A NO-REFERENCE MEASUREMENT METHOD OF BLOCKINESS IN DCT-CODED IMAGES THROUGH EDGE DETECTION AND REGRESSION

A Thesis

by

Koray Ozansoy

Submitted to the

Graduate School of Sciences and Engineering

In Partial Fulfillment of the Requirements for

the Degree of

Master of Science

in the

Department of Electrical and Electronics Engineering

Özyeğin University

January 2014

A NO-REFERENCE MEASUREMENT METHOD OF BLOCKINESS IN DCT-CODED IMAGES THROUGH EDGE DETECTION AND REGRESSION

Approved by:

Asst. Prof. H. Fatih Uğurdağ, Advisor
Department of Electrical and
Electronics Engineering
Özyeğin University

Prof. Tanju Erdem
Department of Electrical and Electronics
Engineering
Özyeğin University

Dr. M. Furkan Kıraç
Department of Computer Science
Özyeğin University

To My Family:

Nursel Ozansoy, Önder Ozansoy,

and Tuğberk Ozansoy

ABSTRACT

This thesis presents a blocking artifact detection algorithm and consists of detailed experimental results. Today, digitally encoded and transmitted images (or videos) are abound. Blocking artifacts are a serious problem in low data rate transmission of such videos (or image sequences), which employ DCT-based compression algorithms. Too much compression due to bandwidth constraints can introduce blocking artifacts. These blocking defects result in an unpleasant viewing experience by end-users.

Online quality measurement techniques for artifacts such as blockiness are required to for broadcasters and internet service providers to inspect their streaming systems. Blockiness measurement can be done through full-reference (FR), reduced-reference (RR), or no-reference (NR) methods. NR is the most efficient one as accessing reference content may not be always possible. NR methods also offer the possibility of real-time inspection.

We propose a novel NR blockiness measurement method. The method is based on DCT and uses a model of the human visual system like some of the literature. However, it is unique as it uses Sobel edge detection algorithm to avoid miscalculations and also employs regression analysis to match mean opinion scores (MOS) of human observers. These unique features of our method let us achieve results that are more correlated to the MOS for LIVE dataset, which is used to train and test our algorithm. This database makes many images available at various levels of blockiness. To measure how well our results match the MOS values in LIVE, we use Pearson and Spearman correlation formulations. The best Spearman correlation obtained in the literature for LIVE dataset is 94% while we get 95%. On the other hand, we get a Pearson correlation of 98% and surpass the best in the literature by 94%.

ÖZETÇE

Bu tezde, bloklanma hatalarını bulan bir algoritma ve detaylı sonuçları sunulmaktadır. Günümüzde, resimlerin ve videoların sayısal kodlamaları ve aktarımları artmaktadır. DCT-tabanlı sıkıştırma tekniklerini kullanan, düşük veri oranına sahip videoların (ya da resim dizilerinin) iletiminde bloklanma hataları ciddi problemler oluşturmaktadır. Bant genişliğindeki kısıtlamadan dolayı uygulanan fazla sıkıştıma hatalarını meydana getirebiliyor. Bu hatalar son kullanıcılarda tatminsizlik yaratmaktadır.

Yayıncılar ve internet servis sağlayıcıları sistemlerindeki bloklanma gibi bozulmaları kalite ölçüm teknikleriyle her zaman değerlendirmek isterler. Bloklanma hataları tam-referanslı, azaltılmış referanslı ya da referanssız ölçüm teknikleriyle yapılabilir. Referanssız ölçüm burada en verimli yöntemdir çünkü, referans veriye ulaşmak her zaman mümkün olmayabilir. Bu yöntem ayrıca gerçek-zamanlı denetlemelere olanak sağlar.

Bu tezde, referanssız bloklanma ölçen yeni bir metod sunduk. Bu metod DCT tabanlı olup diğer literatürlerde de belirtilen insan görsel sistem modelini kullanır. Fakat bu yöntemi özgün kılan, Sobel kenar bulma algoritmasını kullanarak yanlış ölçümlerden kaçınılması ve regresyon analizi ile gözlemcilerin ortalama puanlarının (MOS) deneysel sonuçlarla bağdaştırılmasıdır. Bu özgün yöntemler, LIVE veritabanındaki MOS değerler ile yüksek benzerlikte sonuçlar elde etmemizde yardımcı oldu. LIVE veritabanındaki resimler, algoritma çıktılarının işlenmesinde ve test edilmesinde kullanıldı. Bu veritabanında farklı derecelerde bozulmalara maruz kalmış resimler mevcut. Bizim sonuçlarımızla MOS değerlerinin ne kadar iyi eşleştiğini ölçmek için Pearson ve Spearman korelasyon tekniklerini kullandık. LIVE veritabanı için, literatürdeki en iyi elde edilen Spearman korelasyonu %94 iken, biz %95 elde ettik. Ayrıca, biz %98 Pearson korelasyonu elde ederken literatürde en iyisi %94'tür.

ACKNOWLEDGEMENTS

It was a great experience to work with my advisor Asst. Prof. H. Fatih UĞURDAĞ throughout my thesis. I thank him for his contribution to this work.

Last but not least, I am grateful to my family and thank them for all they have done for me so far.

Koray OZANSOY, January 2014

TABLE OF CONTENTS

ABBREVIATIONS

DCT:	Discrete Cosine Transform
JPEG:	Joint Photographic Expert Group
MPEG:	Moving Picture Experts Group
LIVE:	Laboratory for Image & Video Engineering
HDD:	Hard Disk Drive
CD:	Compact Disc
DVD:	Digital Video Disc
VHS:	Video Home System
CIF:	Common Intermediate Format
HDTV:	High Definition Television
PCM:	Pulse Code Modulation
DPCM:	Differential Pulse Code Modulation
PSNR:	Peak Signal to Noise Ratio
QOS:	Quality of Service
BIQI:	Blind Image Quality Index
BLIINDS:	Blind Image Integrity Notator Using DCT Statistics
DIIVINE:	Distortion Identification-Based Image Verity and Integrity Evaluator
HVS:	Human Visual System
NQM:	Noise Quality Measure
VSNR:	Visual Signal to Noise Ratio
SSIM:	Structural Similarity
MM-SSIM:	Multi-Scale Structural Similarity
UQI:	Universal Quality Index
MICT:	Media Information and Communication Technology
LAR:	Locally Adaptive Resolution
MOS:	Mean Opinion Scores
FPGA:	Field-Programmable Gate Array
ASIC:	Application-Specific Integrated Circuit

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

b_1: Blocking model 1

b_2: Blocking model 2

μ_1: Average value of blocking model 1

μ_2: Average value of blocking model 2

$\varepsilon_{i,j}$: White noise of blocking model 1 with zero mean

$\delta_{i,j}$: White noise of blocking model 2 with zero mean

$b_{(i,j)}$: Definition of new shifted block

s: 2-D auxiliary matrix

β: Amplitude of 2-D auxiliary function of s

$r_{(i,j)}$: Residual function

q_1: Identity matrix 1 to generate shifted block

q_2: Identity matrix 2 to generate shifted block

S: 2-D DCT transform of s

B: 2-D DCT transform of b

B_1: 2-D DCT transform of b_1

B_2: 2-D DCT transform of b_2

Q_1: 2-D DCT transform of q_1

Q_2: 2-D DCT transform of q_2

B_+: Symbol for summation of B_1 and B_2

B_-: Symbol for subtraction of B_1 and B_2

F_+: Symbol for summation of Q_1 and Q_2

F_-: Symbol for subtraction of Q_1 and Q_2

v: First row of the S

R: 2-D DCT transform of r

A: Activity magnitude in a block

η: Visibility of blockiness

θ: Blockiness metric

M_x: Vertical Sobel mask operator

M_y: Horizontal Sobel mask operator

G_x: Vertical direction derivative

G_y: Horizontal direction derivative

G: Gradient value

$A_{i,j}$: Image data

$B_{i,j}$: Convolved image data

$E(Y|x)$: Expected value of Y at each level of x

β_0: Intercept of regression equation

β_1: Slope of regression equation

y_i: Real value of regression model

b_0: Estimated intercept

b_1: Estimated slope for 1st independent variable

b_2: Estimated slope for 2nd independent variable

$\hat{y}_i$: Estimated value of regression model

e_i: Residual of regression model

r: Pearson or Spearman correlation coefficients

x_i: i^{th} element of the first Pearson variable

$\bar{x}$: Average of the first Pearson variable

y_i: i^{th} element of the second Pearson variable

$\bar{y}$: Average of the second Pearson variable

d: Difference between the two numbers in each pair of Spearman ranking

n: Number of pairs of data

CHAPTER I

INTRODUCTION

1.1 General Overview

There are many distortions that occur when lossy compression is applied to the image data. Some of them are blurriness, noisiness, ringing, and blockiness. Lossy data compression is used to transmit and also store media files like images, video, and audio. Using compression brings us many benefits. First of all, we need much less space to store media on HDDs, CDs, and DVDs. Secondly, transmission of data is more effective and easier. Compressed media needs smaller transmission bandwidth via Internet, terrestrial broadcast, satellite links, and mobile phones. Also, it provides faster downloading and uploading of media files through an Internet connection and authoring environment.

It is obvious that compression is really necessary. For example, to motivate the significance of compression, we can state that even an old fashion VHS video with CIF resolution (at 320x240 pixels/color, 8 bits/pixel, 1.5 color components, 30 fps) would need a bandwidth of 40 Mbps. However, after applying MPEG-1 video compression with a 26:1 compression factor, video bit-rate is reduced to 1.5 Mbps. If we consider the commonly used HDTV frame resolution and rate (at 1920x1080 pixels/color, 8 bits/pixel, 3 color components, 30 fps), we would need a bandwidth of 1.44 Gbps. After applying MPEG-4 video compression with a 72:1 compression factor, video bit-rate is reduced to 20 Mbps. Although compression provides great advantage, there will be sacrifices and some kind of artifacts will occur in media files. The more the compression rate, the more the distortion rate.

The most common type of distortion is blockiness. Blockiness, also named as macroblocking, appears in the image as an abnormal pixel blocks. For higher compression rate, more data is unrecoverable during decoding process. So it causes losing some content in media files and creates large pixel blocks on frames. This is annoying for end-users and affects the quality of images negatively. As you can observe in Figure 1.1, there is no blockiness artifact but in Figure 1.2 some macroblocking can be seen especially around the clouds.

Figure 1.1 Original house image

Figure 1.2 House image with lossy compression

Almost all lossy compression techniques use transform coding. The idea is basically breaking the signal up into vectors then decorrelating samples within this vector and finally quantizing transformed samples independently and with simple quantizer (PCM/DPCM). Considering the transform coding techniques, DCT is the most popular. The DCT technique is used in many compression standards. The picture is divided into 8x8 pixels blocks and is applied DCT to obtain the DCT coefficients. Then, the coefficients are quantized. Coarse quantization creates blockiness between two consecutive 8x8 pixels blocks. Hence, there occurs discontinuities at the block boundaries [3].

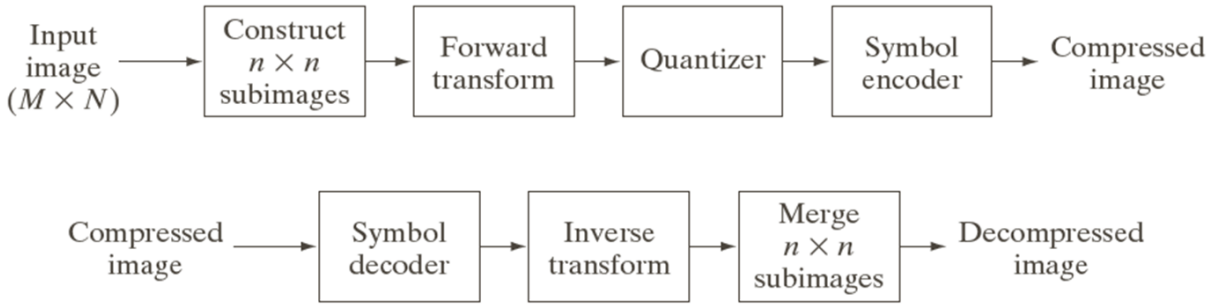

Figure 1.3 Transform coding

1.2 Quality Assessment Techniques

Subjective image quality assessments are becoming unpopular. Subjective characteristics may not be reliable or stable. Human visual system can vary by time or from one individual to another individual. There may be ways to prevent this instability, like increasing the number of expert viewers, taking the mean opinion scores of human observers, etc. Therefore, evaluating images or video datasets is expensive if we consider time.

In recent years, there has been an abundance of work in objective image quality assessment. Some of these techniques require a reference image of the tested image and some do not. The aim is to objectively measure the distortion metrics and give an evaluation opinion. Hence, these techniques can be categorized as full-reference (FR) methods, reduced-reference (RR)

methods, and (NR) no-reference methods. In this thesis, a novel NR method is proposed and is tested on LIVE dataset.

1.2.1 Full-Reference (FR) Techniques

There are numerous FR techniques released in the past decade. They need both reference data and compressed data (i.e., tested image). In [24], eleven FR techniques are evaluated and compared. The most common FR method uses PSNR. The generic PSNR measuring method is applied to the image pixel by pixel. Subjective scores and human psycho-visual system are not considered. Therefore, it is hard to correlate with subjective ratings.

The other techniques mentioned in [24], such as NQM and VSNR, are constructed based on human perception. Their measurements are based on luminance, contrast, and frequency contents. SSIM, UQI, and MM-SSIM use structural measurement. They are all for improving the algorithm and reducing the computation complexity. Also, another target is prediction performance, which needs to be highly correlated with subjective ratings.

FR methods are hard to implement in a real-time measurement system. Therefore, it is not really possible to measure picture quality in transmission systems. In real-world communication environments, image quality algorithms that can use little or no reference signal information are required [19].

1.2.2 Reduced-Reference (RR) Techniques

RR methods also require original data. Unlike FR techniques, RR techniques extract metrics from reference data. These metrics are chosen as features and are measured. RR algorithms use some amount of information from the original data. All parts are not needed for evaluation. Impairments like blockiness, blurriness, or noise estimation are derived from the image or video. After extracting the chosen metrics from the original data, the evaluation process is started. Impairment scores are computed for both original and compressed data. Finally, both scores are compared to arrive at a final result.

These techniques may have some more advantages compared to FR methods. They do not need all the information in an image. Hence, they offer a high-performance solution. Derived features from the original image have much less data than uncompressed information. Thus, transferring these features need lower data rates and it gives advantage for monitoring quality assessment remotely.

Besides this benefit, these algorithms still need reference information for quality assessment. Hence, that remains as a deficit for all RR approaches even though they have advantages over FR methods.

1.2.3 No-Reference (NR) Techniques

Nowadays, image or video transformation via Internet, terrestrial, cable, satellite communication is becoming very widespread. Also, consumers want the broadcaster or internet service provider to deliver high definition streams. In these environments, more quality means more data transfer rate. Thus, it may cause some trouble for QoS while delivering the content to end users. Consequently, a practical solution for testing communication capacity is a necessity. Therefore, a convenient quality assessment tool is needed to get feedback from system for developing perceptual quality [19].

At this stage, NR algorithms come to rescue to solve this problem. NR is a technique that does not require the original image. These algorithms are commonly said to be 'completely blind measurements'.

Even though these algorithms are very handy and useful, they are really hard to implement. Without any information about the original, data assessment is difficult considering human visual system [22]. Many studies have been done, yet no satisfying approach has been published.

In this thesis, a new NR technique is introduced. Through LIVE image database as a test bench, we show that we achieve more correlated results compared to the literature.

1.3 Previous Work

There have been some works in the literature on NR image or video quality assessment. Research groups are still trying to find the least complex and best performance NR algorithm. A significant amount of the effort goes to the search for the best evaluation image or video datasets.

Almost all research groups obtain their results using Pearson and Spearman correlation technique. Subjective and objective results are compared based on Pearson and Spearman

coefficients in the journals and papers published recently. In this paper, these correlations are also used to compare results.

In [16], human visual system is considered, and luminance and texture masking methodologies are used. Similarly in [1], [2], [3], [12], and [13], blocking artifact magnitudes are derived to estimate video quality.

Works in [4], [5], [6], [18] use a mixture of impairments, including blocking defect, to evaluate the images. In these works, things such as blocking, corner outlier, ringing artifacts are considered when coming up with an overall score.

In [23], a visual model is constructed to get a distortion value for blockiness. However, this model needs original and compressed images to obtain a numerical value for the visibility of blocking error.

Also, Prof. Alan Bovik, who is currently working as director of LIVE database, has released some algorithms named BIQI, BLIINDS, and DIIVINE in his papers [20], [21], and [22], respectively. BLIINDS and DIIVINE are compared with the FR technique in [19], which is also by Bovik.

In [7], [9], [10], [11], aside from blocking artifact, other impairments are also considered to evaluate image content. An overall result is obtained by trying to detect addition of edges, lost sharpness of edges, lost in motion, etc.

For perceptual video quality comparison, [8] uses [6], [13], and [16] to compare their performance by applying them different video test set. These set of videos were generated with a different kind of video codec. Finally blockiness metric is analyzed for three algorithms frame by frame.

In [17], an algorithm was proposed for detecting blocking artifact in MPEG encoded videos. I-frames and P-frames are evaluated one by one, and eventually Pearson correlation coefficients are calculated.

[1], [14], [15], [16], [19], [20], [21], and [22] use LIVE image database to test their approach with subjective scores. Also correlation techniques are applied between subjective scores and algorithm outputs to compare results.

1.4 Contributions of This Thesis

The main contribution of this thesis is the development of a new NR algorithm that offers better performance from previous work.

In addition to previous approaches, an edge detection method named “Sobel Edge Detection” is used. Natural edges and blocking edges must be distinguished in an image. The “edges” in images can cause wrong numbers in the evaluation process. In this thesis, Sobel is used to find out natural edges. Natural edges must be eliminated before calculating the quality assessment. Better correlation results have been achieved by applying our technique.

Furthermore, non-linear regression analysis is modeled. Regression analysis is a statistical technique. It is used for exploring the relationship between two or more variables. To implement this analysis model, LIVE image database is used, where there are 233 images available. These images are created from 29 original images. By applying non-linear regression method, errors between algorithm outputs and subjective scores are minimized. Hence, better evaluation results have been accomplished.

In “Experimental Results” chapter, it is obvious that correlation results are better than previous work. By adding two new features into the NR algorithm, Pearson and Spearman correlation coefficients are increased. Thus, the optimization process has been successful.

CHAPTER II

DESIGN AND ANALYSIS

2.1 Blind Measurement of Blockiness Defect

2.1.1 Blockiness Model

In DCT-based coding, images are divided into NxN pixel blocks. N is generally equal to 8 in JPEG, MPEG, H.263 coding systems [3]. These blocks are individually processed. Every block has its own signal frequency parts after applying DCT computation. If there is a lot of quantization performed on these signal frequencies, blockiness impairment occurs.

In [26], a blockiness model is defined. DCT-compressed contents are considered as consisting of blocks that are distorted by independent and identically distributed white noise with zero mean and unknown variance. Two subsequent NxN blocks, both vertical and horizontal direction, can be formed as in the below:

$$b_1 = \mu_1 + \varepsilon_{i,j}\,,\; b_2 = \mu_2 + \delta_{i,j} \tag{2.1}$$

b_1 and b_2 are block model with average values μ_1 and μ_2. $\varepsilon_{i,j\ and}\ \delta_{i,j}$ are formed as white noise blocks with zero mean. As explained before, by applying large amounts of quantization on DCT coefficients, blockiness effect appears. If the quantization parameter is increased, $\varepsilon_{i,j}$ and $\delta_{i,j}$ become ineffective. Therefore, a new block can be generated by using the right half of b_1 and left half of b_2. It can be observed how a new shifted block is generated for both horizontal and vertical direction.

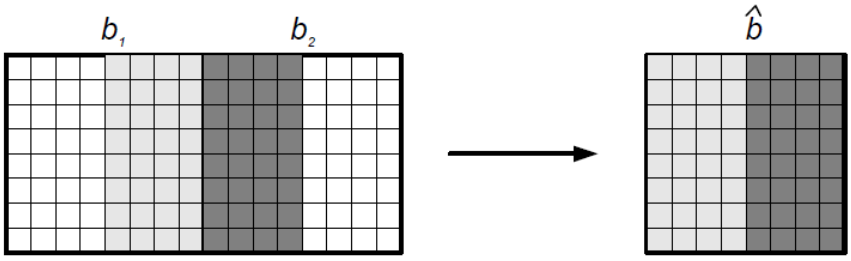

Figure 2.1 Horizontally generation of a new block [3]

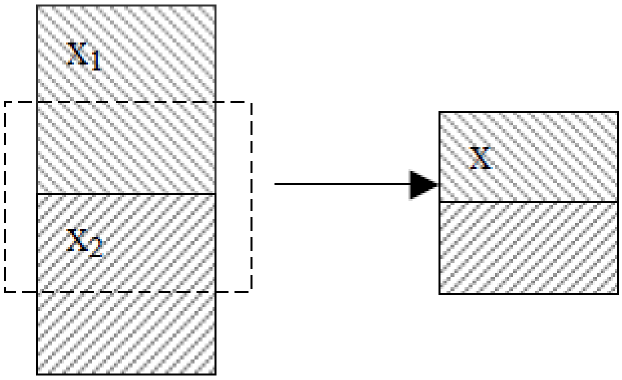

Figure 2.2 Vertically generation of a new block [12]

To construct a new block by using adjacent b_1 and b_2 blocks a 2-D auxiliary matrix, $s_{(i,j)}$ is defined as below [3]:

$$s = \begin{bmatrix} (-1/8) & (-1/8) & (-1/8) & (-1/8) & (1/8) & (1/8) & (1/8) & (1/8) \\ (-1/8) & (-1/8) & (-1/8) & (-1/8) & (1/8) & (1/8) & (1/8) & (1/8) \\ (-1/8) & (-1/8) & (-1/8) & (-1/8) & (1/8) & (1/8) & (1/8) & (1/8) \\ (-1/8) & (-1/8) & (-1/8) & (-1/8) & (1/8) & (1/8) & (1/8) & (1/8) \\ (-1/8) & (-1/8) & (-1/8) & (-1/8) & (1/8) & (1/8) & (1/8) & (1/8) \\ (-1/8) & (-1/8) & (-1/8) & (-1/8) & (1/8) & (1/8) & (1/8) & (1/8) \\ (-1/8) & (-1/8) & (-1/8) & (-1/8) & (1/8) & (1/8) & (1/8) & (1/8) \\ (-1/8) & (-1/8) & (-1/8) & (-1/8) & (1/8) & (1/8) & (1/8) & (1/8) \end{bmatrix} \quad (2.2)$$

The matrix above can be shaped by using MATLAB by using the code below:

```
s1 = (-1/8)*(ones(8,4)); % first part of s
s2 = (1/8)*(ones(8,4)); % second part of s
s = [s1,s2]; % 2D step function
```

So a new generated block can be defined as in the below:

$$b_{(i,j)} = \beta * s_{(i,j)} + \mu + r_{(i,j)} \quad (i, j = 0,\ldots,7) \qquad (2.3)$$

where $|\beta|$ is the amplitude of 2-D auxiliary function of s, μ is average value of shifted block named as b. Also, μ can be defined as local luminance value, r is described as residual function, which defines local activity around the block edge. Blockiness impairment between these two subsequent blocks can be easily detected by measuring the $|\beta|$ value in the function above [3].

2.1.2 Model in DCT-domain

To model blockiness edges, DCT-domain algorithm is used. Coefficients of new generated block are obtained by applying 2D-DCT computation. But first of all two adjacent NxN blocks must be mixed properly. To get a new generated block with a size of NxN pixels again, we need to take advantage of matrixes as in the below [3]:

$$q_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \; q_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{2.4}$$

To generate these matrixes, we can use the below MATLAB code:

```
q1 = [zeros(4,8) ; eye(4,4) , zeros(4,4)];
q2 = [zeros(4,4) , eye(4,4) ; zeros(4,8)];
```

By using the matrixes in (2.4), the generated block can be defined as:

$$b = b_1 * q_1 + b_2 * q_2 \tag{2.5}$$

Using the linear and distributive features of DCT calculation, a new equation can be formed [27]:

$$B = B_1 * Q_1 + B_2 * Q_2 \tag{2.6}$$

where B, B_1, B_2, Q_1 and Q_2 are 2-D DCT's of b, b_1, q_1, b_2 and q_2, respectively. To simplify the 2-D DCT calculations, it is possible to reshape the equation like:

$$B = B_1 * Q_1 + B_2 * Q_2 \tag{2.7}$$

$$= \frac{1}{2}[(B_1 + B_2) * (Q_1 + Q_2) + (B_1 - B_2) * (Q_1 - Q_2)]$$

If we re-write the equation above by substituting the expression as in the below, we can simplify the equation:

$$B_+ = B_1 + B_2$$

$$B_{-} = B_1 - B_2$$

$$F_{+} = Q_1 + Q_2$$

$$F_{-} = Q_1 - Q_2$$

$$B = \frac{1}{2}[(B_{+} * F_{+}) + (B_{-} * F_{-})] \tag{2.8}$$

B can be calculated using the MATLAB code below:

```
for J=1:coordinate_1
     for K=1:coordinate_2-1
          B_generated(:,:,J,K) = (1/2)*((Bpos_value(:,:,J,K) * Fpos) +
(Bneg_value(:,:,J,K) * Fneg));
     end
end
```

Also, 2-D DCT computation of 2-D auxiliary matrix defined in (2.2) has just four non-zero components. This offers great benefit while computing the blockiness artifact. In 2-D step function, the zeros are eliminated, and after then, computation is done. Hence, while running the algorithm, it gives faster response. By using MATLAB commands below, we can obtain 2-D DCT computed matrix as in the below:

```
S = dct2(s); % dct of step function (just four non-zero values)
```

and *S* will be:

$$S = \begin{bmatrix} 0 & -0.9061 & 0 & 0.3182 & 0 & -0.2126 & 0 & 0.1802 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{2.9}$$

To create simple notation, a vector named v is defined. v is the first row of the *S* and it can be described in MATLAB as in the below;

```
v = S(1,:); % taking just first row of the S matrix
```

Hereby, other parameters like μ and β can be calculated as:

$$\mu = B(0,0)/8 \tag{2.10}$$

$$\beta = \sum_{j=0}^{7} v_j * B(0,j)$$

$$= (v_1 * B(0,1)) + (v_3 * B(0,3)) + (v_5 * B(0,5)) + (v_7 * B(0,7)) \tag{2.11}$$

Therefore, the μ parameter can be easily computed in MATLAB:

```
for J=1:coordinate_1
    for K=1:coordinate_2-1
        u_generated(J,K) = B_value(1,1,J,K)/8;
    end
end
```

Likewise, β is defined:

```
for J=1:coordinate_1
   for K=1:coordinate_2-1
      beta_generated(J,K) = v(1,2).*B_value(1,2,J,K) +
      v(1,4).*B_value(1,4,J,K) + v(1,6).*B_value(1,6,J,K) +
      v(1,8).*B_value(1,8,J,K);
   end
end
```

The residual part of (2.3) can be obtained with the relation below:

$$R = B$$

$$R(0,0) = 0$$

$$R(0,i) = R(0,i) - (\beta * v_i) \quad (\text{i} = 0, \ldots, 7) \tag{2.12}$$

In MATLAB, the residuals can be determined as:

```
for J=1:coordinate_1
    for K=1:coordinate_2-1
        R_generated(1,1,J,K) = 0;
        R_generated(1,3,J,K) = 0;
        R_generated(1,5,J,K) = 0;
        R_generated(1,7,J,K) = 0;
        R_generated(1,2,J,K) = B_value(1,2,J,K) - (v(1,2)*beta_value(J,K));
        R_generated(1,4,J,K) = B_value(1,4,J,K) - (v(1,4)*beta_value(J,K));
        R_generated(1,6,J,K) = B_value(1,6,J,K) - (v(1,6)*beta_value(J,K));
        R_generated(1,8,J,K) = B_value(1,8,J,K) - (v(1,8)*beta_value(J,K));
    end
end
```

To compute blockiness artifact, developing the algorithm in DCT-domain is advantageous. Run-time gets shorter when the proposed technique is used. In the following section, how this proposed technique can be implemented in HVS measurement is explained.

2.1.3 Masking Techniques

In section 2.1.2, modeling a new shifted block is explained. After applying 2-D DCT on this new formed block, it is observed that there are only four AC elements mostly shared the signal energy. The first row of the matrix has all the energy as explained above, and these elements are (0, i), where i = 1,3,5,7, respectively [26]. The AC components are shown clearly in Figure 2.3.

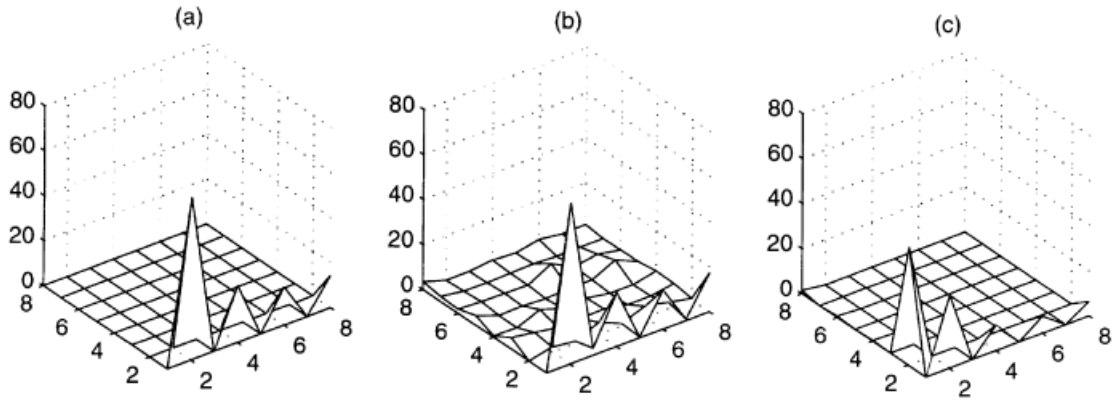

Figure 2.3 (a) $\sigma^2 = 0$ (noise free), (b) $\sigma^2 = 5$, (c) practical case [26]

Hence, it is understandable that blocking impairment is result of the major valued AC components.

In most areas, images or video contents are used by human end-users. Because of this, most of the works in the literature incorporate human psycho-visual systems (HVS). There are a lot of investigations about how a blockiness metric can be measured according to HVS, also which features are more effective on human observation.

Therefore, in this thesis HVS is considered so that we get more accurate results for blocking artifact. There have been two types of masking proposed, and these are most popular for HVS. One of them is texture (activity) masking and other one is luminance (brightness) masking properties. These are very sensitive on human perception [3].

First, activity masking technique is considered and computed. Some of the AC components in the DCT coefficients have more signal power and they cause blocking effect, it is good to take

an advantage of residual part explained in (2.12). The total activity of the residual matrix gives much benefit to obtain blockiness impairment [12]. Also, direction is important because blocking can appear in both way horizontally or vertically. In our proposed algorithm, horizontal activity is computed first and after than image is transposed and applied same operation. By taking transposition of image, vertical activity can be calculated. The activity is explained below:

$$A = \sum_{v=1}^{7} v \sum_{u=0}^{7} |R(u,v)| \tag{2.13}$$

In MATLAB, both vertical and horizontal activity can be determined using residual vector:

```
for J=1:coordinate_1
    for K=1:coordinate_2-1
        for j=1:m
            for k=1:m
             A_generated(J,K) = A_generated(J,K)+
            (j*abs(R_value(j,k,J,K)));
            end
        end
    end
end
```

This code above is run for each vertical and horizontal computation, and after then, an overall evaluation is done.

Secondly, another masking effect (brightness) is included. Local background luminance has effect on visibility of blockiness impairment. This property gives much clue for blockiness artifact calculation. It is possible to design separate vertical and horizontal blocking measurement systems. Visibility (η) in (2.13) can be calculated by using vertical activity or horizontal activity one by one. Otherwise, by mixing the two masking effects, a visibility coefficient can be obtained as below:

$$\eta = \frac{|\beta|}{(1+A)*(1+\left(\frac{\mu}{\mu_0}\right)*\gamma)} \tag{2.14}$$

where $\mu_0 = 150$ and $\gamma = 2$ [23].

It is an easy method to compute blockiness visibility coefficient. Thus, real-time assessments can be applicable by using a simple methodology [12].

Visibility can be simply calculated with the codes below:

```
for J=1:coordinate_1
```

```
    for K=1:coordinate_2-1
       n_generated(J,K) = abs(beta_value(J,K)) / ((1+A_value(J,K)) *
      (1+((gama*u_value(J,K))/u0)));
    end
end
```

So far, blockiness artifact of one shifted DCT block is calculated. However, blockiness metric is required for the whole picture or frame. Hence, to achieve an overall score, all of the DCT blocks must be combined as in:

$$\theta = \sqrt[p]{1/N \sum_{k=1}^{N} \eta_k^p} \tag{2.15}$$

where *N* is the number of entire DCT blocks in an image or frame, θ is a blockiness score and p is experimentally found as to be 4 [26].

2.2 Edge Detection Algorithm

Edges can be mostly found at the boundary of parts. In an image, two different neighboring objects create an edge. Edge detection algorithms are used to find out discontinuities of an image. These methods help to localize the sharp pixel transitions. Also, tracking an object, shape recognition, etc. can be managed by using edge detection applications.

There are some techniques that have been released for edge detection. These algorithms can be classified into two categories as: gradient and Laplacian methods. The gradient methods generally search minima and maxima points in the first derivative of the entire image data. Laplacian techniques try to find zero-crossings in the second derivative of the image to find edges. In Figure 2.4, it can be observed that there is some change in the intensity level and first derivative characteristics by passing through different contrast area of an image segment.

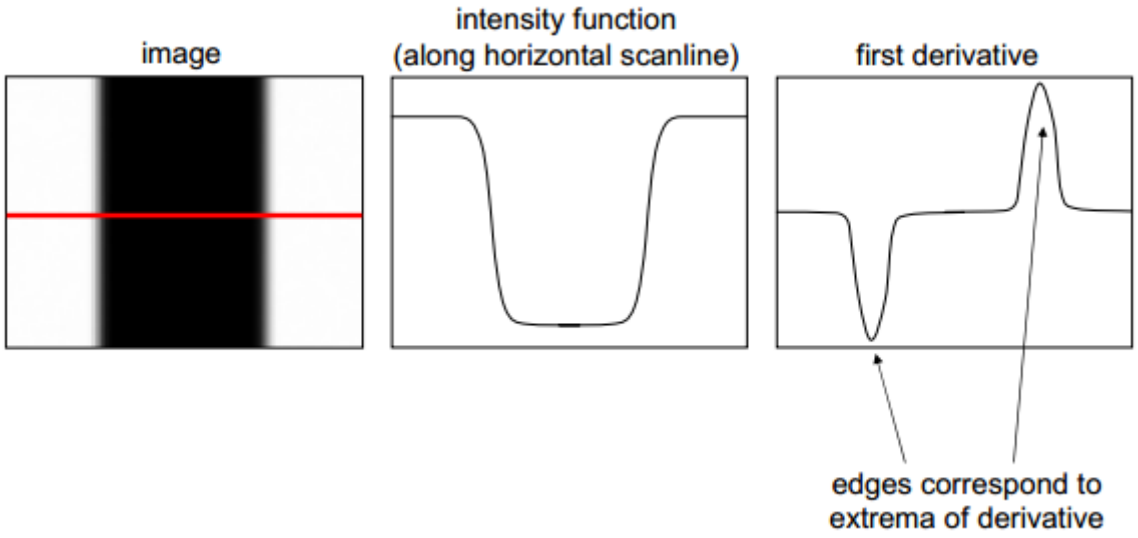

Figure 2.4 Changing intensity function and first derivative characteristics through an edge

Most known gradient methods are:

- Sobel [29,30]
- Prewitt [31]
- Roberts [32]
- Canny [33]

Most common and mostly used algorithm is Sobel operator as it has simplicity and can easily be implemented. Therefore, Sobel algorithm has great advantage in realization for real-time image quality assessment. Also, it has better performance comparing the other techniques.

Sobel operator applies 2-D gradient matrixes on an image. This provides to bring out high spatial frequency regions in the image, which also helps to find out edges. Standard Sobel edge detector has two mask operators. These matrixes consist of a pair of 3x3 blocks as in the below:

$$M_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix}, \; M_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \tag{2.16}$$

These gradient masks are generated to respond to the edges running vertically and horizontally. It is performed for both horizontal and vertical direction to the input image

separately. $\mathrm{M_x}$ is applied for vertical direction, while $\mathrm{M_y}$ is for horizontal. One matrix is the transpose of another one as shown above.

After applying these masks, gradient of the image can be obtained. To do that, convolution needs to be applied to masks and gray-scale image as shown below:

$$G_x = A_{i,j} .* M_x \tag{2.17}$$

where $A_{i,j}$ is image data and G_x is vertical direction derivative. By doing same calculation above, G_y can be calculated considering the horizontal gradient operator.

Convolution masks are often very much smaller than image data. For this reason, masks must be applied by shifting every step of multiplication. From left top to right bottom, every pixel must be swept by sliding the masks over the image data. In [28], convolution is well defined in the Figure 2.5:

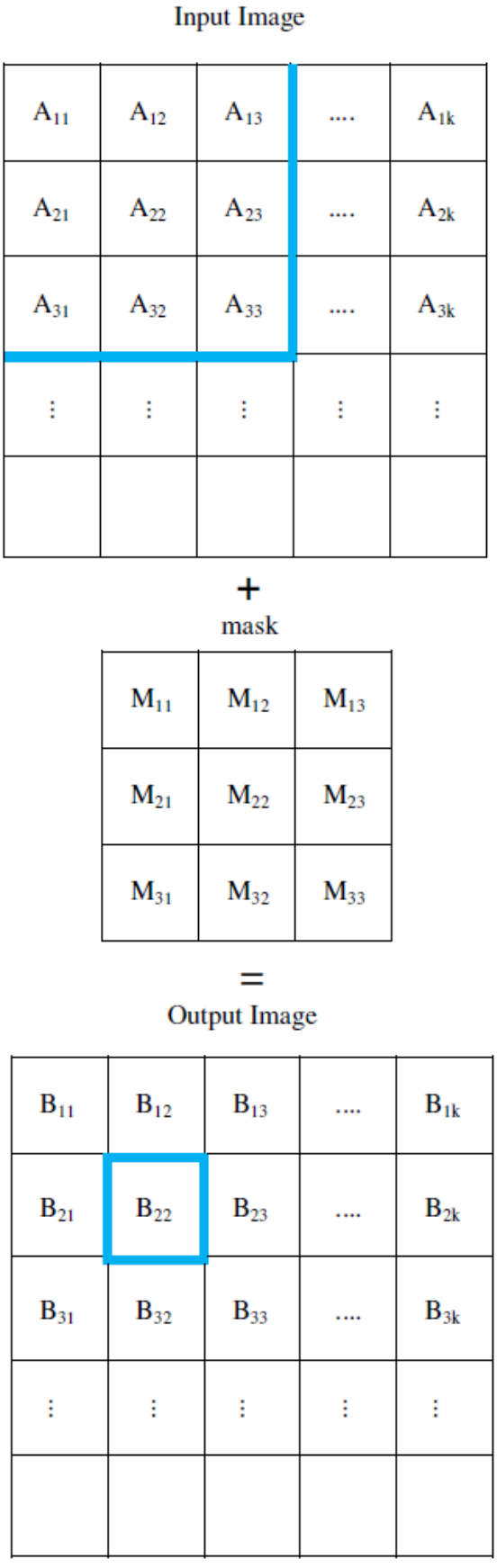

Figure 2.5 Performing convolution mask on the image data [28]

When mask is convolved in image, B_{22}can be calculated as follow:

$$B_{22} = (A_{11} * M_{11}) + (A_{12} * M_{12}) + (A_{13} * M_{13}) + (A_{21} * M_{21}) + (A_{22} * M_{22}) + (A_{23} * M_{23}) + (A_{31} * M_{31}) + (A_{32} * M_{32}) + (A_{33} * M_{33}) \quad (2.18)$$

The masks are performed one by one all over the image. Finally, to find out the gradient value the procedure is followed [28]:

$$G = \sqrt{G_x^2 + G_y^2} \quad (2.19)$$

The procedure is carried out for whole matrixes in image data. The edge detected image can be achieved from Sobel gradient by using a threshold value. In addition, if any Sobel gradient value is less than the predefined threshold value, then it will be replaced with the threshold value.

MATLAB is a powerful tool for computation and image processing applications. In this thesis, to calculate and find out edges in every image a MATLAB built-in function named as 'edge.m' is used. MATLAB command can be found below:

```
BW = edge(I2,'sobel'); %sobel operator edge detection
```

'BW' is logical data and consists 0 or 1 whether there is an edge or not. I2 is double formatted image data. Threshold value is chosen automatically by the 'edge.m' function. Or it can be adjusted by manually as in the below:

```
BW = edge(I2,'sobel','THRESHOLD_VALUE');
```

In the figures below, there are 'Monarch' images from LIVE database and Sobel edge detection algorithm is applied with different threshold values.

Figure 2.6 Gray-scale ‘Monarch.bmp’ image

Figure 2.7 Edge detected image, threshold= 0.3 (missed some details!)

Figure 2.8 Edge detected image, threshold= 0.3 (over-responded!)

Figure 2.9 Edge detected image, threshold= 0.1120 (automatically obtained)

The reason for using edge detection algorithm in this project is to eliminate the edges in the images or frames while calculating the blockiness metric. That is because miscalculations appear during the 2-D DCT computations. Generated by shifting two adjacent NxN blocks that have edges can be misjudged as if they have blockiness impairment. Therefore, before blockiness detection process, the parts of images which contain edges are removed and after than calculations begin. In the table below, percentage of eliminated block over an image for LIVE database can be obtained.

Table 2.1 Percentage of block elimination

	Horizontal %	Vertical %
Minimum	1,85	2,31
Maximum	9,87	9,7
Mean	4,74	5,01

2.3 Regression Method

Although the proposed image quality assessment technique in this thesis is based on 'blind measurement', this algorithm needs to be trained with a dataset. By using a training approach, image quality detection can be done more precisely. Especially, it offers much benefit to estimate quality scores according to human psycho-visual systems. Human perception must be considered while evaluating the image quality. That is because broadcasters and service providers are expected to deliver high-quality contents to their customers.

To verify no-reference algorithms, many image databases are generated. In these databases, images are evaluated by human observers and all human subjects give scores for the distorted images. Therefore, a no-reference quality assessment technique can be analyzed, and algorithm performance can be measured by using these datasets. Also, a training process helps to calibrate score evaluation.

There have been some databases released. These databases consist of distorted images, which are ranked by human subjective. One of them is MICT laboratory [34]. In this dataset, there are 196 images, all of them 24bit/pixel RGB colored and 768x512 pixels. Images are equally separated into two groups with respect to encoding types, JPEG and JPEG2000. 14 images for each are original. Another subjective quality assessment database is IVC [35]. 235 distorted images are generated by using 10 original images. 4 types of distortions are simulated in the

images. These distortion types are JPEG blockiness, JPEG2000 impairment, LAR coding and blurring. Subjective scores are up to 5 discrete quality scales and are provided by human observers for both image databases.

In this thesis, as training dataset, LIVE is used. 29 input images are used to create a database. In total, 233 images are generated with different distortion types. Plenty of papers refer to LIVE image datasets. Also, LIVE is used for training and testing in many works in the literature. Therefore, in this algorithm, in addition to performing training, LIVE is also used for testing.

Regression analysis is a popular technique and is generally used for prediction and forecasting. Therefore, a machine learning method can be applied by using regression analysis. It is a process to estimate relationships between variables. In this thesis, regression analysis performed to get accurate results.

Regression method is used to explore the relationship between two or more variables. It is a statistical technique that is very useful for these types of problems. Statistical technique is for investigating and modeling the relationship between the responses (dependent - Y) and explanatory (independent or predictor - X) variables. This analysis method is used in many research fields including medicine, economics, engineering and management. Regression models that involve one explanatory variable are called *simple* (bivariate) regressions. When two or more explanatory variables exist, the relationships are called *multiple* regressions. Regression models are also divided into *linear* and *non-linear* models, depending on whether the relationship between the response and explanatory variables is linear or nonlinear.

To explain the linear regression method briefly, there is a single regressor or predictor *x* and a dependent or response variable *Y*. The expected value of *Y* at each level of *x* is a random variable:

$$E(Y|x) = \beta_0 + \beta_1 * x \tag{2.20}$$

$$Y = \beta_0 + \beta_1 * x + \varepsilon \tag{2.21}$$

unknown parameters are:

- β_0 : intercept
- β_1 : slope

the assumed model for linear relationship is:

$$y_i = \beta_0 + \beta_1 * x_i + \varepsilon_i \text{ for all observations (i=1, 2, …, n)} \quad (2.22)$$

The error term is not observable and is assumed to be normally distributed with mean of 0 and standard deviation σ.

The fitted model used to predict the expected value of *Y* for a given value of *x* can be stated as:

$$\hat{y}_i = b_0 + b_1 * x_i \quad (2.23)$$

The coefficients are:

- b_0: the estimated intercept
- b_1: the estimated slope

The aim is obtaining smaller residual value. The residual is the difference of real value and the estimated value as in the below:

$$e_i = y_i - \hat{y}_i \quad (2.24)$$

An example for simple regression plot with predicted line and errors is:

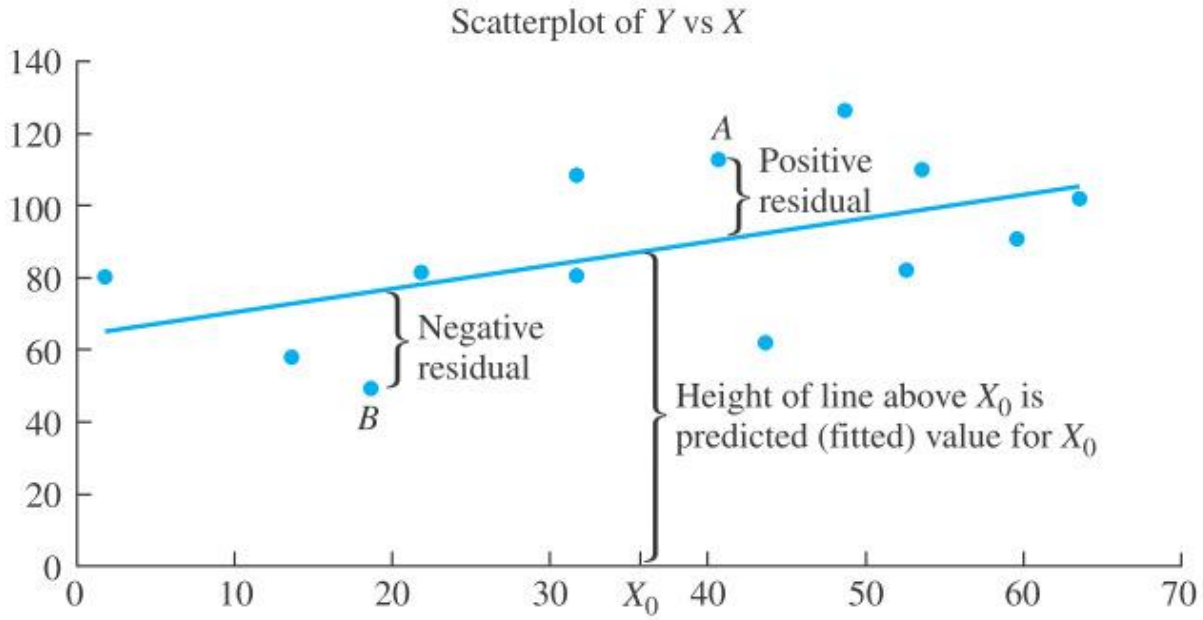

Figure 2.10 A scatter plot obtained from regression model

Microsoft Excel has a powerful add-in and it can be activated in Excel options. Analysis ToolPak application is used to perform regression analysis. Analysis ToolPak consists of a regression analysis tool (usually used in statistical and engineering management). In this thesis, this add-in is also used to implement a machine learning method.

Before applying regression analysis, the characteristic of the relationship between MOS subjective scores and algorithm outputs must be examined. Hence, it gives a lot of clue about regression equation. In Figure 11 and 12, it can be observed that algorithm outputs for horizontal and vertical blockiness artifact searching show non-linear characteristics.

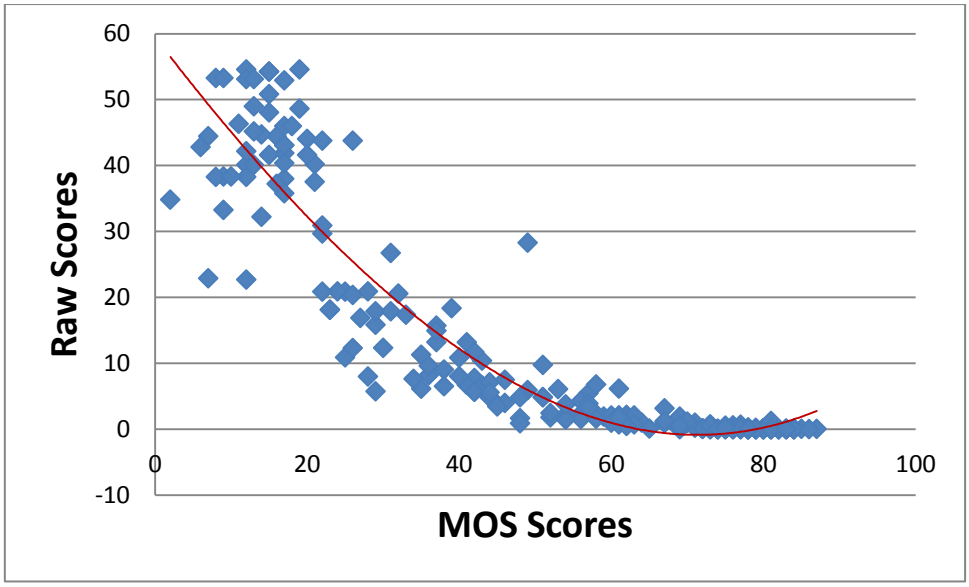

Figure 2.11 Raw blockiness scores vs. MOS scores for horizontal blockiness scanning

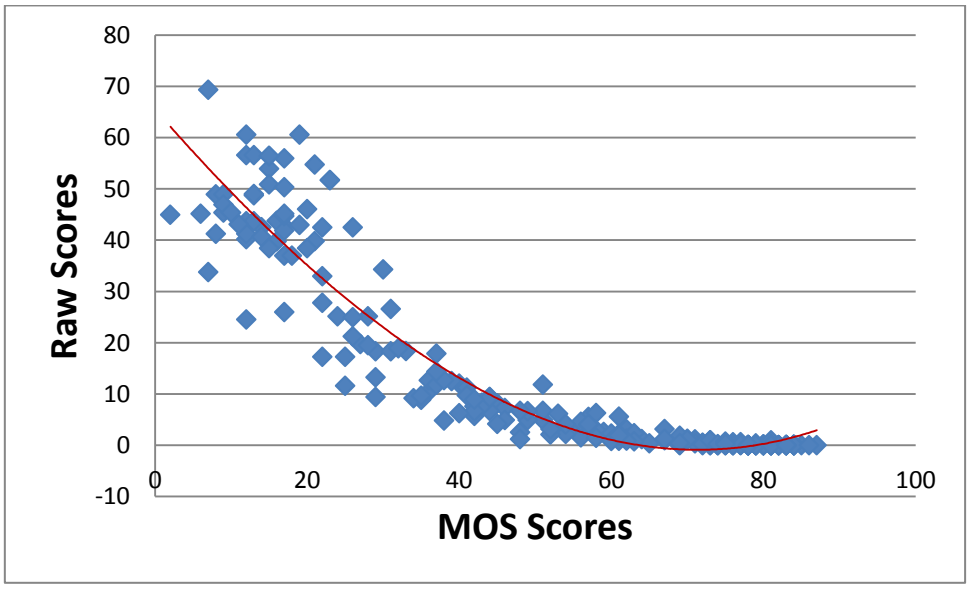

Figure 2.12 Raw blockiness scores vs. MOS scores for vertical blockiness scanning

Due to the fact that both scatter plots exhibit a nonlinear waveform, a non-linear regression analysis must be applied. Linear regression cannot perform well, and hence, image quality assessment is effected badly.

On the other hand, non-linear regression is not available in Microsoft Excel Analysis ToolPak. Also, the non-linear regression method is hard and very complicated to implement. However, if a new variable can be created and adapted to linear relationship, linear regression method that has been explained above can be easily constructed. Therefore, new variables can be replaced with original values, and a new linear equation can be explained.

After experimentation, average subjective scores (MOS values) are found to be proportional with cube root of both raw scores. If we take the cube root of horizontal and vertical raw blockiness metrics, we construct a new linear equation. These new root values are now our new original values. In the Figure 13 and 14, it can be observed that new trend lines show linear characteristics.

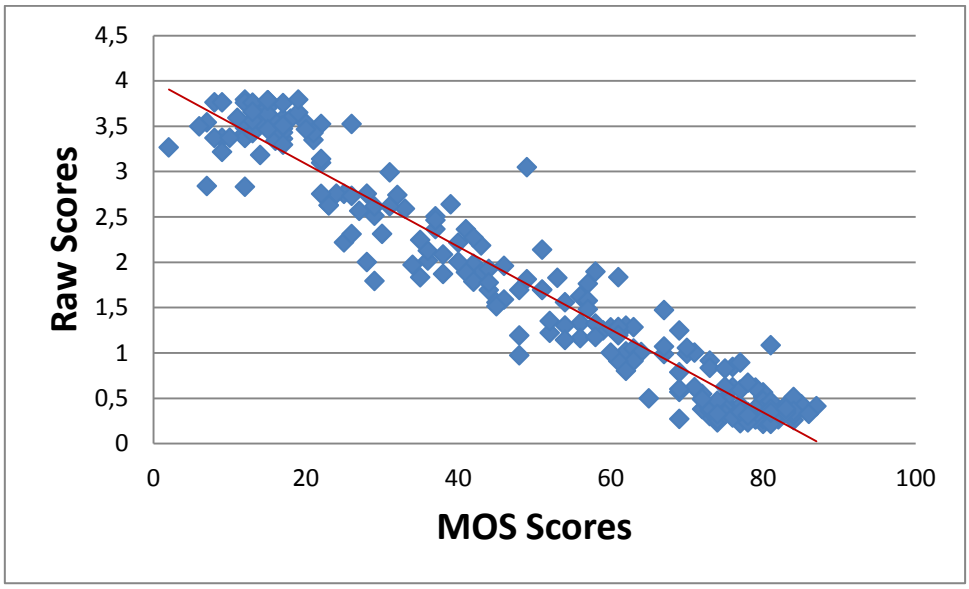

Figure 2.13 New blockiness scores vs. MOS scores for horizontal blockiness scanning

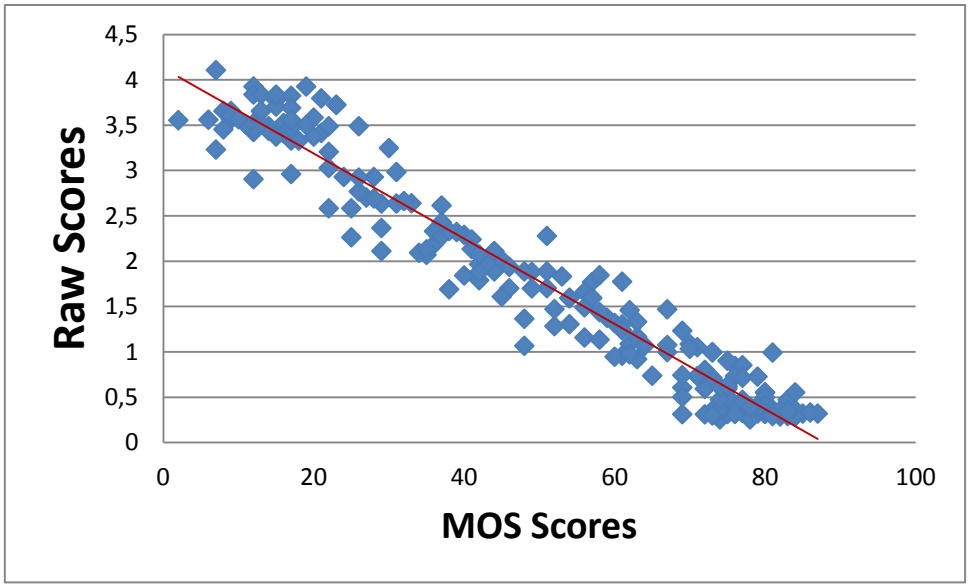

Figure 2.14 New blockiness scores vs. MOS scores for vertical blockiness scanning

After performing regression analysis to new blockiness metrics the coefficients are:

Table 2.2 Coefficient values after applying regression analysis

	Horizontal Coefficients	Vertical Coefficients
b_0	84,5598	84,8184
b_1	-20,8011	-20,4743

By applying new coefficients, scattering plots are become:

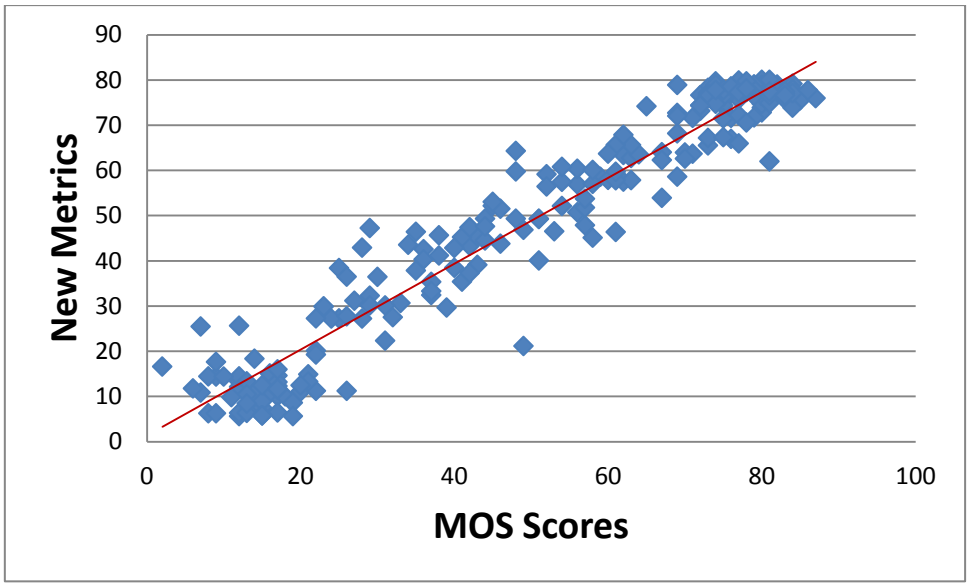

Figure 2.15 Scattering plot for horizontal blockiness scanning after new coefficients

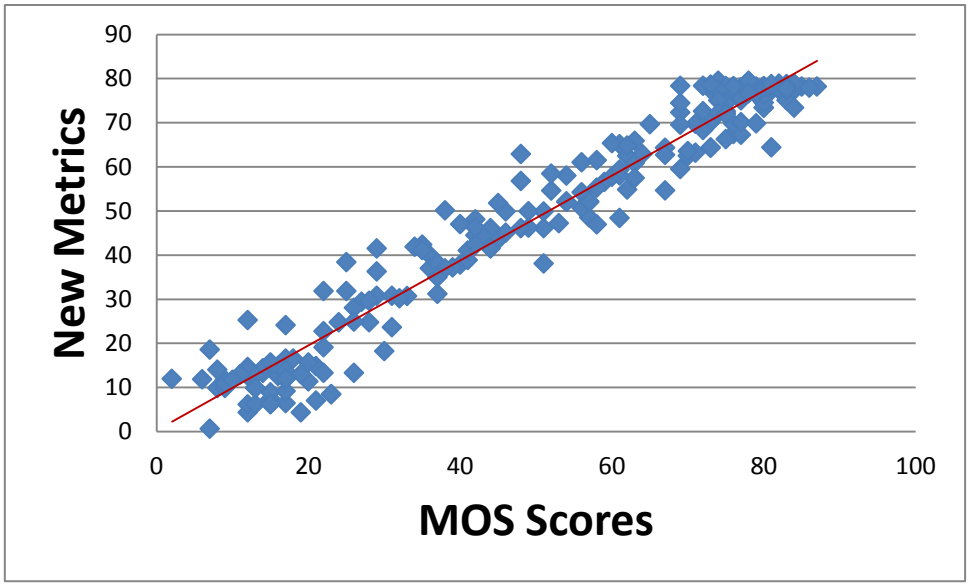

Figure 2.16 Scattering plot for vertical blockiness scanning after new coefficients

After all, we have two calculated metric values. To mix both vertical and horizontal scores, we need to apply multiple regression analysis lastly. Multiple regression analysis is an extension of bivariate regression to include more than one independent variable. It can be expressed like:

$$\hat{y}_i = b_0 + b_1 * x_i + b_2 * y_i \tag{2.25}$$

The coefficients are:

- b_0: the estimated intercept
- b_1: the estimated slope for 1st independent variable
- b_2: the estimated slope for 2nd independent variable

Multiple regression analysis can also be computed by Excel tool. After performing regression, coefficients are $b_0 = 1{,}7957$, $b_1 = 0{,}2077$, $b_2 = 0{,}7878$. Overall metric can be calculated by using new coefficients. Scatter plot for overall score is as in the Figure 2.17.

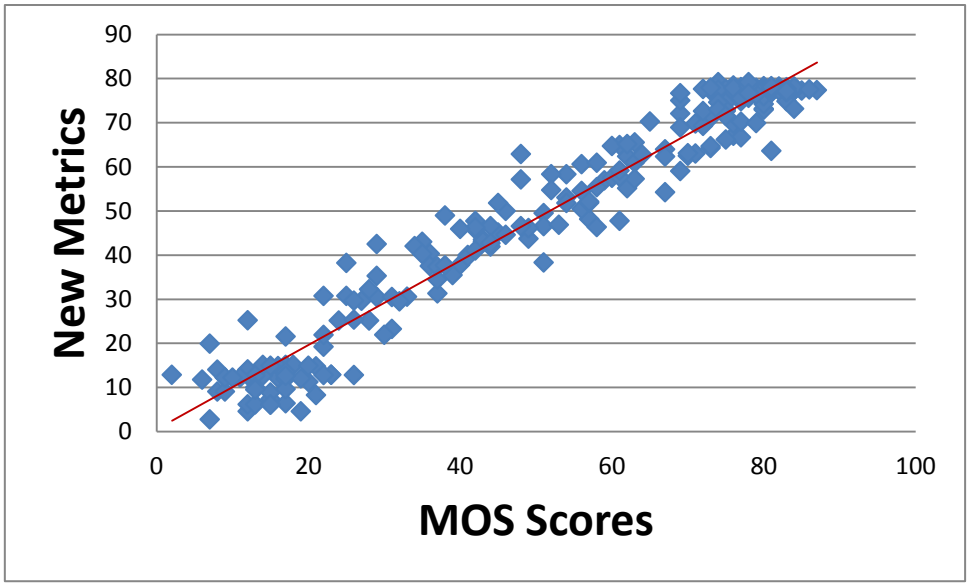

Figure 2.17 Scattering plot for overall metric

In the Experimental Results section, it will be explained that overall metrics have the best correlation values. By using simple and multiple regression methods, more precise results are obtained.

CHAPTER III

EXPERIMENTAL RESULTS

JPEG and MPEGs are DCT-based coding techniques. These coding techniques are lossy types. Because some amount of quantization parameter is applied to image or video data. Therefore, blockiness effect occurs depending on the quantization coefficient. The more quantization is applied on DCT coefficients the more blockiness distortion appears. In other words, blockiness impairment can be seen at the block boundaries because of discontinuity and quantized independently [15].

In this chapter, consistency of the algorithm output is evaluated by using Pearson and Spearman correlation technique. To evaluate scores, still image dataset which is available in the LIVE image quality assessment database release 2 are used. As explained in section 2.3 before, LIVE release 2 database has 233 images. 29 of 233 images are original (uncompressed) and the rest of them compressed with different compression ratios. These images are shown to observers randomly. Observers rank the test images from 1 to 100 degrees. By averaging these scores, MOS values can be obtained.

To measure correlation MOS values and algorithm outputs are used. Types of correlations can be categorized according to their relation between variables as in the below;

- Positive correlation: one of the variables is increased by increasing the other variable.
- Negative correlation: one of the variables is decreased by increasing the other variable.
- No correlation: one of the variables does not have trend by increasing or decreasing the other variable.

In the scatter plots below types of correlations are shown;

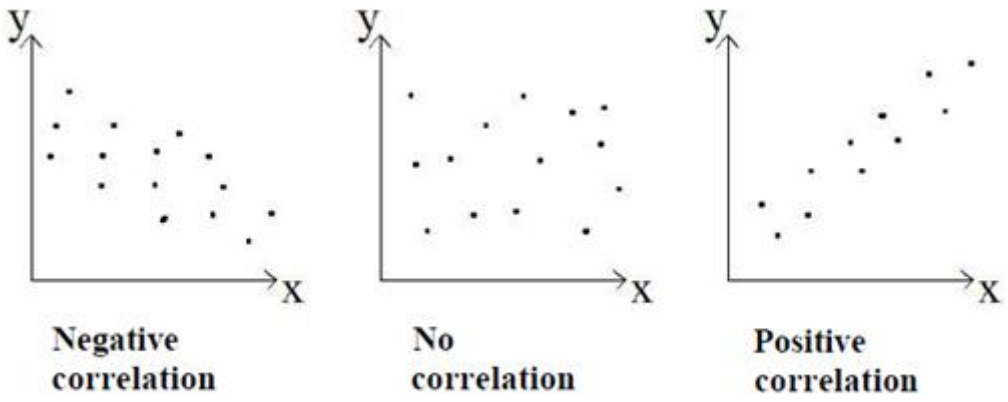

Figure 3.1 Negative, no and positive correlations

3.1 Pearson and Spearman Rank Order Correlations

3.1.1 Pearson Correlation

Pearson's correlation is a technique to measure linear relationship between two variables. Correlation coefficient is denoted as r. r has a limitation as in the below;

$$-1 \leq r \leq 1 \tag{3.1}$$

Positive correlation coefficients mean there is a positive linear correlation. Therefore negative coefficients indicate there is a negative linear relationship. If there is no correlation r value becomes 0. Consequently a value of close to 1 or -1 means strong correlation between the variables. When r is equal to 1 or -1 it is can be said that there is a perfect correlation between variables.

The Pearson's correlation coefficient can be calculated as in the below;

$$r = \frac{\sum_i (x_i - \bar{x}) * (y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} * \sqrt{\sum_i (y_i - \bar{y})^2}} \tag{3.2}$$

where x_i denotes i^{th} element of the first variable and $\bar{x}$ means average value of the all first variables. Likewise y_i denotes i^{th} element of the second variable and $\bar{y}$ means average value of the all second variables.

There are 233 JPEG coded images and some of the images have blockiness defect in LIVE database. To calculate Pearson correlation MOS values of the images are used as one of the variables. Other variable is output score of algorithm for every image.

3.1.2 Spearman Rank Order Correlation

Spearman's rank order correlation is used frequently in statistic. It is a non-parametric measurement technique of correlation. Unlike the Pearson correlation, Spearman does not require a monotonic behavior between variables. In other words, this technique can be also used if there is a non-linear relationship between variables. In the figure below, monotonic and non-monotonic relationships is shown;

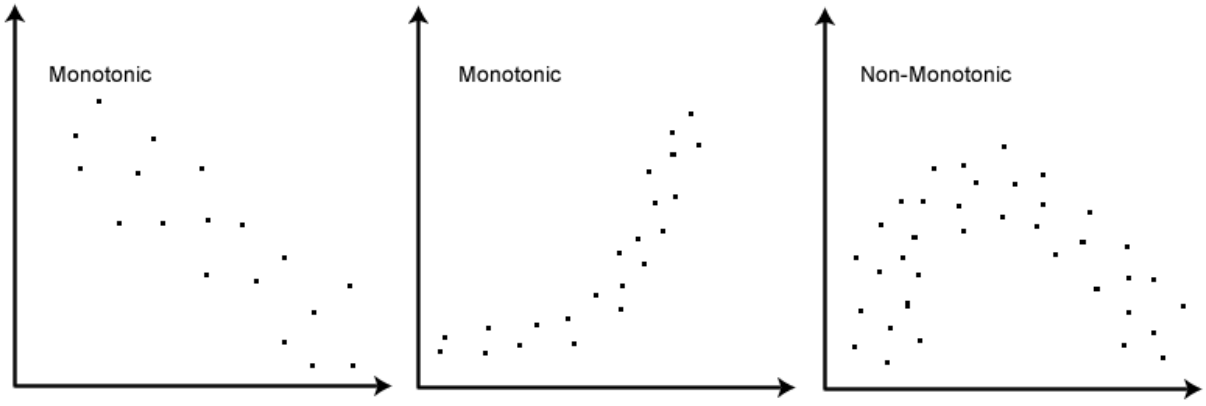

Figure 3.2 Monotonic and non-monotonic relationships

Spearman's correlation technique is based on re-ordering the data. To do that, all of the data are ranked. So, every element of variables is measured at the ordinal level.

Spearman correlation coefficient can be calculated as in the below;

$$r = 1 - \frac{\sum d^2}{n*(n^2-1)} \qquad (3.3)$$

where d is difference between the two numbers in each pair of ranks and n is number of pairs of data.

Similar to Pearson correlation when Spearman correlation coefficient (r) is close to 1 or -1 that means there is a strong correlation between variables. If r is around 0 so correlation is very weak.

3.2 Evaluating the Results with Objective and Subjective Scores

In this chapter, output of algorithm is evaluated by using bits/pixel and MOS values of images. Using bits/pixel value of images provides an objective evaluation. For subjective comparison MOS values are assured from LIVE database.

Evaluation process is divided into 2 categories. Category 1 can be considered as objective assessment. Pearson correlation coefficients are investigated by using bits/pixel ratios and algorithm metrics in this category. Bits/pixel ratio is directly affecting the blockiness defect in an image. In the Figure 3.3 and 3.4, blockiness impairment can be easily seen by reducing the bits/pixel ratio.

Figure 3.3 ‘rapids’ image in LIVE database with bpp=3.815

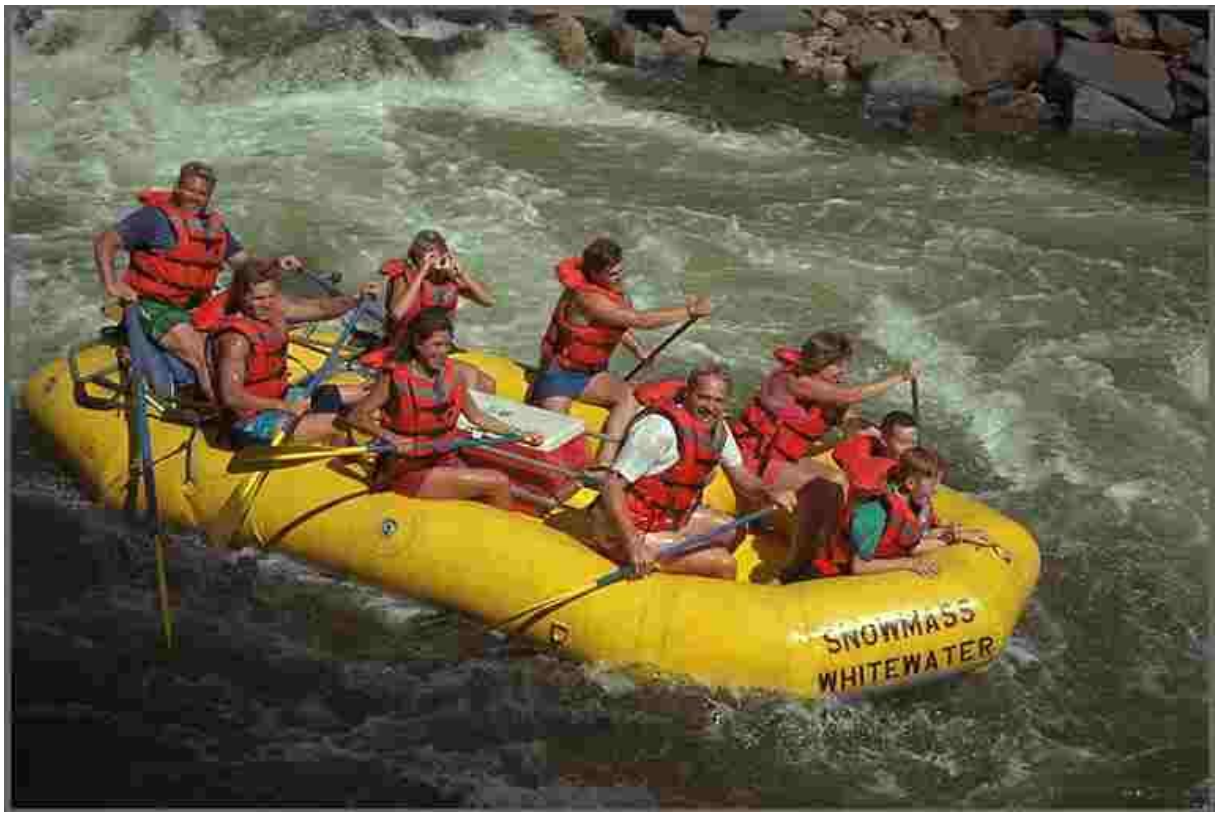

Figure 3.4 ‘rapids’ image in LIVE database with bpp=0.627

The papers which test their approach with LIVE image database and evaluate their result by using Pearson and Spearman correlation technique are gathered in category 2. MOS values and scores obtained from algorithms are compared, so subjective assessment is done.

3.2.1 Category 1

In the category 1, bits/pixel ratio is used for comparing algorithm results. The works in [3] and [12] are considered and compared because these algorithms are using same shifted blocking technique with the proposed approach in this thesis.

The method proposed in paper [3] has blocking metrics as in the table below:

Table 3.1 Measure of blocking artifacts of JPEG-coded images [3]

Images	JPEG coded bit rate (bits/pixel)												Pearson ranking
	0,15	0,2	0,25	0,3	0,35	0,4	0,45	0,5	0,75	1	1,5	2	
Barbara	21,014	15,860	14,130	6,317	3,448	2,548	2,265	1,857	0,830	0,404	0,101	0,018	-0,615
Boats	16,282	10,176	4,288	3,096	2,545	2,011	1,717	1,308	0,594	0,362	0,031	0,011	-0,560
Bridge	19,903	12,762	9,878	8,619	4,900	4,053	2,101	2,092	0,722	0,296	0,012	0,010	-0,661
Goldhill	15,966	8,918	5,970	4,488	2,083	1,499	1,307	1,095	0,253	0,131	0,011	0,010	-0,574
Mandrill	15,673	12,356	9,214	6,346	3,998	3,129	2,806	1,917	0,680	0,365	0,006	0,005	-0,677
Peppers	16,872	10,312	4,281	3,003	2,331	1,842	1,375	1,145	0,354	0,158	0,013	0,011	-0,544
Zelda	12,438	5,284	2,983	1,957	1,343	0,815	0,646	0,531	0,200	0,030	0,013	0,011	-0,501
Average Pearson Rank :													**-0,590**

In Table 3.1, different bits/pixel valued images are evaluated with blockiness metric. After calculating average Pearson correlation coefficient is obtained -0.59.

Table 3.2 Measure of blocking artifacts for our proposed technique

Images	JPEG coded bit rate (bits/pixel)												Pearson Ranking
	0,15	0,2	0,25	0,3	0,35	0,4	0,45	0,5	0,75	1	1,5	2	
Barbara	20,671	35,777	42,530	47,905	53,477	55,555	58,184	60,485	64,075	77,411	77,411	79,098	0,826
Boats	30,906	37,347	43,930	48,494	51,922	56,817	59,043	61,958	63,366	76,880	77,411	76,845	0,845
Bridge	37,428	58,302	69,764	70,640	74,490	72,978	74,489	77,367	77,316	77,310	77,411	77,687	0,502
Goldhill	31,022	39,315	44,184	49,000	52,390	54,396	59,042	61,500	63,471	66,248	77,411	70,046	0,815
Mandrill	32,324	40,157	45,328	49,104	52,906	54,792	56,510	56,508	59,354	69,451	77,411	76,957	0,902
Peppers	13,264	25,753	36,169	41,562	47,103	53,055	54,346	56,098	61,952	78,195	77,411	79,081	0,834
Zelda	31,527	37,214	44,013	51,569	54,373	56,201	57,962	61,227	64,045	69,632	77,411	77,626	0,864
Average Pearson Rank :													**0,798**

In the Table 3.2, proposed algorithm has more correlated result from paper [3] by using same test images and bits/pixel ratios. Average Pearson rank is found as 0.798.

Table 3.3 Measure of blocking artifacts of JPEG-coded images [12]

Images	JPEG coded bit rate (bits/pixel)											Pearson Ranking
	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1	1,5	2	
Barbara	19,250	7,915	6,010	2,232	1,519	1,118	0,854	0,783	0,523	0,018	0,013	-0,618
Pepper	19,390	6,102	3,346	1,794	1,232	0,740	0,634	0,352	0,014	0,011	0,010	-0,557
Cameraman	25,760	7,457	2,888	2,068	1,727	1,362	0,796	0,628	0,397	0,177	0,009	-0,532
Irisflower	15,350	6,885	3,426	2,181	1,491	0,892	0,462	0,452	0,235	0,009	0,007	-0,614
Lab	14,260	5,258	3,061	2,091	1,457	1,118	0,844	0,759	0,566	0,123	0,011	-0,605
Cane	13,910	5,171	0,613	0,385	0,309	0,023	0,020	0,019	0,018	0,015	0,013	-0,501
Eggs	7,650	1,034	0,610	0,142	0,031	0,022	0,020	0,018	0,016	0,014	0,013	-0,447
Plank	5,146	2,161	1,322	0,805	0,490	0,324	0,102	0,100	0,007	0,005	0,004	-0,618
Eveface	5,408	1,848	0,910	0,431	0,299	0,031	0,014	0,012	0,012	0,011	0,010	-0,547
Tower	5,061	2,229	1,314	0,660	0,619	0,271	0,010	0,006	0,007	0,004	0,004	-0,616
Average Pearson Rank :												**-0,566**

In Table 3.3 another approach [12] is considered with different number of images and bits/pixel ratios. Average Pearson score is -0.566 that is very similar result with paper [3].

Table 3.4 Measure of blocking artifacts for proposed technique

Images	JPEG coded bit rate (bits/pixel)											Pearson Ranking
	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1	1,5	2	
Barbara	20,671	35,777	42,530	47,905	53,477	55,555	58,184	60,485	64,075	77,411	79,098	0,916
Pepper	30,906	37,347	43,930	48,494	51,922	56,817	59,043	61,958	63,366	76,880	76,845	0,936
Cameraman	37,428	58,302	69,764	70,640	74,490	72,978	74,489	77,367	77,316	77,310	77,687	0,624
Irisflower	31,022	39,315	44,184	49,000	52,390	54,396	59,042	61,500	63,471	66,248	70,046	0,887
Lab	32,324	40,157	45,328	49,104	52,906	54,792	56,510	56,508	59,354	69,451	76,957	0,961
Cane	13,264	25,753	36,169	41,562	47,103	53,055	54,346	56,098	61,952	78,195	79,081	0,925
Eggs	31,527	37,214	44,013	51,569	54,373	56,201	57,962	61,227	64,045	69,632	77,626	0,931
Plank	29,056	44,571	51,576	56,803	60,639	64,242	66,703	74,719	76,969	77,842	79,180	0,829
Eveface	33,009	47,106	54,348	62,306	75,436	76,591	77,067	77,917	78,071	78,685	78,850	0,702
Tower	23,688	34,879	51,779	55,783	58,888	61,350	62,574	64,552	67,579	74,648	76,116	0,829
Average Pearson Rank :												**0,854**

In Table 3.4, same images with paper [12] are applied to proposed algorithm. Output is 0.854 which is better than algorithm number [12].

Category 1 is explained that proposed algorithm has better performance if bits/pixel ratio is considered. By using bits/pixel ratios, objective evaluation is accomplished. According to

tables above proposed algorithm has more correlated results comparing with other two algorithms.

3.2.2 Category 2

In the second category, the papers which are used LIVE image database and their subjective scores are considered. Furthermore these approaches are also evaluated their results by taking Pearson and Spearman correlation technique.

[1], [14], [15], [16], [20], and [22] papers have Pearson correlation coefficients in their result sections. Also, in the papers of [14], [15], [16], [20], [21], and [22], Spearman coefficients are available to use comparison between proposed technique in this thesis.

Using Microsoft Excel, Pearson and Spearman correlations can be computed. Formulas are easily generated for both correlation techniques. After acquiring output scores of proposed algorithm, Pearson correlation is found as 0.9774 and Spearman correlation is obtained as 0.9518.

Table 3.5 Pearson correlation coefficient comparison

Paper #	Pearson Ranking
1	0,92
14	-0,843
15	0,944
16	-0,727
20	0,9011
22	0,921
Proposed technique	**0,9774**

Table 3.6 Spearman correlation coefficient comparison

Paper #	Spearman Ranking
14	0,838
15	0,937
16	0,925
20	0,8914
21	0,9411
22	0,91
Proposed technique	**0,9518**

In Table 3.5 and 3.6, it is shown that proposed technique has better performance from other approaches. Tables also show that, remarkable improvement has been achieved. Additionally about 0.3 point of improvement has been attained considering the closest approach [15] in Pearson correlation coefficient. For Spearman coefficient there have been about 0.1 point improvement obtained. Consequently, proposed approach has better correlations with subjective test scores acquired from LIVE database.

Sobel edge detection and regression analysis provide great advantage to achieve more correlated result. If edge detection function is disabled, Pearson and Spearman coefficients would be as in the table below:

Table 3.7 Correlations without edge detection and regression methods

	Pearson Ranking	Spearman Ranking
Without Sobel Edge Detection	0,9494	0,9321
Without Regression Model	-0,8842	-0,9279

CHAPTER IV

CONCLUSIONS and FUTURE WORK

In this thesis, a blind measurement technique of blockiness defect has been proposed. In DCT-based coded images, blockiness artifacts have been detected successfully. Objective and subjective evaluation have been applied. Variety of images with different blockiness degrees and bits/pixel ratios have been tested. After an evaluation, it is proved that the proposed model is consistent and balanced. Also, compared to other blockiness measurement approaches, the proposed technique has better performance.

Compared to other approaches, proposed method has more correlated results. In 'Experimental Results' chapter, results show that this method performs better than other methods. The algorithm has been tested by a set of images, quality of which was scored by human observers and also different bits/pixels valued images.

In addition to having the best performance, this algorithm is dynamic and easily implemented. Proposed technique is very capable of being used in real-time applications. It is suitable for online quality controlling and monitoring. Broadcasters or internet service providers can use this technique to inspect their streaming systems. Applying this model in their systems can be a great benefit to examine the QoS.

By merging Sobel edge detection and using a regression model, performance has increased strongly. These methods provide more correlated results with blockiness metrics of algorithm output and subjective scores. Also, objective evaluation is supported and improved much.

As future work, it is planned to do an FPGA design. If a hardware tool is made to measure blockiness metric it will provide much benefit in field tests.

BIBLIOGRAPHY

[1] R. Muijs and I. Kirenko, "A no-reference blocking artifact measure for adaptive video processing", *European Signal Processing Conference*, Antalya, Turkey, 2005.

[2] F. Yang, S. Wan and Y. Chang et al., "A novel objective no-reference metric for digital video quality assessment", *IEEE Signal Processing Letters*, 12 (10), pp.685-689, 2005.

[3] S. Liu and A. C. Bovik, "Efficient DCT-Domain Blind Measurement and Reduction of Blocking Artifacts", *IEEE Trans.on CSVT* 12(12), 1139–1149, 2002.

[4] J. Caviedes and J. Jung, "No-reference metric for a video quality control loop", *Proc. Int. Conf. Information Systems, Analysis and Synthesis*, vol. 13, 2001.

[5] C. Keimel, T. Oelbaum and K. Diepold, "No-reference video quality evaluation for high definition video", *IEEE International Conference, Acoustics, Speech and Signal Processing*, 2009 (ICASSP 2009), Taipei, Taiwan, pp. 1145-1148. *IEEE*, Washington DC, 2009.

[6] M. Farias and S. Mitra, "No-reference video quality metric based on artifact measurements", *Proc. IEEE ICIP*, vol. 3, pp. 141–144, 2005.

[7] S. Wolf and M. H. Pinson, A. A. Webster and G. W. Cermak, "Objective and Subjective Measures of MPEG Video Quality", *139th SMPTE Technical Conference*, New York City, November 21-24 1997.

[8] S. Winkler, A. Sharma, and D. McNally, "Perceptual video quality and blockiness metrics for multimedia streaming applications", *Proc. International Symposium on Wireless Personal Multimedia Communications*, Aalborg, Denmark, pp. 547–552, invited paper, September 9–12, 2001.

[9] S. Winkler and P. Mohandas, "The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics," *IEEE Transactions on Broadcasting*, Vol. 54, No. 3, pp 1-9, September 2008.

[10] S. Wolf and M. Pinson, "Video quality measurement techniques," *NTIA Report 02-392*, June 2002. Available: www.its.bldrdoc.gov/n3/video/documents.htm

[11] S. Wolf and M. Pinson, “Video quality measurement techniques”, U.S. Dept. of Commerce, National Telecommunications and Information Administration, 2002. Available: http://purl.access.gpo.gov/GPO/LPS31510

[12] A. C. Bovik and S. Liu, “DCT-domain blind measurement of blocking artifacts in DCT-coded images”, *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 3, pp. 1725 –1728, May 2001.

[13] Z. Wang, A. C. Bovik and B. L. Evans, “Blind measurement of blocking artifacts in images”, *Proc. ICIP*, vol. 3, pp. 981–984, Vancouver, Canada, 2000.

[14] Md. Mehedi Hasan, K. Ahn, O. Chae, ”Measuring Blockiness of Videos using Edge Enhancement Filtering”, *International Conference on Signal Processing, Image Processing and Pattern Recognition*, CCIS 260, pp. 10-19, 2011.

[15] Z. Wang, H.R. Sheikh and A. C. Bovik, “No-Reference Perceptual Quality Assessment of JPEG Compressed Images”, *IEEE International Conference on Image Processing*, pp. 477–480, 2002.

[16] H.R. Wu and M. Yuen, “A generalized block edge impairment metric for video coding”, *IEEE Signal Processing Letters*, vol. 4, no. 11, pp. 317-320, November 1997.

[17] K. T. Tan and M. Ghanbari, “Blockiness Detection for MPEG2-coded Video”, *IEEE Signal Processing Letters*, Vol. 7, No. 8, pp. 213-215, August, 2000.

[18] P. Le Callet, C. Viard-Gaudin, S. Pechard, and E. Caillault, "No reference and reduced reference video quality metrics for end to end QoS monitoring," *IEICE Transactions on Communications*, vol. E89-B, No. 2, pp. 289-296, 2006.

[19] Z. Wang and A. C. Bovik, “Reduced- and no-reference image quality assessment: The natural scene statistic model approach,” *IEEE Signal Process. Mag.*, vol. 28, pp. 29–40, November 2011.

[20] A. K. Moorthy and A. C. Bovik, “A two-step framework for constructing blind image quality indices”, *IEEE Signal Processing Letter*, vol. 17, no. 5, pp. 513–516, May 2010.

[21] M. A. Saad, A. C. Bovik, and C. Charrier, “Model-based blind image quality assessment using natural DCT statistics,” *IEEE Trans. Image Process*, to be published, 2011.

[22] A. K. Moorthy and A. C. Bovik, “Blind image quality assessment: From natural scene statistics to perceptual quality”, *IEEE Trans. Image Process*, vol. 20, no. 12, pp. 3350–3364, December 2011.

[23] S. A. Karunasekera and N. G. Kingsbury, “A distortion measure for blocking artifacts in image based on human visual sensitivity", *IEEE Trans. on Image Processing*, vol. 4, pp. 713-724, June 1995.

[24] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "A comprehensive evaluation of full reference image quality assessment algorithms", *Proc. ICIP*, pp. 1477-1480, 2012.

[25] H.R. Sheikh, Z. Wang, L. Cormack and A.C. Bovik, “LIVE Image Quality Assessment Database”. Available: http://live.ece.utexas.edu/research/quality.

[26] B. Zeng, “Reduction of blocking effect in DCT-coded images using zero-masking techniques”, *Signal Processing 79,* 205-211, 1999.

[27] S. F. Chang and D. G. Messerschmitt, “Manipulation and Compositing of MC-DCT Compressed Video”, *IEEE Journal on Selected areas in Communications*, vol. 13, no.1, January 1995.

[28] S. Gupta and S. G. Mazumdar, “Sobel Edge Detection Algorithm”, *International Journal of Computer Science and Management Research*, vol. 2, February 2013.

[29] I. Sobel, “An Isotropic 3×3 Gradient Operator, Machine Vision for Three – Dimensional Scenes”, *H. Freeman Academic Pres*, NY, 376-379, 1990.

[30] I. Sobel, “Camera Models and Perception”, Ph.D. thesis, Stanford University, Stanford, CA, 1970.

[31] J. Prewitt, “Object Enhancement and Extraction, Picture Processing and Psychopictorics “, B. Lipkin and A. Rosenfeld, Ed., NY, Academic Pres, 1970.

[32] L. G. Roberts, “Machine Perception of Three-Dimensional Solids, in optical and Electro-Optical Information Processing”, J. Tippett, Ed., 159-197, MIT Press, 1965.

[33] J. Canny, “A Computational Approach to Edge Detection”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8, 679-700, 1986.

[34] Z. M. Parvez Sazzad, Y. Kawayoke, and Y. Horita, "Image Quality Evaluation Database". Available: http://mict.eng.u-toyama.ac.jp/mictdb.html

[35] P. Le Callet and F. Autrusseau, “Subjective quality assessment”, IRCCyN/IVC database. Available: http://www.irccyn.ec-nantes.fr/ivcdb/

[36] K. Seshadrinathan, R. Soundararajan, A. C. Bovik and L. K. Cormack, "Study of Subjective and Objective Quality Assessment of Video", *IEEE Transactions on Image Processing*, vol.19, no.6, pp.1427-1441, June 2010.

[37] K. Seshadrinathan, R. Soundararajan, A. C. Bovik and L. K. Cormack, "A Subjective Study to Evaluate Video Quality Assessment Algorithms", *SPIE Proceedings Human Vision and Electronic Imaging*, January 2010.

VITA

Koray Ozansoy was born in 1987 in İzmir, Turkey. He graduated from İzmir Karşıyaka Anatolian High School in 2005. He received the B.Sc. degree from Dokuz Eylül University in Electrical & Electronics Engineering (2006-2011). He worked one and a half year as part-time in application and component department at Vestel Dijital and has been working at Vestel Elektronik as a hardware development engineer in R&D department for the last two and a half years.