reconstructed images exhibit severe artifacts at block boundaries for high compression ratios. The blocking effect in JPEG images can be characterized into five kinds as follows. Stair case noise along the image edges, grid noise in the monotone areas, corner outliers in the corner points of the $8 \times 8$ DCT blocks, ringing artifacts near the strong edges, and corruption of edges across block boundaries [2], [9], [11], [16], [22], [23].

Due to the huge data requirements for multimedia, the attention is focused toward getting more compression and less visual defects. To remove the blocking effect, several deblocking techniques have been proposed in the literature as post process mechanisms after JPEG compression, depending on the perspective from which the deblocking problem is dealt with. The easiest way of looking at this problem is to low-pass the blocky JPEG image [1]. This approach will reduce the effect of high-frequency tendency but the image will be blurry and some details will be wiped out. Going further step in complexity and applying a simple nonlinear smoothing to the pixels will add another obstacle to the solution [2]. A more sophisticated approach involves segmentation and smoothing that will reduce the ringing artifacts due to sharp variations [3]. Due to the importance of high-frequency details that may be overwhelmed with blocking noise, another restoration process that uses linear predictors to estimate the corresponding boundary pixels in the regions of interest is devised [4]. The attempt in this method is to classify small local boundary regions according to their intensity distribution and to employ this information in designing proper predictors. Despite the importance of this approach, the blocking effect at high compression rates not only is resulted by block boundaries; moreover, images with sharp variations cannot be easily configured with low order predictive filters.

To improve the accuracy of the classified patterns an iterative method for block removal using block classification and space-frequency filtering is proposed [5].

Because of the defined location of the blocking noise within the image area, methods that employ both edge detection techniques for details classification and spatial adaptive filtering for smoothing have the potential for solving this problem [6]–[9]. The first serious attempt in this direction is proposed in [6], where in this algorithm an efficient way of eliminating both grid noise and staircase noise has been proposed. The major drawbacks of this method are the effect of speckle noise on the image gradient and the limitation in using the directional filter for only horizontal and vertical directions. In this case diagonal edges are not properly handled. A more computationally involved approach combines most of the previously mentioned techniques in a selective way in order to design an estimated spectrum adaptive post-filter (ESAP) [7]. This filter estimates 2-D pixel-adaptive bandwidths directly from the de-quantized DCT coefficients. It then combines this information with directional Canny edge detectors to control a 2-D spatially adaptive
nonlinear post-filter to reduce the DCT blocking artifacts. Based on the HVS tolerance to quantization errors in the high-frequency regions, ESAP performs directional filtering parallel to the edges and no filtering across the edges, subject to filter design constraints. This method adaptively filters low frequency regions to minimize blocking and restores the natural smoothness of the image. The nonedges are post-filtered with separable nondirectional adaptive low-pass filters while the edges are post-filtered with nonseparable directional low-pass filters. Although different aspects of the blocking effect are elegantly exploited, the proposed approach of [7] inherits the following problems: it requires a very expensive computational procedure, it relies on a DCT analysis of the DCT of the image to determine the adaptive filter bandwidth, and it does not handle the different types of noise that contribute to blocking effect such as ringing artifacts and edge corruptions. The ideas of [6] and adaptive filtering scheme of [7] are thoroughly discussed in [8].

In [8], a less expensive edge detector that has smoothing capabilities is used to determine global and local features of the image. Then (1-D) directional filter, and (2-D) adaptive filter are used to preserve natural edges and remove blockage. Although the work of [8] was able to address some of the defects of the previously mentioned approaches some points are still to be resolved. Applying two filtering steps (1-D directional filtering and 2-D adaptive filtering) without compensation results in a less detailed image. Using fixed threshold value focuses the use of this approach for specific types of images (less robust performance). The smoothing and noise removal improvements are limited within and on the borders of the block where the correlation among adjacent blocks is not subjugated, this will reduce the ability of the algorithm from restoring edges corrupted by blocking noise and ringing artifacts. Therefore, the decoded images using this approach that don’t share the same statistical characteristics will appear either blurry or heavily smoothed with some reduction in its SNR. Advances in Internet accessibility and network speed boost the importance of using MPEG for video processing. Since MPEG inherits both blocking noise and quantization noise from JPEG, similar ideas of [6]–[8] are utilized in [9], [10] as a solution for these problems. The use of (1-D) filter in the horizontal and vertical directions of the block boundaries does not solve the ringing artifacts that corrupt diagonal edges. Also the use of (2-D) signal adaptive filter within the block will not preserve the continuity of edges between consecutive blocks, which may add another form of image artifacts.

The theory of projections onto convex sets (POCS) provides another avenue for solving the blocking effect; where several research groups have effectively implemented it [11]–[14]. The idea there is to constrain the distribution of the DCT coefficients of the decoded image after smoothing to be within the quantization levels. Although the improvements that these approaches add to the decoded images are noticeable, they don’t resolve the interactions of different types of noise. Moreover, two main issues are not fully explored in these approaches: the choice of the smoothing filter and the existence of a convergence criterion. In the case of well-defined convergence condition, the cost of the iterative procedure is not affordable.

The features of the wavelet theory add another tool of exploration to the blocking problem; several ideas based on soft threshold approach in the wavelet domain are successfully implemented for deblinking JPEG coded images [15]–[19]. The core of these techniques is to make use of Donho’s algorithm [20] for denoising Gaussian noise and modifying the denoised wavelet coefficients to remove the effect of other types of noise. Direct application of Donho’s algorithm for removing blocking effect can be found in [15]. In [16] a further step has been adopted by presenting a simple and efficient denoising algorithm that exploits correlations among cross-scale wavelet coefficients to extract edge information and protect these edges during threshold operation. The importance of edge information provides another method for enhancement. The effect of blocking in wavelet domain is investigated in [17]. Regarding to the observations in [17], the blocking effect is characterized as one of the following:

1) small modulus maxima at block boundaries over smooth regions;
2) noises or irregular structures near strong edges;
3) corrupted edges across block boundaries.

To handle these effects, first a segmentation procedure is performed to discriminate the texture regions of an image based on local regularity variance and keeping regions that have low regularity variation. Then singularities on the remaining regions and small modulus maxima at block boundaries are removed. Due to threshold operation, a simple linkage procedure is adopted for correcting corrupted edges using the phase of modulus maxima as well as the amplitude of strong edges. Finally, the image is reconstructed using the projection onto convex set (POCS) technique on the processed WTMM of the JPEG decoded image. However, the WTMM deblocking algorithm is an iterative algorithm that requires a long computation time to reconstruct the processed WTMM to obtain the deblinked image. In [18], the singularity of an image is detected by computing sums of the wavelet coefficients inside the so-called ‘directional cone of influence’ in different scales of the image. In this way the edge and smooth regions of an image are effectively identified irrespective of the discontinuities introduced by the blocking effect. Besides, advantage over WTMM, a simple inverse wavelet transform is required to reconstruct the processed wavelet coefficients to obtain the deblinked image. Another simple algorithm is found in [19] where the artifact reduction operation is applied to only the neighborhood of each block boundary in the wavelet transform on the first and second scales. The blocking effect is distinguished as a stepwise discontinuity at block boundaries. Each block boundary is classified into one of the shade region, the smooth edge region, and the step edge region. Step edge regions are removed based on an image dependent threshold values.

The significance of deblocking in improving the visual quality of the decoded image motivates other approaches to be considered. A pre-processing technique that classifies image information based on their frequency distribution and design of an adaptive coding based on a subjective criterion is proposed in [21]. A further step in this direction is shown in [22] in which an optimal decoding scheme that takes into account observed data, a priori knowledge on the solution, and precise spatial location of blocking artifacts are considered. A robust
algorithm that uses correlation between the intensity values of boundary pixels of two neighboring blocks is instantiated by [23]. Specifically, authors realize the fact that quantization of the DCT coefficients of two neighboring blocks increases the expected value of the mean squared difference of slope (MSDS) between the slope across two adjacent blocks, and the average between the boundary slopes of each of the two blocks.

In general, most of the aforementioned techniques that use different methodologies are based on identifying image details by segmentation and smoothing image textures by some sort of low-pass filtering. In all the aforementioned techniques, regardless of the accuracy of the applied segmentation procedure, the crispiness of the image are either completely destroyed or deteriorated. In this work, we propose a combination of crispiness and statistically adaptive-deblinking technique that is geared toward preserving image texture crispiness and dissolving the effect of blocking on edge details. This approach is based on modeling the degradation of the decoded image into additive and multiplicative noise as well as estimation of the image local statistics for each block. The blocking effect in this approach is significantly diminished and smoothed images are produced. Fig. 1 illustrates the major steps involved in the proposed post-processing algorithm.

The paper is organized as follows. In Section II, steps that are involved in combined crispiness and statistical differencing approach for noise removal are explained. In Section III, global and local feature enhancement techniques are illustrated. In Section IV, several computer simulation results are introduced to make a comparative study between the proposed approach and projection onto convex sets (POCS), directional filtering supported with both adaptive restoration and outlier detection, and wavelet deblinking approaches. Finally, in Section IV conclusions are given.

II. COMBINED CRISPINESS AND STATISTICAL DIFFERENCING FOR NOISE REMOVAL IN JPEG CODED IMAGES

Image crispiness can be defined as those textures that have a frequency span not only on the low frequency side of the spectrum but on the high-frequency domain as well. To maintain image crispiness both smooth features and edge textures are to be preserved. Any compression scheme depends in nature on the assumption that the low frequency band has the highest energy and can be easily coded with less number of bits. For efficient compression high-frequency details are either use heavy threshold or they are completely destroyed. In either case the image crispness is not maintained. The goal of this work is to preserve as much as possible the crispiness of the image by reducing the effect of compression on high-frequency details and differentiating image crispiness from blocking noise. To satisfy the minimum crispiness requirement, our algorithm is geared toward extracting the high-frequency details of the observed image that are overwhelmed by the quantization noise. The algorithm is based on the following observations: firstly, image crispiness is not limited to one side of the spectrum and so feature continuity is a main requirement for determining image crispiness. Secondly, quantization noise can be modeled as additive noise. Although this model is valid in the transform domain; i.e., for the DCT coefficients, it will be a strong assumption if we have an *a priori* knowledge about the noise distribution. Since such distribution is definitely non-Gaussian the assumption is statistically weak. The quantization noise is correlated within each $8 \times 8$ block and independent from one block to another. Since inverse DCT is a linear operator, this particular noise remains additive in the spatial domain as well. Due to the dependency of the DCT coefficients from block to block it can be statistically modeled as a multiplicative portion of a generalized Gaussian distribution. To clarify both observations and assumptions let the original image be represented by $X$ and the decoded image by $Y$. Then, the decoded image is modeled by $Y = X + N$, where $N$ represents the effect of quantization noise.

Recovering actual image details that are hidden within the quantization and blocking noise can achieve major improvement in the enhancing process. To facilitate achieving this goal, the observed image is first high pass filtered by filter $\hat{W}$, this step is instantiated for two reasons to focus the high-frequency data of the observed image and to smooth the high-frequency points within the kernel transfer function. In this case, the characteristics of the filtered image are not a rough estimate of the high-frequency spectrum of the decoded image. The filtered image is then scaled by a factor $k$ that depends on the compression ratio of the quantization table and the percentage of image details. A statistical differencing approach is used that counters the effect of the statistical differences between the observed image and the scaled version of the filtered image. In this case, the subtraction will occur for those data points that are severely affected by the compression process while the image crispiness that propagates in the high-frequency domain will be smoothed rather than decayed. The subtraction will not remove image details; instead it will amplify their existence by removing some of the distortions that were inherited from both compression and quantization. Since exact locations of image features are not deteriorated, the modeling process combined with adaptive filtering and proper linkage process will assure their continuity.

In this stage, the processed image can be represented by

$$Z = Y - kWY$$

and it is related to the original image by

$$Z = (I - kW)X + (I - kW)N = EX + \bar{N}$$

where $E$ represents a transformation kernel that affects both the initially coded image $X$ and the additive noise $N$; $\bar{N}$ is the trans-
formed noise. The transformation effect on each pixel of the initially coded image can be viewed as follows:

\[ z = \zeta x + \eta \]  
\[ (3) \]
where \( z \) is the processed pixel and \( x \) is its corresponding pixel related to the initially coded image \( X \). In this formulation, it is assumed that \( \zeta \) is a random variable with unity mean and variance \( \sigma^2 \), where the variance depends on the correlation from block to block, and the effect of compression on the scaled version of the filtered image. Increasing the compression rate means \( k \) should be high that means great distortion affects the coded data and so higher values of the data are considered noise and so it cannot be considered as smoothed features that propagate from low to high frequency. Hence; the condition for crispiness of features is not satisfied and so it should be subtracted.

On the other hand, at low compression rate, the data are less distorted and so low value of \( k \) factor is required. This means that the DCT coefficients are less correlated and so a low variance is expected. Regarding this discussion, using the statistical differencing approach will make the initially weak additive assumption a reasonable approximation to use. The additive part of (3) \( \eta \) can be modeled as a random variable with zero mean and variance \( \sigma_a^2 \). The previous assumptions are achieved from the fact that the processed image \( Z \) should be equal to \( X \) if there are no compression \((k=0)\) and no quantization noise. In this approach, the initially preprocessed image is therefore modeled as an original image corrupted with both multiplicative and additive noise. The purpose of the multiplicative term is to ensure that \( k \) depends on the actual characteristics of the decoded image. The problem now is converted into estimation problem based on the model of (3). With the first order constrain, the estimate of the actual pixel \( x \) from the observed pixel \( z \) is

\[ \hat{x} = \alpha z + \beta. \]  
\[ (4) \]
By using the minimum mean-square-error, \( \alpha \) and \( \beta \) are determined by minimizing

\[ e^2 = E[(\hat{x} - x)^2] = E[((\alpha \zeta - 1)x + (\alpha \eta + \beta)]^2).\]  
\[ (5) \]
The result of this minimization is

\[ \begin{align*}
\alpha &= \frac{\sigma^2_X}{\sigma^2_X + \sigma^2_a + \sigma^2 \left( \frac{\sigma^2_X}{\mu^2_X} \right)} \\
\beta &= (1 - \alpha) \mu_X
\end{align*} \]  
\[ (6) \]
This estimate makes \( \hat{x} \) a convex combination of the preprocessed pixel value \( z \) and the mean of the original image \( \mu_X \) as follows.

\[ \hat{x} = \frac{\sigma^2_X}{\sigma^2_X + \sigma^2_a + \sigma^2 \left( \frac{\sigma^2_X}{\mu^2_X} \right)} z + \frac{\sigma^2_a + \sigma^2 \left( \frac{\sigma^2_X}{\mu^2_X} \right)}{\sigma^2_X + \sigma^2_a + \sigma^2 \left( \frac{\sigma^2_X}{\mu^2_X} \right)} \mu_X. \]  
\[ (7) \]
For each pixel, by using a given neighborhood region, the local mean and the variance are estimated as follows:

\[ \begin{align*}
\bar{\mu}_{Y} &\approx \bar{\mu}_Y \approx \frac{1}{NM} \sum_{[n_1,n_2] \in \text{window}} Y(n_1,n_2) \\
\sigma^2_{Y} &\approx \sigma^2_Y \approx \frac{1}{NM-1} \sum_{[n_1,n_2] \in \text{window}} [Y(n_1,n_2) - \bar{\mu}_Y]^2.
\end{align*} \]  
\[ (8) \]
In (8), \([n_1,n_2] \in \text{window}\) represents the \( M \times N \) neighborhood of the pixel under consideration. In this work, it is found that the \( 3 \times 3 \) neighborhood works well. The additive noise variance in (7) is estimated by using the average of the noise variance, in (8), within each neighborhood

\[ \sigma^2_a = \frac{1}{NM} \sum_{[n_1,n_2] \in \text{window}} \sigma^2_z(n_1,n_2). \]  
\[ (9) \]
If we assume that the estimate of the additive noise variance \( \sigma^2_a \) is close to the quantization noise, then the estimate of the multiplicative noise variance can be obtained from the high pass filtered image \( Y_f \) and the scaling factor \( k \) as follows:

\[ \sigma^2 = \frac{k^2 \sigma^2_{Y_f}}{\bar{\mu}^2_X + \bar{\mu}^2_X}. \]  
\[ (10) \]
Otherwise, by assuming \( \sigma^2_X = \sigma^2_Y \) the estimate for \( \sigma^2 \) becomes

\[ \sigma^2 = \frac{k^2 \sigma^2_a - \sigma^2}{(\bar{\mu}^2_X + \bar{\mu}^2_X)}. \]  
\[ (11) \]
Both estimates in (9) and (10) or (11) are obtained based on the actual data from the encoded image. The major source of error in these estimates is the assumption that the original image statistics are the same as the JPEG encoded image. The first estimate of \( \sigma^2 \) in (10) gives better results than the second one in (11) because in the first case the estimate is obtained by using the filtered image. The overall result at this stage has higher PSNR than other methods because all the parameters are adaptively estimated based on the local image characteristics. Although this process is based on minimum mean-square-error, the overall approach is not a globally optimum procedure. This is due to the linear constraint used in (4) and using rough estimates of the unknown parameters and nondeterministic behavior of the noise.

In this stage, however, noise features are somewhat distinguishable from image features and further processing will lead to the smoothness and removal of the blocking effect.

III. GLOBAL AND LOCAL EDGE FEATURE ENHANCEMENT

The first step of the blocking removal approach was directed toward minimizing the effect of the additive portion of the noise and part of the multiplicative portion of the quantization noise. When blocking noise contaminates the crispiness of the image features, the statistical differencing step provides a mechanism for distinguishing that effect without trying to prune both image features and noise. The major goal there is to differentiate statistically the noise pattern from the actual image details by maintaining the smoothness constraint of the image features. To maximize the noise removal behavior of the algorithm further steps
are required, where each step will target specific type of noise. Our approach, here, will focus on using the estimated image for further processing steps that yield global and local image features, due to smoothing constraint; image intensities at features locations are filtered to alleviate the contamination effect of the blocking noise.

A. Gradient Amplitude and Phase

In order to determine the global image features and the local image details, an estimate of the image derivative in both dimensions is necessary. Several methods can be used to achieve the image gradient, and Sobel operator is the less expensive edge detector that provides information about pixel’s contrast and directivity. Sobel operator not only gets the derivative of the image but also smooths the image in horizontal and vertical directions reducing the effect of random noise from corrupting the estimate of edge contrast. The gradient of the image is obtained by applying the Sobel kernels to the estimated image ($\hat{X}$) in both directions.

$$\begin{align*}
G_h(r,c) &= \hat{X}(r,c) \ast H_h(r,c) \\
G_v(r,c) &= \hat{X}(r,c) \ast H_v(r,c).
\end{align*}$$  \hspace{1cm} (12)

In (12), $\hat{X}$ is the processed image from the statistical differencing stage, $\ast$ is the linear convolution operator, and $H_h(r,c), H_v(r,c)$ are Sobel kernels in horizontal and vertical directions. The amplitude and the phase of the gradient are obtained by

$$\begin{align*}
G_A(r,c) &= \sqrt{(G_h^2(r,c) + G_v^2(r,c))} \\
G_p(r,c) &= \tan^{-1}\left(\frac{G_v(r,c)}{G_h(r,c)}\right).
\end{align*}$$  \hspace{1cm} (13)

Gradient edge operators generally suffer from two major problems; the first problem is the production of noisy results [24]. The importance of using the adaptive statistical differencing reduces the hazards of such problem because that step dramatically reduces the effect of noise on less representative image details. The other problem with gradient edge operators is that they may produce thick edges. In this case both gradient and direction are required to estimate the feature location.

B. Global and Local Edge Detection

In order to classify the image crispiness, Sobel edge detector is used to differentiate between an edge area and a monotone area. Using a given threshold on the amplitude of the gradient generates an edge map. If the amplitude is greater than the threshold an edge is detected otherwise it is assumed as a monotone area. Two different edge maps are obtained by using the edge detector. The first one is the global edge map that depends on the overall image characteristics. The second one is the local edge map that depends on the pixel characteristics within each $8 \times 8$ block of the image. The global edge map is obtained by using the global threshold $T_g$ on the gradient. The global threshold value is obtained empirically as follows.

$$T_g = 2 \left( \frac{\sum \sum G_A(r,c)}{N_r N_c} \right)$$  \hspace{1cm} (14)

where $[N_r, N_c]$ is the image size. The importance of adapting the threshold to image characteristics is crucial in this stage. The use of fixed threshold level could lead to destructive effects where both global and local edge maps will be used in an adaptive way to differentiate between different types of details. Although the use of an adaptive threshold that depends on probabilistic distribution of the gradients is more convincing, the use of the aforementioned empirical formula shows comparable results on the global level and better results on the local one.

Local edge map threshold is obtained by using the block statistics and the global threshold value as follows:

$$T_n = \left(1 - \frac{\sigma_n}{m_n}\right) \cdot T_g$$  \hspace{1cm} (15)

where $\sigma_n$ and $m_n$, are the standard deviation and the mean of the $n$th block of the gradient image. This means that for a monotone block $\sigma_n$ will be zero and so the block is processed as a global image block. While if the block has a local edge, then the value of $\sigma_n$ will be greater than zero. This will result in a threshold value that is less than $T_g$ and hence producing a detailed local map that has not been detected as a global edge. To preserve the block details, only the internal $6 \times 6$ pixels within each block are used. This approach will relatively reduce the effect of the distortion on the block borders and will protect the detailed edges from blurring. In the case where the block variations are significantly high ($\sigma_n > m_n$), then the local edge threshold value will be negative indicating that the processed pixel is definitely an edge pixel. Practically such situation is less probable and if it occurs it is properly handled by (15).

C. Minimizing Grid Noise and Monotone Noise

Once global and local edge maps are generated, locations of both local and edge details can be used to minimize a specific type of noise related to the blocking effect.

An adaptive filter is used to reduce the effect of both the grid noise in the relatively monotone area and the staircase noise along the image edges without affecting the image details. The filtering process is divided into two major steps.

1) Directional filtering: A 1-D directional filter aligned parallel to the edge is ideally suited for the reduction of staircase noise, which is visible along the image edges [8]. This operation is performed on all points on the global edge map to reduce the staircase noise. Due to using Sobel kernels a 3-coefficient directional filter is used to smooth the edges in eight directions $0, \pm 45, \pm 135$ and $180$. The filter is a weighted average filter that gives the highest weight to the processed pixel and aims at reducing the effect of random noise on strong edges. Giving a reasonably high value to the preprocessed pixel is instantiated due to the statistical differencing step where strong image features are more pronounced and less affected by noise. The directional filter that is used in this work is $[1/6, 4/6, 1/6]$, which is the same filter used in [8]. However, using other 3-coefficients filter will produce the same results keeping in mind that the processed pixel is weighted higher than the other two pixels.
The effective direction is obtained from shifting the gradient angle \( G_e(r,c) \) by \( 90^\circ \). To obtain the edge direction, \( G_e(r,c) \)’s are quantized into steps of \( 45^\circ \) and shifted by \( 90^\circ \). The filtering process is performed only on the pixels of the estimated image that has a location on the global map. This process will enhance the pixels on the global edges and will remove the noise that affects these pixels.

2) 2-D adaptive low-pass filtering: Due to the distribution of blocking noise in the high-frequency portion of the image spectrum, supervised smoothing can significantly reduce its contribution [23]. In [7], [8] a 2-D adaptive approach aiming at reducing the grid noise has been proposed. In this step a 2-D adaptive low-pass filtering procedure similar in principle to the one used by [8] is used on the estimated image to reduce the effect of grid noise as follows.

a) A \( 5 \times 5 \) low-pass filter kernel is used.

b) If the center of the kernel is on a global or local edge point, no low-pass filtering is performed. The processed pixel represents an edge pixel that means the existence of two statistically different regions and so any further smoothing will blur the corresponding edges.

c) If there is no edge point (global or local) in the window, a definitely monotone region is detected and so simple averaging can be performed. The importance of this step is determined by the achievements of the statistical differencing step where great portion of blocking noise is diminished and if there is still some residuals left in monotone regions further averaging will filter them. The averaging process is performed by using

\[
\text{Kernel}_1 = \begin{bmatrix}
0 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 \\
0 & 1 & 1 & 1 & 0
\end{bmatrix},
\]

(16)

d) If any edge is located in the kernel window and not in its center, a masking filter is used as follows. If the edge is on one of the eight pixels around the center of the kernel, the weights of the kernel of the edge pixel and its neighbors are set to zero. The weights of the kernel are changed according to the position of the edge pixel in the filter window to protect image details from corruption and to maintain details continuity.

D. Corner Outlier Detection and Replacement [27]

The corner outlier is characterized by a pixel that has either much larger or much smaller value than its neighbors. Corner outlier is a possible defect in decoded JPEG images because of the application of sequential blocking in DCT where adjacent blocks are decoded independently. The detection scheme of the outlier pixels is performed by applying a \( 2 \times 2 \) window on the cross-point of the block boundaries in the smoothed image. After detecting the edge corner, the pixel is replaced by a weighted-average of the four pixels in the window, giving the edge pixel half of the weight. The threshold value that determines the edge is 20% of \( T_2 \) in (14).

E. Ringing Artifacts Near Edges and Corruption of Edges

The ringing artifact that is resulted from steep edges and edge corruption due to global threshold approach is another problem to overcome. This artifact is considered exasperating by most of the considered subjects. In this case, our goal is aiming at extracting the correlation between consecutive pixels in the global or local edge map and using this information to predict the continuity of the feature and to classify image crispiness from corrupted noise. Smoothing procedure of previous step eliminates the effect of blocking artifacts at internal portion of the image \( 8 \times 8 \) blocks and so the only region that needs to be studied are the boundaries between consecutive \( 8 \times 8 \) blocks. Therefore, edge pixels on the boundaries of \( 8 \times 8 \) blocks can be modeled by

\[
c(r,c) = \alpha r + \beta c + \delta rc + \gamma + \mu(r,c)
\]

(17)

where

- \( c(r,c) \): edge pixel of the smoothed estimated image;
- \( \mu \): noise that corrupts the edge pixel and because of the several processing steps on the decoded image can be considered as an additive noise with zero mean and standard deviation \( \sigma_e \);
- \( r,c \): \( [r,c] \in M + 1 \times M + 1 \) window centered at \( (r,c) = (0,0) \).

The least square solutions for estimated parameters in (17); i.e., \( \alpha, \beta, \delta, \gamma \), that minimize the sum of the square of error between the modeled pixel \( c(r,c) \) and the smoothed estimated one are given by

\[
\begin{align*}
\hat{\alpha} &= \frac{\sum r e(r,c)}{\sum r^2} \\
\hat{\beta} &= \frac{\sum c e(r,c)}{\sum c^2} \\
\hat{\delta} &= \frac{\sum r c e(r,c)}{\sum r^2 c^2} \\
\hat{\gamma} &= \frac{\sum e(r,c)}{\sum 1}
\end{align*}
\]

(18)

Parameter \( \hat{\gamma} \) represents the fitted value of the center pixel within the \([2M+1, 2M+1] \) neighborhood. Equation (18) shows that the estimated parameters are dependent on the noise and they are unbiased estimators for \( \alpha, \beta, \delta, \gamma \). The assumption of normally distributed noise implies that the estimated parameters are normally distributed. Using these results, the estimation square error divided by noise variance \( \frac{\varepsilon^2}{\sigma^2} \) is distributed as chi-squared random variable with \( (M+1)^2 - 4 \) degrees of freedom. The estimation square error is determined by

\[
\varepsilon^2 = \sum_{r} \sum_c [c(r,c) - \hat{X}_{\text{smooth}}(r,c)]^2.
\]

(19)
The detection and estimation procedure for the edge corrupted by noise is determined as follows. At \( r = c = 0 \), the central pixel is estimated by

\[
e(0,0) = \hat{c} + r(0,0).
\]  

(20)

The centered pixel is taken as a noisy pixel if

\[
|\hat{X}_{\text{smooth}}(0,0) - \hat{c}| > 5\sqrt{1 + \frac{1}{(2M+1)^2 - 1}} e^2. \tag{21}
\]

The detected noisy pixel is then replaced by its fitted value \( \hat{c} \). In this algorithm, those corrupted pixels are elegantly handled and replaced by a proper value within the window region. The best window size used is \( 5 \times 5 \) which is statistically enough to differentiate noisy pixel from feature pixel to be preserved.

Weak edges when corrupted with blocking noise may be deleted during the deblocking process. To preserve those edges a \( 3 \times 3 \) window is used to scan the processed image. If two edges are located within the window and the center pixel is not recognized as an edge, the angles of those edges are compared. If those edges have same angles, it means that the center is a missing edge and the mean value of those edge pixels is given to the centered pixel. In this proposed approach, the noise that affects strong and weak edges is removed and all the details that are lost during the adaptive filtering are strongly preserved.

IV. SIMULATION RESULTS

To evaluate the performance of the proposed algorithm with the existing deblocking methods, we block coded all images using a high compression ratio. The bit rate used in this study was between 0.1 and 1.0 bits/pixel. Different types of images are compressed using standard JPEG code, which was distributed by the Portable Video Research Group.

A. Image Quality Control

In general, there are two types of distortion measure used for picture quality prediction. The first type, the raw error measure, gives simple mathematical deviations between the original and the de-blocked images. The most widely used measure is the mean square error or the peak signal to noise ratio PSNR. Although, it quantifies the error mathematically, it does not necessarily reflect the observer’s visual perception of the error. The second type of distortion measure is based on human perception. For appropriate picture quality prediction, a distortion measure that quantifies human perception is devised. This measure should rate the dissatisfaction of the viewer in observing de-blocked images [26]. One way of determining such measure is subjective testing. Subjects view a series of de-blocked images and assess them based on noticeable artifacts. Although such approach is tedious and time consuming, it provides a uniform measure of dissatisfaction when used with a large number of images. Subjective testing targeting special type of artifacts may provide a robust measure of image quality. A mean opinion score (MOS) is a numerical measure for the subjective performance of the algorithm, where each artifact is rated based on Table I [26].

Simple distortion scales, such as the PSNR or weighted signal to noise ratio WSNR are usually used as an evaluation measure for comparison with other techniques. Such measures are good indicators for random errors but not for structured or correlated errors. An objective measure that scales the overall dissatisfaction of the viewer is to be determined by analyzing the different types of artifacts separately and weight them according to the human sensitivity. Blocking artifacts are the major source of error in transformed images and so it is the most annoying part. Although the main goal of this study is to establish a framework for alleviating deblocking artifacts, a simple objective measure targeting the structure format of the artifacts is devised. The objective measure is similar in principle to the one proposed in [25], Fig. 2 shows the processing steps for this approach. Both the de-blocked image and the original image are fed into the estimator. For each image both vertical and horizontal edges are determined. The amplitude of vertical and horizontal edges is then computed. The percentage edge error (PEE) is then determined by

\[
\text{PEE} = \frac{E_{\text{ORG}} - E_{\text{DB}}}{E_{\text{ORG}}} \times 100\%
\]

where \( E_{\text{ORG}} \) is the edge strength of the original image and \( E_{\text{DB}} \) is the edge strength of the de-blocked image. The PEE measures how close the de-blocked image details from the original image; negative PEE means that the de-blocked image has more details than the original image and so blocking artifacts are not completely removed. On the other hand, positive PEE means that the de-blocked image is over smoothed and possibly image details are corrupted. Therefore, PEE not only measures the details of the image quality but it also gives the weight by which the deblocking algorithm deteriorates these details. The edge strength is determined by

\[
E_S = \sum_{i=1}^{N} \sum_{c=1}^{M} (E_I(r,c))
\]

(23)

where \( E_I(r,c) \) is the edge intensity value of the analyzed image. The PEE is a strong indicative measure of the structural percentage errors due to blocking. The decreased value of PEE is a strong measure of reducing the blocking effect. In this study simple kernels are used to achieve horizontal and vertical edges. It is worth mentioning that the intention here is to use an objective measure that evaluates image edges rather than optimizing the measure. The used kernels are

\[
\begin{bmatrix}
F1 & = & \frac{1}{3} \begin{bmatrix} 1 & 0 \\
F2 & = & \frac{1}{3} \begin{bmatrix} 1 & -1
\end{bmatrix}
\end{bmatrix} \\
\end{bmatrix}
\]

(24)
Fig. 2. Processing procedure to quantify the edge intensity of the image. Image is passed into two parallel edge detector, the first detector aims at extracting vertical edges while the second is for horizontal edges. The amplitude of the edge image is then computed. The pixel of the amplitude image is considered an edge intensity if it exceeds a threshold value shown in the final block of the figure.

Fig. 3. Woman image and the post-processed images: From left to right; the original image, JPEG decoded image, and the de-blocked image using the proposed algorithm.

where \( F_1 \) is the low-pass kernel and \( F_2 \) is the high pass kernel.

**B. Image Deblocking Procedure**

The image used in this case is the 256 \( \times \) 256 woman picture from Matlab, which is decoded using the JPEG standard with variable compression factor. Fig. 3 shows the original woman image along with the JPEG decoded image with compression ratio of 36 : 1, and the de-blocked image. The high pass kernel used in this process is

\[
w = \frac{1}{33} \begin{bmatrix} -1 & -1 & -1 \\ -10 & 27 & -10 \\ -1 & -1 & -1 \end{bmatrix}.
\]

The high pass filter coefficients are chosen to best match the image high-frequency characteristics. The coefficients are experimentally determined to maximize the PSNR of the image woman. The main effect of the kernel in the deblocking process is to enhance the appearance of those details that are covered by blocking noise. Therefore, any kernel that has the ability to enhance details can be used to provide the crispiness of textures that are dominated by noise. Moreover, the kernel may be optimized to suite special type of images that maximize the reduction in de-blocked image disturbance. In the next stage, the statistical differencing approach is performed to give an approximate estimate to those details that have been lost or smeared during the coded process. This process also provides the adaptive filtering stage with tools that differentiate between the image details and noise corruption. The value used for the factor \( k \) in the statistical approach was 0.325. The relation between the statistical factor and the bit rate of the JPEG decoded images is shown in Fig. 4. This figure is obtained experimentally by compressing set of test images and finding the value of \( k \) that provides the best performance. The curve shows the aforementioned correlation in statistical differencing approach between the \( k \) value and the bit rate. The curve can be used as reference for any target bit rate. Using spline interpolation will offer the best value of \( k \) for a given bit quota.

In this stage, the gain in PSNR is between 2 to 3 dB for uniformly distributed images. Generally, most of the images have these criteria and so the improvements are considerably high. After this stage, the resultant image is fed into the adaptive filtering stage that involves generating both local and global maps.
of the detected edges. This process will literally smooth the contaminated image and improve the visual quality so that the effect of noise is no longer observable. Although, in this stage, the smoothness is threshold dependent but we could come up with an empirical relation for thresholds in terms of the compression factor of the JPEG standard. This method performs well for all the tested images.

In the fourth stage of the algorithm, an outlier detector and replacement scheme, adopted from [27], is utilized to improve the visual quality of the image. In the fifth stage of the algorithm, a detection procedure for the quantization noise on global edges and replacement based on correlation information within a $5 \times 5$ neighborhood is carried out. In this part, those pixels that are dominated by noise are detected and replaced by their correspondent estimates. Since the image is smoothed initially, the threshold value used in this section is halfway of the distribution bandwidth. In this stage, the gain in the PSNR is between 0.05 to 0.2 dB. Weak edges that didn’t survive due to the consecutive de-noising processes are finally regularized by linkage of the gapped edges both globally and locally. Fig. 5 shows the corresponding progress throughout the stages of the proposed algorithm. Fig. 6 shows the effectiveness of the algorithm in de-noising and regularizing near strong edges by zooming on a selective part of both blocky and de-blocked images with high bit rate.

C. Evaluation of the Proposed Approach

Two objective measures are induced to study the effect of both random and structural errors on the de-blocked images. Furthermore, a subjective quality tests that uses the MOS has been conducted to evaluate the perceptual quality of the decoded images. A comparative study is also performed to evaluate the algorithm with the several existing techniques proposed in the literature. Table II shows the performance of the proposed algorithm in comparison with other conventional methods for JPEG-encoded images that are compressed using the recommended quantization table of Joint Photographic Expert Group. As it is clearly noticeable from Table II, the performance of the proposed algorithm has high PSNR in most of the tested images. Lena image
TABLE II
Comparison of the Performance of the Proposed Algorithm With Conventional Methods. BPP Is the Bit Per Pixel, SAF, LCP Is the Deblocking Algorithm of [8], P OCS, 13 Is the Deblocking Algorithm of [11], W TM, MM, Is the Deblocking Algorithm of [17], and CECSD Is the Proposed Algorithm. ΔPSNR Is the Gain of the Corresponding Algorithm Over the JPEG-Coded Approach in dB.

<table>
<thead>
<tr>
<th>IMAGE</th>
<th>BPP</th>
<th>SAF, LCP</th>
<th>POCS, 13</th>
<th>W TM, MM</th>
<th>CECSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Woman</td>
<td>0.257</td>
<td>0.57</td>
<td>0.12</td>
<td>0.56</td>
<td>3.81</td>
</tr>
<tr>
<td>Pout</td>
<td>0.223</td>
<td>0.03</td>
<td>0.12</td>
<td>0.56</td>
<td>4.90</td>
</tr>
<tr>
<td>Kids</td>
<td>0.294</td>
<td>0.15</td>
<td>0.12</td>
<td>0.56</td>
<td>4.90</td>
</tr>
<tr>
<td>Bone marrow</td>
<td>0.493</td>
<td>0.14</td>
<td>0.12</td>
<td>0.56</td>
<td>4.90</td>
</tr>
<tr>
<td>Lena</td>
<td>0.202</td>
<td>0.07</td>
<td>0.12</td>
<td>0.56</td>
<td>4.90</td>
</tr>
<tr>
<td>Geometry</td>
<td>0.33</td>
<td>0.15</td>
<td>0.12</td>
<td>0.56</td>
<td>4.90</td>
</tr>
<tr>
<td>Canoe</td>
<td>0.077</td>
<td>0.14</td>
<td>0.12</td>
<td>0.56</td>
<td>4.90</td>
</tr>
<tr>
<td>Tartan</td>
<td>0.37</td>
<td>0.23</td>
<td>0.12</td>
<td>0.56</td>
<td>4.90</td>
</tr>
<tr>
<td>Baboon</td>
<td>0.368</td>
<td>0.12</td>
<td>0.12</td>
<td>0.56</td>
<td>4.90</td>
</tr>
</tbody>
</table>

TABLE III
Comparison of Proposed Algorithm Performance With Conventional Deblocking Techniques Shown in Table II Using Percentage Edge Error PEE.

<table>
<thead>
<tr>
<th>IMAGE</th>
<th>SAF, LCP PEE (%)</th>
<th>POCS, 13 PEE (%)</th>
<th>W TM, MM PEE (%)</th>
<th>CECSD PEE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Woman</td>
<td>22.54</td>
<td>53.58</td>
<td>55.88</td>
<td>5.54</td>
</tr>
<tr>
<td>Pout</td>
<td>-49.59</td>
<td>-39.59</td>
<td>-64.04</td>
<td>-35.58</td>
</tr>
<tr>
<td>Kids</td>
<td>-92.27</td>
<td>-122.58</td>
<td>55.88</td>
<td>5.54</td>
</tr>
<tr>
<td>Bone marrow</td>
<td>-35.64</td>
<td>-35.64</td>
<td>-64.04</td>
<td>-35.58</td>
</tr>
<tr>
<td>Lena</td>
<td>-30.27</td>
<td>-30.27</td>
<td>-64.04</td>
<td>-35.58</td>
</tr>
<tr>
<td>Geometry</td>
<td>-74.38</td>
<td>-74.38</td>
<td>-64.04</td>
<td>-35.58</td>
</tr>
<tr>
<td>Canoe</td>
<td>-97.43</td>
<td>-97.43</td>
<td>-64.04</td>
<td>-35.58</td>
</tr>
<tr>
<td>Tartan</td>
<td>-38.05</td>
<td>-38.05</td>
<td>-64.04</td>
<td>-35.58</td>
</tr>
<tr>
<td>Baboon</td>
<td>-38.05</td>
<td>-38.05</td>
<td>-64.04</td>
<td>-35.58</td>
</tr>
</tbody>
</table>

Table IV shows the results of the PEE for the same images and the performance of the proposed approach and the other de-blocking techniques. Results noticeably show that the PEE for the proposed approach is very low compared with other techniques, where the mean PEE is less than any of the considered techniques. The main observation here, is that structural features are preserved and less affected by the deblocking effect. Other approaches resolve the deblocking by over smoothing the image details.

To show the robustness of the proposed approach, five independent viewers are asked to evaluate the decoded images based on Table I. They are asked specifically to rate the following: edge level, edge length, background luminance, continuity of image features, and overall image. The MOS value is determined by the following:

\[
\text{MOS}(k) = \frac{1}{M_o M_f} \sum_{i=1}^{M_o} \sum_{j=1}^{M_f} \text{SC}(i, j, k) \tag{25}
\]

where
\[
\text{SC}(i, j, k) \quad \text{evaluation score for feature (j) of image (k) by viewer (i)};
\]
\[
M_o \quad \text{number of observers};
\]
\[
M_f \quad \text{number of features}.
\]

Table IV shows the MOS ratings for the Lena image where our PSNR in our case does not gain the highest value among the tested images. Table V shows the overall MOS ratings for all the tested images.

V. CONCLUSION

In this paper, a more comprehensive approach to de-blocking JPEG decoded image is proposed and compared with other ex-
isting deblending algorithms. The proposed approach is able to smooth the decoded image while preserving the image details. The approach utilizes the advantages of the previous image enhancement techniques and overcomes some of their defects. The improvement in PSNR, PEE, and the visual perception of the image demonstrate the effectiveness of the considered technique. Some images may perform better than others due to the variations in their statistical characteristics and image formation. Therefore, a considerable improvement in a specific class of images may need optimization of other parameters that are not explored here. The ability to combine this algorithm with other existing image enhancement techniques makes it a useful technique for deblending JPEG encoded images.

REFERENCES