CHAPTER 4

TV REGULARIZATION METHOD

4.1 INTRODUCTION

MPEG International Standard has been used as the format for compressing and storing still images and moving respectively. MPEG-2 has been a widely accepted video standard for various applications ranging from DVD to Digital TV Broadcast. A large variety of products based on the MPEG-2 standard are available in the market. The most important goal of MPEG-2 was to make the storage and transmission of digital AV material more efficient. The new H.264/AVC standard has an even broader perspective to support high and low bit-rate multimedia applications on existing and future networks.

The advantage in terms of better quality at a lower bit rate is why H.264 is fast replacing MPEG-2. In the DCT-based compression methods such as JPEG, MPEG and H.264, the compression artifacts occur by quantization of DCT coefficients. The quantization of low frequency coefficients generates blocky noise and the quantization of high frequency coefficients generates mosquito noise.

When the data transmission channel bandwidth is narrow, the data rate is lowered and quantization level is dropped. As a result, the compression artifacts are increased. In DCT coding, the original image is divided into blocks and the DCT is processed in each block, integrating signals in all
blocks into low bands and quantizing the DCT coefficients so that compression performance can be increased. However, higher the compression rate is set, the lower the quality of reconstructed images because of increased blocky noise. This is because there is no correlation between blocks as a result of quantizing low band signals, and since the high band signals are quantized, reconstructed images become blurry and include mosquito noise.

In (Fogg et al 2009), a model was obtained that gives the numerical value depending upon the visibility of the blocking artifacts in compressed images and thus requires original image for comparison with reconstructed image. In practice the original images will not be available. In this model, blocky image is modeled as a non blocky image interfering with a pure blocky signal. Blocking artifacts measurement is accomplished by estimating the power of blocky signal. The weakness of is to assume that the difference of the pixel value in block boundary is caused only by blocking artifacts. This assumption decreases computation complexity but the measured value does not confirm to truth for the two adjacent blocks with a gradual change in pixel value. The variation of pixel value across block boundary was modeled as a linear function. This method is not accurate especially for the adjacent block with a large change of pixel value across the block boundary (Stephen Warrington et al 2009).

Blockiness is the most prevailing artifact in Block Discrete Cosine Transform (BDCT) code image (Zhai et al 2008) and video (Mitcheel 2007) under low bit-rate conditions, due to the independent transformation and quantization of image blocks. In order to ameliorate the perceptual picture quality, numerous blockiness reduction algorithms have been proposed during the last two decades. Analytically, the blockiness artifacts can be divided into two categories (Kim et al 1998; Al-Fohoum & Reza 2006): the grid noise in monotone areas and the edge related artifacts, such as staircase noise along
edges, ringing around strong edges and corner outliers. The grid noise in monotone area, which is often referred to as “blockiness" in general sense, is a kind of structural artifact.

It attracts much human attention and is thus the most annoying artifact (Rao & Wu 2005) and that is the reason why most deblocking algorithms deal with grid noise only. However, since edges are fundamental cognitive clues in image for a superior perceptual quality, the edge-related artifacts should also be appropriately addressed (Gasho & Ramamurthi 1986).

The remarkable phenomenon is that all blocky noise and mosquito noise are separated into the texture component. Several post-processing methods have been applied to remove these artifacts but they could not address to the problem of the texture component loss. The proposed method achieves efficient reduction of blocky noise and mosquito noise without texture component being lost. The usual TV will be changed in a weighted TV that regularizes blocks’ edges without regularizing images’ true edges. Although the algorithm converges in infinite time, one obtains best PSNR with a very few number of illustrations, leading therefore to a fast method.

4.2 IMPLEMENTATION OF IMAGE DENOISING & ITS RECONSTRUCTION

In this chapter, a new compression artifact removable method is developed. The Total Variation (TV) regularization method is utilized, and a texture component is obtained in which both the blocky noise and mosquito noise are decomposed. These artifacts are removed by using selective filters controlled by the edge information from the structure component and the motion vectors. The experimental results show much better performances for removing compression artifacts compared with the conventional method. Modern consumer television sets need a high quality format conversion,
image enhancement, and artifact reduction. Especially artifacts due to block based coding schemes may degrade image quality and make artifact reduction mandatory.

A promising approach to high quality filter design is given by formulating the filter task via a variation problem and by the use of partial differential equations to solve it. Here the design of a generalized total variation regularization filter is described. This filter can be applied in many video signal processing steps in consumer television as pre- or post-processing filter, for example as pre-processing filter prior to a motion estimation to improve image quality or analysis results. The usage in coding artifact reduction without knowledge of decoding parameters is presented in greater detail and the results are also presented. The block diagram for TV Regularization is given in Figure 4.1.

![Figure 4.1 Block diagram for TV Regularization](image-url)
Here, TV Regularization decomposition, Motion compensation, Sobel Filter, Gaussian Filter and DEF are used. MC is the use of motion vectors to improve the efficiency of the prediction of peel values. The prediction uses motion vectors to provide offsets into the past and/or future reference pictures containing previously decoded peel values that are used to form the prediction error signal.

The Sobel filter consists of two kernels which detect horizontal and vertical changes in an image. If both are applied to an image, the results can be used to compute the magnitude and direction of the edges in the image. The Sobel operator is used in image processing, particularly within edge detection algorithms. Technically, it is a discrete differentiation operator, computing an approximation of the gradient of the image intensity function.

At each point in the image, the result of the Sobel operator is either the corresponding gradient vector or the norm of this vector. The Sobel operator is based on convolving the image with a small, separable, and integer valued filter in horizontal and vertical direction and is therefore relatively inexpensive in terms of computations. In electronics and signal processing, Gaussian filter is a filter whose impulse response is a Gaussian function. Gaussian filters are designed to give no overshoot to a step function input while minimizing the rise and fall time.

This behavior is closely connected to the fact that the Gaussian filter has the minimum possible group delay. DEF is one of the most well-known methods of reducing blocky noise, which is defined on ITU-T H263–Annex J. In DEF, by filtering both sides of 2 pixels between the blocks in the reconstructed images with blocky noise, blocky noise is smoothed out.
A stream compressed by MPEG-2 is decomposed into the structure component and texture component by using the regularization. And all signals in the structure component are filtered by using sobel filter, and threshold is set to extract edge components. As a result of using the proposed method, the blocky noise and mosquito noise are removed with edge preservation.

4.2.1 Decomposition of Mosquito Noise

Mosquito Noise is a temporal, statistically random appearance of flicker in individual pixels along image edges of consecutive frames of a video. The noise level in the areas that are along the edges is significantly higher than in other areas of the image. Pixel flicker has to last over a number of consecutive frames for that to be recognized as MN. MN appears particularly along slow moving edges. This is a consequence of two major shortcomings of common video coding algorithms; ringing and motion compensation. MN can be described as a temporal subset of these two spatial video coding artifacts.

As mentioned in this definition, the major manifestation of MN is the change of pixel luminance values between consecutive frames, but in color images MN can also appear and is then characterized by a color flicker of pixel values in small spatial areas, generally not in a larger extent than one or two pixel sizes. The following explanations are based on current video coding algorithms such as MPEG-2 or H.264/MPEG-4 AVC without loss of generality of the previous definition.

A. Spatial Aspect of Mosquito Noise

The major cause of MN is ringing, which an artifact is resulting from compression algorithms implemented in video codecs. MN mainly originates in I-frames inside the Group of Pictures (GOP) and in residual
coding of motion compensated frames. With increasing compression rate accompanied by an increasing quantization of coefficient values (e.g. coefficients of the discrete cosine transformation) of the transformed blocks, undershoots and overshoots of pixel values also referred to as ripples along image edges increase and will be visible in the image.

B. Temporal Aspects of Mosquito Noise

The second reason for the appearance of MN is Motion Compensation (MC) that is used in current codecs. On one hand MC performs a great service in order to reduce the bandwidth needed to broadcast video or to reduce storage requirements. On the other hand this technique introduces new video artifacts and MC shifts and spreads the already existing artifacts (like ringing) inside the image space of consecutive frames.

It occurs when reconstructing the image and approximating discarded data by inverting the transform model (IDCT). "Mosquitoes" can also be found in other areas of an image. For instance, the presence of a very distinct texture or film grain at compression will also introduce mosquito noise.

Figure 4.2 Mosquito noise level measurement
Figure 4.2 shows the test pattern in (a) and the profile differences in (b). To highlight the mosquito noise artifact in H.264 compressed video sequences, test sequences of moving and rotating spiral were designed. Due to the ripple-characteristics of the ringing artifact, which is the main contributor to mosquito noise, a new test pattern (Dartboard pattern) was developed as shown in Figure 4.2. The related video sequence is composed of different gray values instead of black and white to provide foot room and headroom for undershoots and overshoots.

To achieve better MN measurement results, the pixel value differences are filtered out between profiles of consecutive frames of the uncompressed video sequence. These differences originate from the use of a polar profile. Furthermore, a threshold is applied to consider only pixel differences that are typical for mosquito noise and that which do not originate from any other compression artifact.

To assess MN, a rotation-invariant polar re-sampled version of the circular test pattern was used. Computations only incorporate areas of proposed test pattern around edges in the ROI as shown in Figure 4.2. This approach allows the rotation and shifting of the test pattern through the video as shown in Figure 4.3 without losing the ability to compare equivalent parts of the test pattern in consecutive frames.

To stress the motion compensation algorithm of the encoder, the structure was spun between consecutive frames by one degree and shifted one pixel position in both the x- and y-directions as shown in Figure 4.3. Thereby, slow varying and fast moving parts are obtained in the image, so that global motion estimation can hardly be applied by the codec.
If the MC itself produces inaccurate block matches, high pixel differences in the very near area along the image edges can occur. Due to the fact that these MC mismatches had a major influence on the image edge itself and a minor influence on the surrounding pixels, this image distortion does not contribute much to the level of MN in the image sequence.

This approach to measure MN highly differs from the MN mitigation. By reducing the overall ringing level in a video sequence, the MN level will decrease automatically. But due to the fact, that MN is just a subset of ringing; the measurement has to take both the spatial and the temporal aspect of this video artifact into account.

C. MOSQUITO NOISE MEASUREMENT

The pixel intensity value difference for each profile of consecutive frames is calculated as:

Figure 4.3 Movement of dartboard pattern in test sequence
\[ p(\alpha, I)_{\text{diff}} = |p(\alpha, I)_t - p(\alpha, I)_{t-1}| \]  

(4.1)

where,

\( I \) = Positions of a pixel value inside a measurement profile,

\( p(\alpha, I)_t, p(\alpha, I)_{t-1} \) are threshold ROIs of conjugate measurement profiles of consecutive frames.

### 4.2.2 Decomposition of Blocky Noise

A Blocky noise is a distortion that appears in compressed video material as abnormally large pixel blocks. Also called "macro blocking," it occurs when the encoder cannot keep up with the allocated bandwidth. It is especially visible with fast motion sequences or quick scene changes. Video uses lossy compression, and the higher the compression rate, the more content is removed.

In medical field, lossless compression techniques are used because while recovering the image some of the information may be lost if lossy compression is used. But in video and audio some information lost is acceptable so, in such a situation lossy compression techniques can be applied. Also, the bandwidth usage is less in transmission. At decompression, the output of certain decoded blocks makes surrounding pixels appear in average together and look like larger blocks (as shown in Figure 4.4)
4.3 DESIGN CHARACTERISTICS OF THE PROPOSED ALGORITHM

4.3.1 TV Regularization method

DENOISING

-reconstruct u from the given data f.

Basic assumptions:

1. Noise characterized by fast oscillations.
2. Image consists of well separated uniform regions.

Assume that the noise satisfies

\[ \| n^\delta \|_2^2 = \int_\Omega |n^\delta(X)|^2 \, dx \approx \delta^2. \]  

(4.3)
Minimize

\[ \| D_u \| (\Omega) : \approx \int_{\Omega} \sqrt{x} \, |dx| \]  \quad (4.4)

Subject to the constraint

\[ \| u-f \|_2^2 = \int_{\Omega} (u(x)-f(x))^2 \, dx = \delta^2. \]  \quad (4.5)

Choose some final target variation \( 0 < \Theta < 1 \) and update \( \alpha \) until

\[ \text{var}_n (v_\alpha/\alpha) \approx \Theta. \]  \quad (4.6)

if \( \text{var}_n (v_\alpha/\alpha) (x) < \Theta \): decrease \( \alpha(x) \).

If \( \text{var}_n (v_\alpha/\alpha) (x) > \Theta \): increase \( \alpha(x) \)

Chosen update:

\[ \alpha_{\text{new}}(x) = \alpha(x) \cdot \text{var}_n(v_\alpha/\alpha)(x) + \Theta / 2 \Theta \]  \quad (4.7)

For better stability: convolve \( \alpha_{\text{new}} \) with some smooth kernel \( p \).

### 4.3.2 Non-Constant Regularization method

Instead of a number \( \alpha \in (0, +\infty) \), use a continuous regularization function

\[ \alpha : \Omega \rightarrow (0, +\infty), \]

and minimize

\[ T(U; \alpha) = 1/2 \int_{\Omega} (u(x)-f(x))^2 \, dx + \int_{\Omega} \alpha(x) \, |D U| \]  \quad (4.8)
4.4 EXPERIMENTAL ENVIRONMENT

The phrase Peak Signal-To-Noise Ratio, often abbreviated as PSNR, is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. Because many signals have a very wide dynamic range, PSNR is usually expressed in terms of the logarithmic decibel scale. The PSNR is most commonly used as a measure of quality of reconstruction of loss compression codec’s (e.g., for image compression). The signal in this case is the original data, and the noise is the error introduced by compression.

When comparing compression codec it is used as an approximation to human perception of reconstruction quality, therefore in some cases one reconstruction may appear to be closer to the original than another, even though it has a lower PSNR (a higher PSNR would normally indicate that the reconstruction is of higher quality). It is most easily defined via the Mean Squared Error (MSE) which for two \( m \times n \) monochrome images \( I \) and \( K \) where one of the images is considered a noisy approximation of the other as in Equation (3.8)

\[
\text{PSNR} = 10 \log_{10} \left( \frac{\text{MAX}_I^2}{\text{MSE}} \right)
\]

Here, as in equation (3.9) \( \text{MAX}_I \) is the maximum possible pixel value of the image. When the pixels are represented using 8 bits per sample, this is 255. More generally, when samples are represented using linear PCM with \( B \) bits per sample, \( \text{MAX}_I = 2^B - 1 \). For color images with three RGB values per pixel, the definition of PSNR is the same except the MSE is the sum over all squared value differences divided by image size and by three. The PSNR value measurement graph is shown in Figure 4.5.

Typical values for the PSNR in lossy image and video compression are between 30 and 50 dB, where higher is better.
Acceptable values for wireless transmission quality loss are considered to be about 20 dB to 25 dB. PSNR is derived by setting the Mean Squared Error (MSE) in relation to the maximum possible value of the luminance. For an n-bit value it is defined as follows:

\[ M_{i=1} \& N_{j=1} [x(i, j) - y(i, j)]^2 \frac{1}{M} \cdot \frac{1}{N} \quad (4.9) \]

\[ 2n - 1 \quad \text{PSNR} = 20 \log_{10} \sqrt{\text{MSE}} \quad (4.10) \]

where,

\[ x(i, j) \] and \[ y(i, j) \] are the original and processed signals at pixel (i, j) and \( M, N \) are the picture dimensions. The resultant is a single number expressed in decibels (dB). PSNR is the most widely used metric for evaluating 2D video quality. However, its correlation with the HVS perception of quality is not strong.

A. OUTPUT ILLUSTRATIONS:

Figure 4.6 shows the original image of Lena
Figure 4.6 Original image of Lena

Figure 4.7 shows the noisy image of Lena. Generalization of the TV process is performed to adaptive power constraints. Noise is strong in the smooth region and weaker in textured regions.

Figure 4.7 Noisy image of Lena

Figure 4.8 shows the Adaptive TV image of Lena with the reduced total-variation of the image. It filters out noise while preserving edges. Textures and fine-scale details are also removed. The effectiveness of the two
The proposed filtering method (for reduction of the blocking effect with an increase, or a slight reduction, of PSNR value) is clearer in cases of small numbers of the DCT coefficients, i.e. in cases corresponding to high compression ratios.

![Figure 4.8 Adaptive TV image of Lena](image)

**Figure 4.8 Adaptive TV image of Lena**

**MOSQUITO NOISE**

Figure 4.9 shows the original image of Barbara.

![Figure 4.9 Original image of Barbara](image)

**Figure 4.9 Original image of Barbara**

Figure 4.10 shows the noisy image of Barbara where the Video compression artifacts include cumulative results of compression of the
comprising still images, for instance ringing or other edge busyness in successive still images appear in sequence as a shimmering blur of dots around edges, called mosquito noise, as they resemble mosquitoes swarming around the object.

![Figure 4.10 Noisy image of Barbara](image1.png)

**Figure 4.10 Noisy image of Barbara**

Figure 4.11 shows the Adaptive TV image of Barbara image that gives the blur image if Barbara with the preserving components of mosquito’s swarming around the picture.

![Figure 4.11 Adaptive TV image of Barbara](image2.png)

**Figure 4.11 Adaptive TV image of Barbara**
BLOCKY NOISE

Figure 4.12 shows the sample original image of Lena

![Figure 4.12 Sample original image of Lena](image)

Figure 4.12 Sample original image of Lena

Figure 4.13 shows the noisy sample original image of Lena with distortion that appears in compressed video material as abnormally large pixel blocks.

![Figure 4.13 Noisy sample original image of Lena](image)
Figure 4.14 shows the Adaptive TV sample original image of Lena image. It is especially visible with fast motion sequences or quick scene changes. Video uses lossy compression, and the higher the compression rate, the more content is removed. This effect becomes even more pronounced when there is some fast motion or quick camera movement.

![Adaptive TV sample original image of Lena](image)

**Figure 4.14 Adaptive TV sample original image of Lena**

### B. GRAPHICAL REPRESENTATION

Figure 4.15 shows the Adaptive TV, which on x-axis represents noise variance and y-axis represents signal to noise ratio.

![Noise variance Vs Signal to noise ratio](image)

**Figure 4.15 Noise variance Vs Signal to noise ratio (Adaptive TV)**
Figure 4.16 shows the Mosquito noise, which on x-axis represents noise variance and y-axis represents signal to noise ratio.

![Figure 4.16 Noise variance Vs Signal to noise ratio (Mosquito noise)](image)

Figure 4.16 Noise variance Vs Signal to noise ratio (Mosquito noise)

Figure 4.17 shows the represents the Blocky noise, which on x-axis represents noise variance and y-axis represents signal to noise ratio.

![Figure 4.17 Noise variance Vs Signal to noise ratio (Blocky Noise)](image)

Figure 4.17 Noise variance Vs Signal to noise ratio (Blocky Noise)
The aim of the method is to keep the signal to noise ratio value high. In the above graphs, SNR has been shown for various noise variances. Both SNR and PSNR are used to measure the quality of image after reconstruction.

C. SIGNAL TO NOISE RATIO

Given the variance of the signal $\sigma_s^2$ and the variance of the noise $\sigma_n^2$, SNR can be calculated as,

$$\text{SNR}=10 \log_{10}(\sigma_s^2/\sigma_n^2)$$

(4.11)

After rearranging $\sigma_s^2 = \sigma_n^2 10 \text{SNR}/10$

(4.12)

Signal-to-noise ratio is defined as the power ratio between a noise signal (meaningful information) and the background (unwanted signal):

$$\text{SNR} = \frac{P_{\text{signal}}}{P_{\text{noise}}}$$

(4.13)

where,

$P$ is average power. Both signal and noise power must be measured at the same and equivalent points in a system and within the same system band width. If the signal and the noise are measured across the same impedance then the SNR can be obtained by calculating the square of the amplitude ratio:

$$\text{SNR} = \frac{P_{\text{signal}}}{P_{\text{noise}}} = (A_{\text{signal}}/A_{\text{noise}})^2$$

(4.14)

$$\text{SNR}_{\text{dB}} = 10 \log_{10} \left( \frac{P_{\text{signal}}}{P_{\text{noise}}} \right) = P_{\text{signal}} \text{ dB} - P_{\text{noise}} \text{ dB}$$

(4.15)

Which may equivalently be written using amplitude ratios as:

$$\text{SNR}_{\text{dB}} = 10 \log_{10} (A_{\text{signal}}/A_{\text{noise}})^2 = 20 \log_{10}(A_{\text{signal}}/A_{\text{noise}})$$

(4.16)
The performance of SNR is improved by reducing the noise. All real measurements are disturbed by noise. This includes electronic noise, wind, vibration, temperature etc., depending on what is measured and sensitivity of the device. It is possible to reduce the noise by controlling the environment otherwise when the characteristics of the noise is known it is possible to filter the noise.

Figure 4.18 represents both the Mosquito noise and the Blocky noise, for which x-axis represents Signal to Noise Ratio (SNR) and y-axis represents Peak Signal to Noise Ratio (PSNR).

![Figure 4.18 PSNR Vs Signal to Noise Ratio for Mosquito and Blocky Noises](image)

Figure 4.19 shows the Adaptive TV, which on x-axis represents PSNR and y-axis represents Frame index. Figure 4.19 (a) is the original frame and 4.19 (b) is the noisy frame. Figure 4.19 (c) is the filter PSNR. After filtering the PSNR value, it increased to 40.8332. Figure 4.19 (d) shows that the PSNR value after deblurring. It increased to 43.42. There is a certain
improvement in PSNR value. This shows the quality of the Figure 4.19 (d) is high.

![Figure 4.19](image)

**Figure 4.19** (a) Video Frame (b) Noisy Frame (c) Filter PSNR (d) Deblurred Frame

Figure 4.20 shows the Mosquito noise, which on x-axis represents PSNR and y-axis represents Frame index.

![Figure 4.20](image)

**Figure 4.20** Frame index Vs PSNR (Mosquito Noise)
Figure 4.21 shows the Blocky noise, which on x-axis represents PSNR and y-axis represents Frame index.

![Figure 4.21 Frame index Vs PSNR (Blocky Noise)](image)

**D. PSNR**

PSNR is usually expressed in terms of the logarithmic decibel scale. The signal in this case is the original data, and the noise is the error introduced by compression. When comparing compression codecs it is used as an approximation to human perception of reconstruction quality, therefore in some cases one reconstruction may appear to be closer to the original than another, even though it has a lower PSNR as defined in Equation (3.8) and (3.9) (a higher PSNR would normally indicate that the reconstruction is of higher quality). The PSNR results are shown in Table 4.1. The PSNR value of the proposed method is increased with an increase in the corresponding gain value.
Table 4.1 PSNR results

<table>
<thead>
<tr>
<th>Frame Index</th>
<th>PSNR (ADF Method)</th>
<th>PSNR (Filtering Process) (Goto et al 2011)</th>
<th>PSNR (Proposed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frame Index 1</td>
<td>35.121 dB</td>
<td>40.8577 dB</td>
<td>43.4623 dB</td>
</tr>
<tr>
<td>Frame Index 2</td>
<td>35.220 dB</td>
<td>40.8714 dB</td>
<td>43.4603 dB</td>
</tr>
<tr>
<td>Frame Index 3</td>
<td>35.650 dB</td>
<td>40.8332 dB</td>
<td>43.4238 dB</td>
</tr>
</tbody>
</table>

The difference between PSNR of the proposed method and the conventional methods is graphically shown in Fig 4.22

4.5 CONCLUSION

Hence, a new noise removable method for MPEG2 compression artifacts is proposed. In the proposed method an image is decomposed into a structure component and texture component by utilizing TV Regularization
method. As a result, decomposition compression artifacts are included in texture component. Then mosquito noise is decreased by using Gaussian filter around the edges which are detected from the structure components and motion vectors. Blocky noise is decreased by DEF method without losing small texture components which are important for keeping fine picture quality.

In the experimental results there was less blocky noise and mosquito noise in the images reconstructed in the proposed method than those of reconstructed using the conventional method and the image quality in the proposed method is much higher than that in conventional method at all bit rates.

The goal is to improve the visual quality, so perceptual blur and ringing metrics are used in addition to PSNR evaluation and the value is approx. 43%. The experimental results show much better performances for removing compression artifacts compared with the conventional methods.