Automatic Measurement of Digital Video Quality

Wei Dai

A Thesis

in

The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Applied Science at Concordia University
Montreal, Quebec, Canada

April 2004

© Wei Dai, 2004

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisisitons et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-612-91016-4
Our file Notre référence
ISBN: 0-612-91016-4

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou aturement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this dissertation.

While these forms may be included in the document page count, their removal does not represent any loss of content from the dissertation.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de ce manuscrit.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

ABSTRACT

Automatic Measurement of Digital Video Quality

Wei Dai

Compression of digital video introduces artifacts (i.e., typical types of degradations), such as Blocking, Blurring, Ringing. Nowadays, digital video systems have been applied widely to various areas, and this has led to a rising demand for video quality metrics. In the past ten years, many automatic metrics of video quality have been developed; among them, the Wolf-Pinson metric is the better one. Based on the Wolf-Pinson metric, an Artifact-Measured-based Automatic Metric (AMAM) is proposed in this thesis, which measures four artifacts (Blocking, Ringing, Blurring and Block Flashing). Among them, Block Flashing is an artifact not previously measured, in the proposed metric, the Masking property of Human Visual System (HVS) is taken advantage to measure the Blocking and Ringing in order to improve performance.

In this thesis, the performance of the automatic metrics is evaluated by using the Pearson linear correlation coefficient, Mean Square Error and Spearman rank-order correlation coefficient. The simulation results indicate that the performance of the AMAM is significantly better than those of other simulated metrics (including the Wolf-Pinson metric) for MPEG-2 test video sequences.

Keywords: digital video system, automatic metric, Human Visual System (HVS), Wolf-Pinson metric, artifacts, Masking phenomenon, Block Flashing.

ACKNOWLEDGEMENTS

First I would like to express my deepest appreciation to my supervisor Dr. William E. Lynch, for his invaluable help, patience, guidance and support. I would also like to thank all the professors with whom I have interacted during my studies at Concordia University.

I would also like to thank my friends for their help, companion and fellowship. In particular, thanks for Mr. Chenyu Pai.

Finally, I express my deepest gratitude to my parents and my wife without their love, support and encouragement I would never have reached this level.

Thanks also for financial supports from the Natural Sciences and Engineering Research Council of Canada.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

PSNR	Peak Signal-to-Noise Ratio
MSE	Mean Squared Error
Mb	Mega byte
Gb	Gega byte
MPEG	Moving Pictures Expert Group
MPEG-1	Video Compression Standard
MPEG-2	Video Compression Standard
MPEG-4	Video Compression Standard
VQEG	Video Quality Expert Group
DVQ	Digital Video Quality
HVS	Human Visual System
DSCQS	Double Stimulus Continuous Quality Scale
DSIS	Double Stimulus Impairment Scale
SSCQE	Single Stimulus Continuous Quality Evaluation
IDM	Impairment Detection Method
2CABM	Two Context-Artifact-Based Metric
3CABM	Three Context-Artifact-Based Metric
EWPM	Extensive Wolf-Pinson Metric
Q-factor	Quantization factor
QPC	Quality Primitive Combination
S-T region	Spatial-Temporal region
ρ_1	Pearson linear correlation coefficient

ρ_2	Spearman rank-order correlation coefficient
AMAM	Artifact-Measured-based Automatic Metric
2D DCT	Two Dimension Discrete Cosine Transform
DC	Direct Current
AC	Alternating Current
Eqn.	Equation

Chapter 1 Introduction

1.1 Video compression

Digital video systems have many advantages over analog systems. Specifically, digital video can be easily stored, error-protected, multiplexed and mixed. However, to transmit, store and process a raw digital video signal requires a huge amount of resources. For example, a one-second video sequence at 30 frames/sec and with resolution of 704×480 pixels, requires as much as 120Mb/s bit rate (i.e. 120 million bits every second). If a two-hour video sequence with the above format is stored, it would need about 108 Gb (i.e. 1.08×10^5 Bytes) space. This large bandwidth and massive storage requirement is often not practical. In order to address this problem, video compression techniques have been developed. They can be classified into lossless compression and lossy compression. Lossless compression can gives perfect video quality. Compared with lossless compression, lossy compression obtains higher compression efficiency by sacrificing video quality. The efficiency of video compression may be measured by a compression ratio. The video compression ratio is defined as follows [1]:

$$Compression\ ratio = \frac{Compressed\ video\ size}{Orignal\ video\ size} \qquad (1.1)$$

Currently, major video compression standards, MPEG-1, MPEG-2 and MPEG-4 have been produced by the Moving Pictures Expert Group (MPEG) [2]. Because the above compression standards use lossy compression techniques, some typical types of degradations (such as Blurring, Blocking, and Ringing) are introduced into compressed video produced with these standards. These typical kinds of degradations are called artifacts. This thesis focuses on measuring typical compression artifacts and relating them to quality measurements given by the Human Visual System (HVS). Note that degradation caused by noise is not studied here.

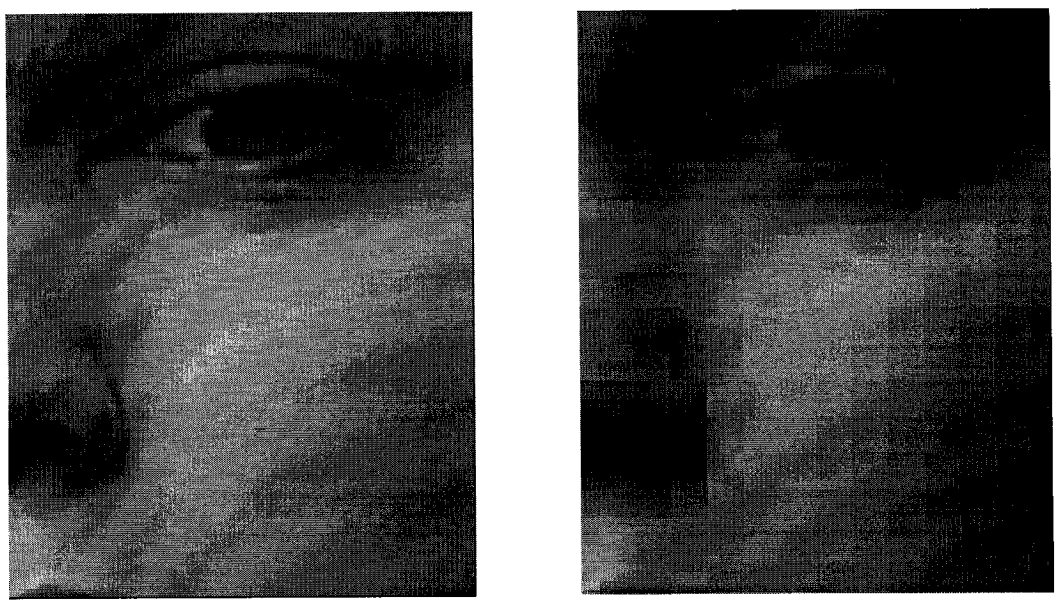
1.2 Artifacts of compression video

Standard lossy video compression techniques such as the MPEG-1, MPEG-2 and MPEG-4 result in the following several artifacts [3] [4]:

- Blocking: Degradation of the image quality, characterized by the visibility of the underlying block structure [3].
- Blurring: Distortion over the entire image region, characterized by the loss of sharpness of spatial details [3].
- Ringing (i.e., Spatial Edge Noise): Busyness near the edges of the objects,

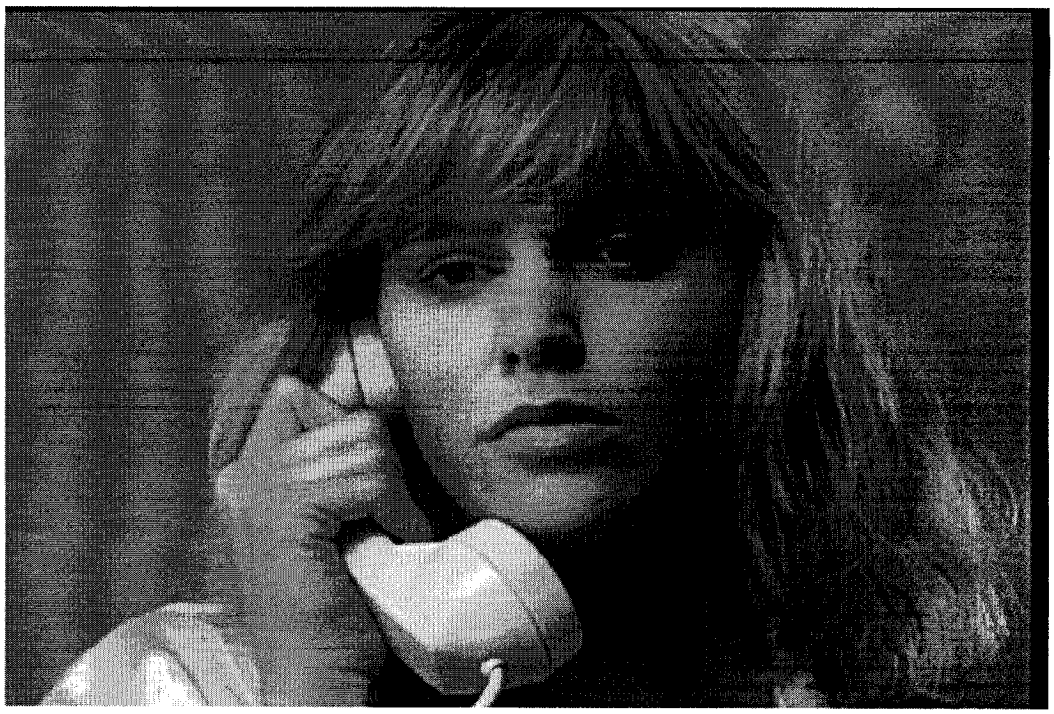
which caused by spatially varying distortion close to the edges [3].

Figure 1.1, Figure 1.2 and Figure 1.3 show examples of Blocking, Blurring and Ringing artifacts, respectively.



(a) (b)

Figure 1.1 Part of Susie's face
(a) The original image (b) The Blocking artifact in compressed image



(a)

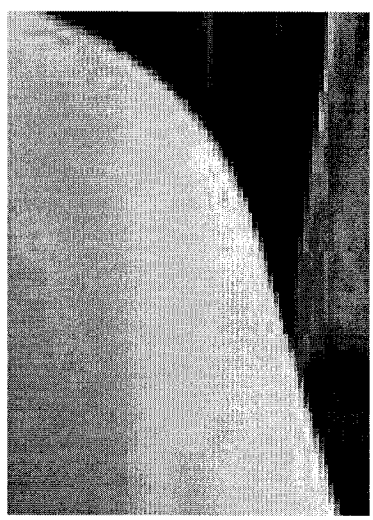
(b)

Figure 1.2 Susie image

(a) The original Susie image (b) The Blurred Susie image

(a)

(b)



(c)

Figure 1.3 Lena image [37]
(a) The whole original Lena image
(b) Original image of Lena's shoulder
(c) The Ringing artifact in the compressed image of Lena shoulder

In addition to the above artifacts, an artifact Block Flashing is observed and defined as follows:

- Block Flashing: Some blocks flash when a compressed video sequence is played.

Similar to Blocking artifact, Block Flashing is also caused by block-based compression technique (such as MPEG-1, MPEG-2 and MPEG-4). However, the Block Flashing differs from the Blocking in that it is a temporal and spatial artifact. The Block Flashing artifact is described in detail in Section 5.2.

1.3 Video Quality Measurement

Nowadays, because digital video systems have been applied widely, there is a rising demand for video quality metrics. Methods to evaluate video quality, can be classified into two classes, automatic and non-automatic. Non-automatic metrics are implemented by having human viewers. If done properly they are accurate but because of the requirement for human intervention they are slow and expensive. In contrast, automatic metrics do not require human intervention, but rather use computational machinery to analyze video signal and predict the video quality.

Currently, the most commonly used automatic metrics are Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR) since they are easily calculated. Performance of MSE and PSNR was tested in [5][6]. Test results suggest that measures of MSE and PSNR do not correlate well with video quality. The MSE and PSNR are defined as follows [3]:

$$MSE = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left[i(m,n) - j(m,n)\right]^2 \quad (1.2)$$

$$PSNR = 10 \times \log\left(\frac{255^2}{MSE}\right) \quad (1.3)$$

where, M and N are numbers of rows and columns of an image respectively; *i(m,n)* and *j(m,n)* are original image pixels and degraded (compressed) image pixels located at (m,n), respectively.

Weaknesses of MSE and PSNR are not surprising since MSE and PSNR are simple mathematical calculations and neglect the much more complex behaviors of the HVS. They suffer from the following weaknesses [7]:

1. Equal values of MSE and PSNR for two degraded video sequences do not mean the equivalent video quality.
2. MSE and PSNR do not give the types of artifacts and their respective locations.
3. An increase in the MSE does not always reflect a decrease in video quality.

To overcome the weaknesses of MSE and PSNR, many automatic metrics [8]-[34] incorporated with the properties of the HVS are developed. This class metrics can further be classified to metrics based on emulating the HVS [19]-[25], and metrics based on measuring expected artifacts and relating them to human quality ratings [26]-[29].

1.4 Problem Statement

In the automatic metrics that incorporate the properties of the HVS, the Wolf-Pinson metric [28] [29] has the better performance [12]. In this metric, after measures of artifacts are calculated, they are formed into an overall measure. However, the Wolf-Pinson metric does not consider the impacts of the artifact's context on measures of the artifacts. Specifically, in the HVS, there is Masking property, which is the reduction in the visibility of one image component because of the presence of another [32]. Thus, the Blocking may be masked in the blocks with more spatial details (i.e., non-flat blocks), while the Blocking in the blocks with less spatial details (i.e., flat blocks) is easily perceivable by the HVS. The Ringing may be masked in texture blocks, while it may be easily observed near sharp edges in non-flat blocks as well as near weak edges of flat blocks [34].

In this thesis, based on the Wolf-Pinson metric, a new automatic metric, Artifacts-Measured-based Automatic Metric (AMAM) is proposed, which takes advantage of the Masking property to detect the Blocking and Blurring. In order to improve performance of automatic metric further, except that the above innovation, an artifact Block Flashing is defined and detected. In the AMAM, after measures of the Blocking, Blurring, Ringing and Block Flashing are calculated, they are formed into an overall measure of the test video sequences.

1.5 Figure of merit

Qualitatively, the automatic quality metrics developed in this thesis, should predict,

given a degraded video and the original from whence it came, the quality that a group of human users would say the degraded video had. As a benchmark several degraded sequences that have been rated by human users are used. How close the developed metric is to these human observations must be judged.

The performance of automatic quality metrics may be evaluated with respect to two aspects: prediction accuracy and prediction monotonicity in relation to the human observations [14]. The prediction accuracy is the ability of the metric to predict the non-automatic quality score (i.e., quality score obtained by human visual rating) in a minimum average error. The Pearson linear correlation coefficient (denoted as ρ_1) and Mean Square Error (MSE) are two common methods used to measure the average error.

The relationship between automatic quality scores (i.e., quality scores calculated by using an automatic quality metric) and non-automatic quality scores (i.e., quality scores produced by human rating) should ideally be monotonic[14]. This attribute is called prediction monotonicity. The Spearman rank-order correlation coefficient (denoted as $\rho 2$) is a sensitive measure of monotonicity [35].

In this thesis, performance of the simulated metrics (including the proposed metrics) is evaluated with the Pearson linear correlation coefficient, Mean Square Error and Spearman rank-order correlation coefficient.

1.6 Organization of the thesis

This thesis is organized as follows:

Chapter 2 covers the background of the video quality measurement, which includes artifact classification, non-automatic quality measurement technique and automatic quality metrics. Among them, the Wolf-Pinson metric [28] is introduced in detail.

Chapter 3 briefly describes Impairment Detection Method (IDM) by Bishtawi.[37]

Chapter 4 illustrates five automatic quality metrics, and demonstrates their simulation results. Among them, the performance of Three Context-Artifact-Based Metric (3CABM) is analyzed in detail.

Chapter 5 proposes an Artifact-Measure-based Automatic Metric (AMAM) that may measure four artifacts, Blurring, Blocking, Ringing and Block Flashing. The performance of this metric is also evaluated in this chapter.

Chapters 4 and Chapter 5 represent my work.

Chapter 6 gives conclusions and future work direction.

Chapter 2 Video Quality Measurement Techniques

2.1 Introduction

In this chapter, the background knowledge of video quality measurement is introduced. This chapter is organized as follows. Video Compression Standards are introduced in Section 2.2. Then, Section 2.3 presents video quality measurement. Section 2.4 states non-automatic video quality measure. Afterwards, Section 2.5 describes Watson's DVQ metric [22] briefly. Finally, the Wolf-Pinson metric [28] is illustrated in Section 2.6.

2.2 Video Compression Standards

In Chapter 1, the video compression standards MPEG-1, MPEG-2 and MPEG-4 are described briefly. In this section, their applications are illustrated further. MPEG-1 is intended for video compressed coding at the bit rate of 1.2 MB/s. Video CD utilizes MPEG-1 standard. MPEG-2 extends the compression technique of MPEG-1, and includes all of the MPEG-1's functions. Originally, MPEG-2 video is primarily targeted at compressed coding of interlaced video at the bit rate range of 4 to 9 Mb/s. However, the applications of MPEG-2 have been extended to cover video of various bit rates. Digital Television, HDTV (i.e., high definition TV), and DVD apply the MPEG-2 standard. MPEG-4 is intended for the support of very lower bandwidth

applications, such as video telephone. Its bandwidth requirement is in the tens to hundreds of Kb/s range.

In this thesis, the compressed video sequences (i.e., test video sequences) are limited to MPEG-2 video sequences.

2.3 Video Quality Measurement Techniques

As mentioned in Chapter 1, video quality metrics may be categorized into two classes, non-automatic and automatic. In the upcoming sections, these two classes of metrics are illustrated in detail.

Non-automatic metrics require human viewers to evaluate the quality or difference of two video clips [23]. In the cases considered here, one of these two video clips is the source and the other is processed in some manner such as MPEG-2 or MPEG-4 compression coding. Non-automatic metrics if done properly, usually achieve accurate results for any given evaluation. However, these methods are costly and time-consuming, and they cannot be used to evaluate video quality in real time.

In contrast, automatic metrics do not require human viewers, but rather automatically analyze the video signal and predict the video quality. Automatic metrics may be classified into full-reference, partial-reference (reduced-reference) and no-reference methods. In this thesis, all of discussed metrics are the full-reference method. The full-reference methods require both the reference (original) and test (compressed) video sequences, while the partial-reference methods need only partial information of the reference sequence and the full test sequence; no-reference

methods only require the test video sequence.

Currently, the most commonly used automatic metrics are MSE and PSNR, but they may not reflect the actual perceptual quality correctly. The MSE and PSNR belong to full-reference.

In the last two decades, a great deal of effort [8]-[34] has been made to develop automatic metrics of video quality, which incorporate properties of the Human Visual System (HVS). The Video Quality Expert Group (VQEG) [13] formed in 1997 is a group of experts working in this field. The intention of VQEG is to evaluate and standardize new automatic metrics of video quality. In the Phase I test of VQEG, there were ten metrics tested; the results showed that no metric was entirely successful [12]. After the Phase I test, VQEG keeps working on the Phase II test [14]-[16] for full-reference quality assessment, reduced-reference assessment and no-reference quality assessment. The target of the Phase II test is to achieve more applicable results than those of the Phase I [14].

Most current video quality automatic metrics incorporate properties of the HVS to estimate the video quality either by attempting to emulating the HVS [21]-[24], or by measuring expected artifacts [25]-[29]. In this thesis, these artifact measures are called primitives.

Although automatic metrics based on emulating the HVS are the most general and potentially the most accurate ones, the HVS is extremely complex and many properties of the HVS are not well understood [24]. In this class of metrics, the Digital Video Quality (DVQ) metric [21] [22] by Watson is one of the proponents in

Phase I test of the VQEG[12].

The metrics based on measuring artifacts do not rely on emulating the HVS, while they measure expected artifacts according to a priori knowledge about the degradation present. This class of metrics while not as versatile, may perform well in some given applications. So far, one of the better automatic metric based on measuring artifacts is the one by Wolf and Pinson [28] [29]. In this metric, after the artifacts are measured, they are formed into an overall measure using a linear function. In this thesis, artifacts caused by compression are considered.

2.4 Non-automatic Quality Assessment

Non-automatic quality ratings have been standardized in ITU-R (formerly CCIR) Recommendation 500 [31]. In this standard, the viewing conditions, criteria for viewers and test scene selection, evaluation procedures, and analysis methods are described. There are three methods of non-automatic quality ratings mentioned in the standard, namely, Double Stimulus Continuous Quality Scale (DSCQS), Double Stimulus Impairment Scale (DSIS) and Single Stimulus Continuous Quality Evaluation (SSCQE). Among the above quality scales, the DSCQS and DSIS are full reference quality metric, while the SSCQE belongs to no reference quality metric. Of these three methods, the DSCQS is the most reliable method, thus VQEG uses the DSCQS for non-automatic ratings [12] [14] [15]. The outline of DSCQS is described as follows.

2.4.1 Double Stimulus Continuous Quality Scale

In the DSCQS method, viewers are showed two video sequences, one of which is a source sequence and the other is a test sequence. The presentation of the video sequences is illustrated in Figure 2.1.

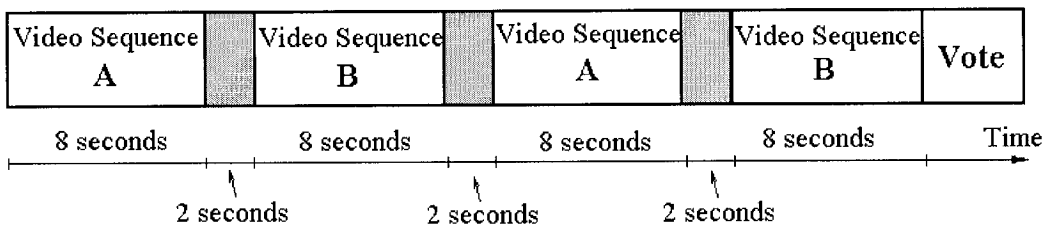


Figure 2.1 The presentation outline of the video sequence for the DSCQS [23]

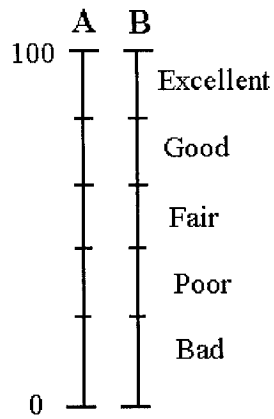


Figure 2.2 Rating scale for the DSCQS method [23]

The source video sequence is unimpaired, while the test video sequence may be impaired (or unimpaired). The source and test video sequences are presented twice in an alternating fashion in each trial. In different trials, the sequences' presentations are random, as well as viewers are not told which is the source and which is the test

sequence. Viewers evaluate the quality of both video sequences according to a continuous quality scale as shown in Figure 2.2. Afterwards, based on evaluation, an equivalent numerical score from 0 to 100 is given [23].

2.5 Watson DVQ Metric

The Watson DVQ metric is a full-reference method. The output of the DVQ is a number that indicates the impairment degree of the test video sequence related to the original video sequence. The DVQ value of the original video sequence is 0. A higher DVQ value means worse quality. Figure 2.3 shows the block diagram of the Watson DVQ metric [22].

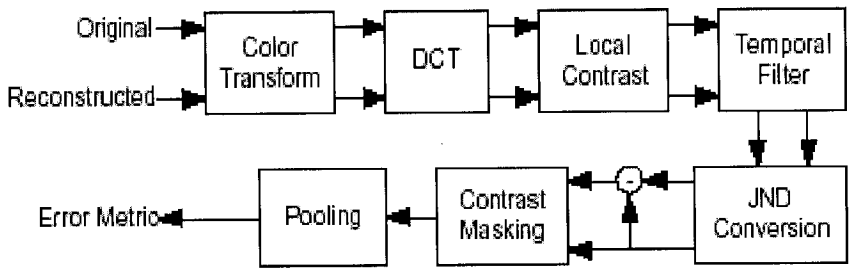


Figure 2.3 The block diagram of the Watson's DVQ metric [22]

The purpose of each component in Figure 2.3 is listed as follows.

- Color transformation: the video signal (i.e., YC_BC_R color signal) is first converted to an opponent color space since the HVS is an opponent-color system [23][24].

- DCT: The DVQ transforms the input signal to frequency domain using a Two Dimension Discrete Cosine Transform (2D DCT) since the HVS treats different frequencies differently,
- Local Contrast: The HVS may only see patterns if luminance difference is large in bright area. A low-pass time-filter is utilized to model this effect.
- Temporal Filter: The temporal filter is a low-pass filter since the HVS reacts like a low-pass filter in high frequency area of temporal direction.
- Just-Noticeable-Difference (JND) Conversion: Frequency components are divided by using the corresponding spatial thresholds described in [21] becasue the HVS has deferent sensitivities at different spatial frequencies.
- Contrast Masking: The reduction in the visibility of one image component because of the presence of another, especially ones that appear at the same location and share the same spatial frequency.
- Pooling: To use a Minkowski metric to combine all JNDs to form a measure of visual error.

The Watson DVQ metric is one of the metrics that emulate the HVS. In general, most of these metrics are complex and time consuming in particular in analyzing the spatial spectrum of video [22]. Watson accelerates this analysis by using the 2D DCT. This may provide a powerful advantage, since the 2D DCT may have already been calculated in the video compression process. On the other hand, because the HVS is an extremely complicated system, and current understanding of the HVS is limited, several critical issues about the HVS are still under investigations [18]. Therefore, for

the automatic metrics that emulate the HVS, such as the Watson DVQ metric, does not always perform better than other metrics.

2.6 Wolf-Pinson Metric

The Wolf-Pinson metric [28] is a full-reference method based on measures of the three artifacts, i.e. Blurring, Blocking and Ringing. The block diagram of this metric is shown in Figure 2.4.

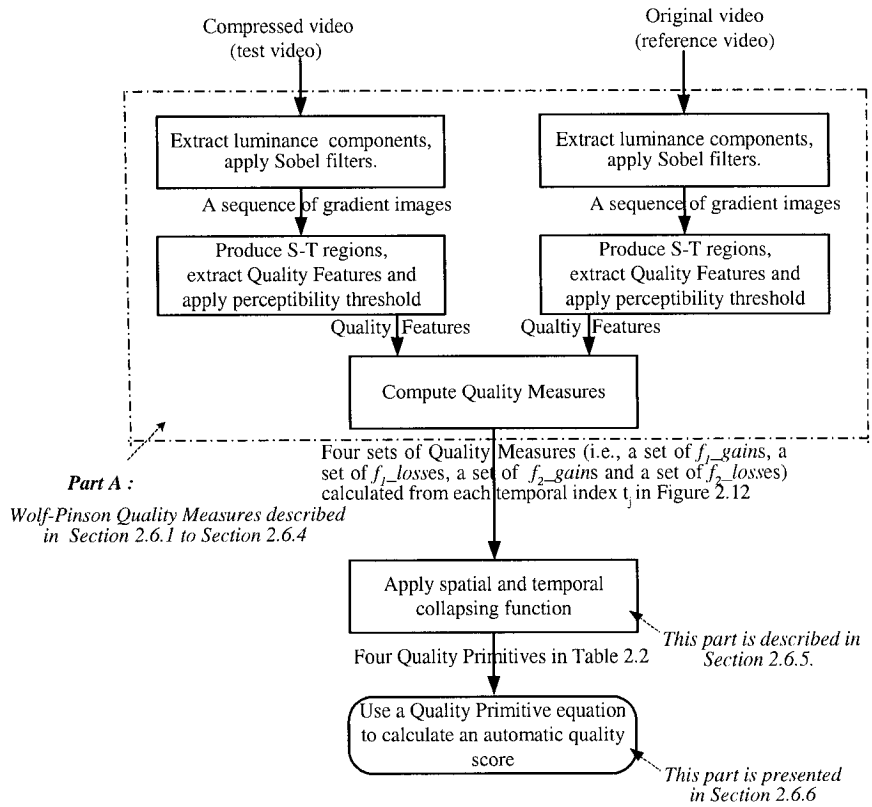


Figure 2.4 The block diagram of the Wolf-Pinson Metric [28]

In this metric, the luminance components of the original and compressed video sequences are processed separately with the 3×3 Sobel filters, and then the Quality

Features, f_1 and f_2 are extracted from the Spatial-Temporal regions (S-T regions) of these processed video sequences. Afterwards, a perceptibility threshold is applied to these extracted Quality Features. Then, by comparing the Quality Features extracted from the compressed video sequence with those extracted from the original video sequence, four types of the Quality Measures, namely, a set of f_1*_gain*s, a set of f_1*_loss*es, a set of f_2*_gain*s and a set of f_2*_loss*es are computed. Here, the Quality Features and Quality Measures are discussed in detail in Section 2.6.3 and Section 2.6.4, respectively. After that, spatial and temporal collapsing functions are applied to the above four collections of Quality Measures to form the four Quality Primitives, *Qp_*f_1*_gain*, *Qp_*f_1*_loss*, *Qp_*f_2*_gain* and *Qp_*f_2*_loss*. Finally, with a linear equation given by Wolf and Pinson [28], an automatic quality score is calculated.

The major steps of the Wolf-Pinson metric are described in the subsequent sections. Here, all the materials are referenced [28][29].

2.6.1 Application of Sobel filters

In the Wolf-Pinson metric [28], the luminance components of original and compressed video sequences are extracted because only these components are required in the measurement. Actually, the luminance components extracted are the luminance images of video frames.

After the luminance images are extracted, these images are processed with two 3 $\times$3 Sobel filters, respectively. The Sobel filters are common edge detection FIR filters that include a horizontal Sobel filter and a vertical Sobel filter as shown in

Table 2.1.

-1	-2	-1
0	0	0
1	2	1

(a)

-1	0	1
-2	0	2
-1	0	1

(b)

Table 2.1 The 3×3 Sobel filter [36]

(a) The Horizontal Sobel filter. (b) The Vertical Sobel filter.

In a luminance image, for a given pixel located at row i, column j, its luminance value is denoted by p(i, j). The output of the horizontal Sobel filter for the p(i, j), denoted by H (i, j), which is expressed mathematically as

$$H(i,j) = P(i-1,j+1) - P(i-1,j-1) + 2\times P(i,j+1) - 2\times P(i,j-1) + P(i+1,j+1) - P(i+1,j-1) \quad (2.1)$$

where, the P(i-1, j+1), P(i-1, j-1), P(i, j+1), P(i, j-1), P(i+1, j+1) and P(i+1, j-1) are the luminance values of six neighboring pixels, respectively.

Similarly, the output of the vertical Sobel filter for the p(i, j), denoted by V(i, j).

The V (i, j) is expressed mathematically as

$$V(i,j) = P(i+1,j) - P(i-1,j) + 2\times P(i+1,j) - 2\times P(i-1,j) + P(i+1,j) - P(i-1,j+1) \quad (2.2)$$

The output of the Sobel filters for P(i, j), denoted by R(i, j), is the luminance gradient magnitude. The direction of the gradient is denoted by θ(i, j). The luminance image processed by the Sobel filters is called a gradient image, which is composed of the gradient magnitude R(i, j)s. The R(i, j) and θ (i, j) are expressed respectively as follows [28]:

$$R(i,j) = \sqrt{H(i,j)^2 + V(i,j)^2} \quad (2.3)$$

$$\theta(i,j) = \tan^{-1}\left[V(i,j)/H(i,j)\right] \quad (2.4)$$

As an illustration, Figure 2.5 and Figure 2.6 show the luminance images of the first frame in an original video clip " Susie " and a compressed video clip "Susie.24.mpeg2", respectively. "Susie.24.mpeg2" is produced by encoding "Susie" with an MPEG2 encoder (the quantization factor is 48). The resolution of these two video clips is 720×480. Each video clip consists of 30 frames per second, 10 seconds in duration.

Figures 2.7 and 2.8 show the gradient images achieved from the original "Susie" image (i.e., Figure 2.5) and the compressed "Susie" image (i.e., Figure 2.6) by using the 3×3 Sobel filters. In these two gradient images, the edges in the luminance images (i.e., Figure 2.5 and Figure 2.6) are extracted and enhanced.

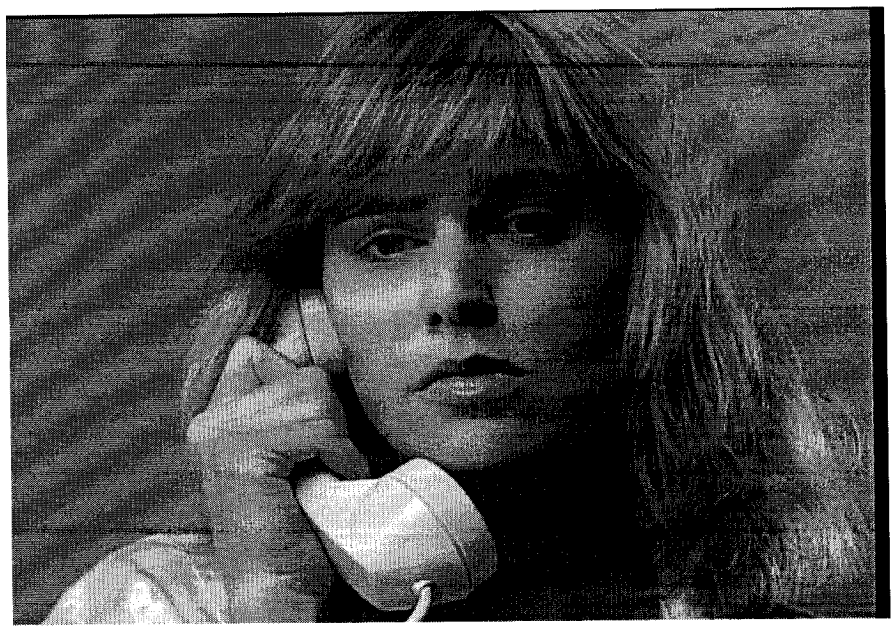


Figure 2.5 The luminance image of the first frame in the original video clip "Susie"

Figure 2.6 The luminance image of the first frame in the compressed video clip"Susie.24.mpeg2"

Figure 2.7 The gradient image produced from the original "Susie" image (i.e., Figure 2.5) with the 3×3 Sobel filters.

The Figure 2.7 results from the luminance image (i.e., Figure 2.5) whose edges are extracted and enhanced.

Figure 2.8 The gradient image produced from the compressed "Susie" image (i.e., Figure 2.6) with the 3×3 Sobel filters

The Figure 2.8 results from the luminance image (i.e., Figure 2.6) whose edges are extracted and enhanced.

2.6.2 S-T regions

After the luminance images are processed with the Sobel filters, a sequence of gradient images are achieved. Then, these gradient images need to be divided into S-T regions with the size of 8 horizontal pixels × 8 vertical pixels × the temporal width (△t), as shown in Figure 2.9.

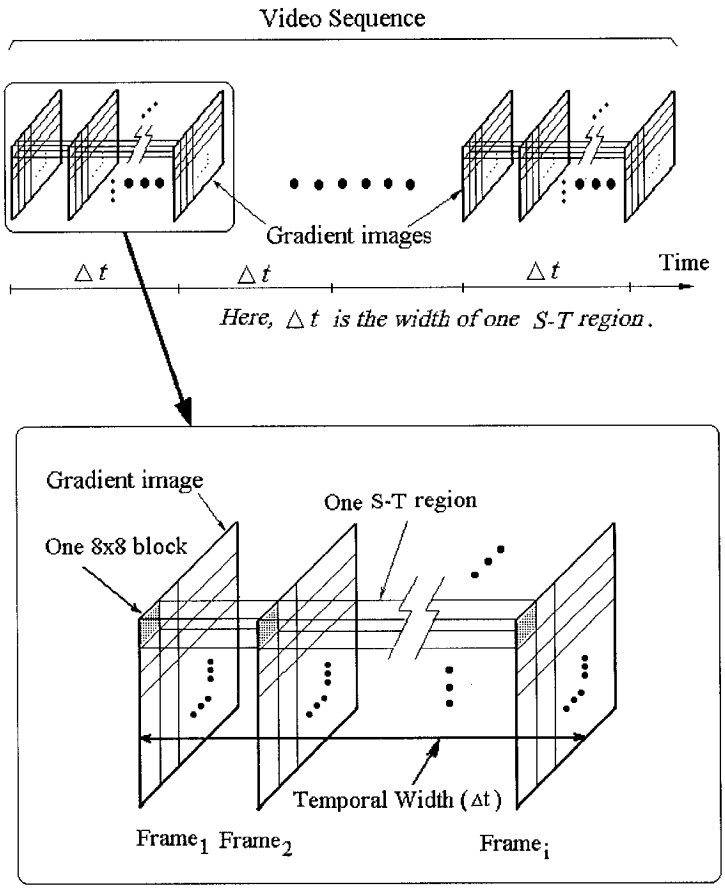


Figure 2.9 Illustration of the S-T regions

For MPEG-2 video sequences, Wolf and Pinson found that an S-T region of 8 horizontal pixels × 8 vertical pixels ×1 video frame (i.e. each 8×8 block in the Figure 2.9) works well for spatial distortion measures [28].

Note that, for five simulated metrics described in Chapter 4, the size of an S-T region is 8 horizontal pixels × 8 vertical pixels × 1 video frame. In this case, each 8×8 block of video frames is an S-T region.

2.6.3 Extraction of Quality Features

After the S-T regions are produced, the Quality Features f_1 and f_2 are extracted from each of these regions. The Quality Features are defined as a quantity of information calculated from an S-T region of a video sequence. [28]

The first Quality Feature f_1 is calculated as the standard deviation of the all R(i,j,t)s within a given S-T region, and then adjusted by a perceptibility threshold p (i.e., if the value of the standard deviation is less than p, the f_1 is set equal to p.).[28] The p value is given to 12 by [29]. The f_1 is represented mathematically as follows:

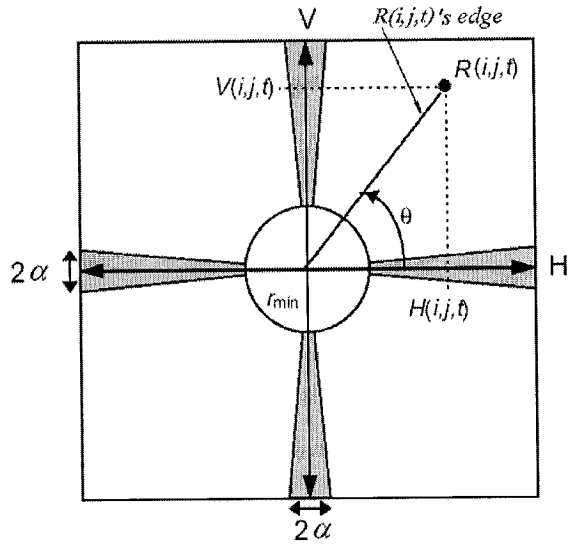
$$f_1 = \{ stdev\,[R(i,j,t)] \}\Big|_p \qquad (2.5)$$

where, the R(i,j,t)s are in an S-T region.

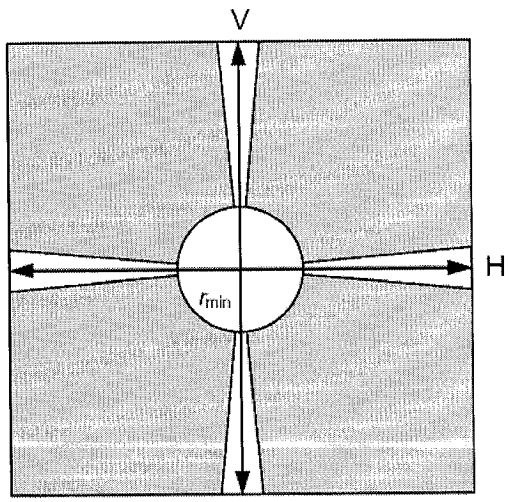
The Quality Feature f_1 reflects variations in the overall amount of spatial activity within a corresponding S-T region. For example, the f_1 value extracted from the compressed video sequence is less than that from the original video sequence because the Blurring distortion reduces the amount of spatial activity in the compressed video

sequence, while, the f_1 value extracted from the compressed video is getting greater than that from the original because the Ringing increases the amount of spatial activity in the compressed video sequence.[28]

The Quality feature f_2 is sensitive to the angular variations of spatial activity within a corresponding S-T region.



(a)



(b)

Figure 2.10 Gradient distribution pictures [28]

(a) Horizontal and Vertical gradient distribution
(b) Non Horizontal and Vertical gradient distribution

In Figure 2.10 (a), the grey region contains most of the R(i,j,t)s that have horizontal or vertical edges, while the R(i,j,t)s that have diagonal edges are zeroed [28]. Here, a small angular, α is given by Wolf and Pinson[28]. When the angular between a R(i,j,t)'s edge and the axes is less than α, this R(i,j,t)'s edge is also viewed as horizontal (or vertical) approximately.

In Figure 2.10 (b), the grey region contains most of the R(i,j,t)s that have diagonal edges, while the R(i,j,t)s that have horizontal or vertical edges are zeroed [28]. In Figure 2.10 (a), and (b), the gradient magnitudes R(i,j,t)s less than r_{min} are also zeroed in order to ensure accurate angular computations. In [28], the recommended values of r_{min} and α are given by Wolf and Pinson, which are 20 and 0.05236 radians, respectively.

Here, the horizontal and vertical gradient distribution (i.e., Figure 2.10 (a)) is denoted as *HV*, while the non Horizontal and Vertical gradient distribution (i.e., Figure 2.10 (b)) is denoted as *NHV*. The *HV* and *NHV* may be represented mathematically as follows [28]:

$$HV = \left\{ \begin{array}{l|ll} R(i,j,t) & R(i,j,t) \geq r_{\min} \ \ and \ \ \frac{n\pi}{2} - \alpha < \theta(i,j,t) < \frac{(n+1)}{2} + \alpha & (n = 0,1,2,3) \\ 0 & otherwise & \end{array} \right\} \quad (2.6)$$

$$NHV = \left\{ \begin{array}{l} R(i,j,t) \\ 0 \end{array} \middle| \begin{array}{ll} R(i,j,t) \geq r_{\min} \ \ and \ \ \frac{n\pi}{2} + \alpha < \theta(i,j,t) < \frac{(n+1)}{2} - \alpha \ \ (n = 0,1,2,3) \\ otherwise \end{array} \right\} \quad (2.7)$$

where, the R(i, j, t)s are in an S-T region, and the θ(i, j, t)s are the R(i, j, t)s direction computed with Eqn. 2.4.

The Quality Feature f_2 is defined as the ratio of the mean of R(i, j, t)s in the HV to the mean of R(i, j, t)s in the NHV, where these two means are adjusted by a thresholds q (i.e., if the mean is less than q, the numerator value or denominator value is set equal to q.).[28] The q value is given to 3 by [29]. The f_2 is represented mathematically as follows:

$$f_2 = \left\{ mean\left[R(i,j,t) \middle| R(i,j,t) \in HV \right] \right\} \Big|_q \Big/ \left\{ mean\left[R(i,j,t) \middle| R(i,j,t) \in NHV \right] \right\} \Big|_q \quad (2.8)$$

where, the R(i, j, t)s are in an S-T region.

As mentioned above, the Quality Feature f_2 is sensitive to the angular variations of spatial activity of the S-T regions. For instance, if erroneous horizontal or vertical edges are introduced to a compressed video sequence, such as Blocking distortion, then the f_2 of the compressed video sequence is greater than the f_2 of the original video sequence [28].

After the Quality Features are extracted from each of the S-T regions of the original and compressed video sequences respectively, the Quality Measures may be computed.

2.6.4 Calculation of Quality Measures

The Quality Measures are defined as the measures of video quality distortions, which are calculated by comparing the Quality Features extracted from the original video sequence with those extracted from the compressed video sequence. In the calculation, four Quality Measures of each S-T region are computed, which are denoted as f_1_loss, f_1_gain, f_2_loss and f_2_gain.

In the Wolf-Pinson metric, gain and loss of spatial activity in the compressed video sequence need to be examined separately since they produce fundamentally different effects on quality perception.[28] Of the four Quality Measures, f_1_gain and f_2_gain are measures of the gain of spatial activity, while f_1_loss and f_2_loss are measures of the loss. These four Quality Measures are defined below [28]:

$$f_1\_gain = pp\left\{\log\left[\frac{f_{1\_compressed}}{f_{1\_original}}\right]\right\} \tag{2.9}$$

$$f_1\_loss = np\left\{\frac{f_{1\_compressed} - f_{1\_origianl}}{f_{1\_original}}\right\} \tag{2.10}$$

$$f_2\_gain = pp\left\{\log\left[\frac{f_{2\_compressed}}{f_{2\_original}}\right]\right\} \tag{2.11}$$

$$f_2\_loss = np\left\{\frac{f_{2\_compressed} - f_{2\_origianl}}{f_{2\_original}}\right\} \tag{2.12}$$

where, the $f_{1\_compressed}$ and $f_{2\_compressed}$ are the Quality Features extracted from the S-T regions in the compressed video sequence, while the $f_{1\_original}$ and $f_{2\_original}$ are the Quality Features extracted from the corresponding S-T regions in the original video sequence.

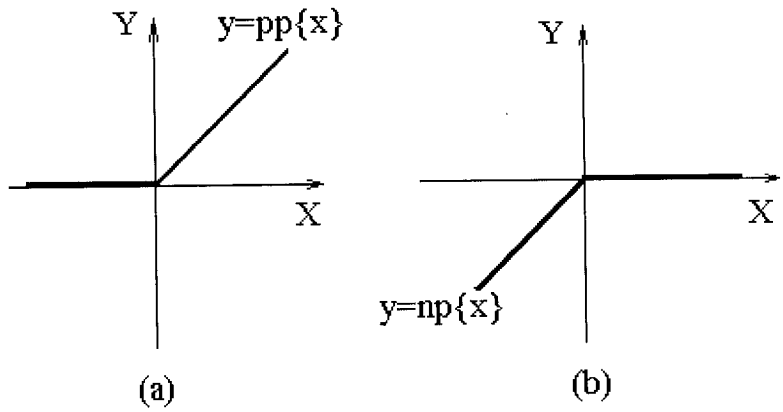


Figure 2.11 Function curves of y=pp{x} and y=np{x}
(a) Function curve of y=pp{x} (b) Function curve of y=np{x}

Figure 2.11 shows function curves of y=pp{x} and y=np{x}. Here, the *pp* is the positive part operator. That is, the f_{1_gain} and f_{2_gain} are replaced with zero, if the value inside the bracket is negative. The *np* is the negative part operator. That is, the f_{1_loss} and f_{2_loss} are replaced with zero, if the value inside the brackets is positive. Therefore, the f_{1_gain} and f_{2_gain} are not negative, while the f_{1_loss} and f_{2_loss} are not positive.

Because each S-T region produces four Quality Measures, four sets of the Quality Measures, (i.e. a set of f_{1_gain}s, a set of f_{1_loss}es, a set of f_{2_gain}s and a set of f_{2_loss}es) are computed from the temporal index t_j (j =1, 2,..., i) shown in Figure 2.12. In the next step, based on these four sets of Quality Measures, four Quality Primitives (i.e. $Qp\_f_{1\_gain}$, $Qp\_f_{1\_loss}$, $Qp\_f_{2\_gain}$ and $Qp\_f_{2\_loss}$) of the whole

compressed video sequence are computed. Here, the Quality Primitives are defined as the artifact measures of whole compressed sequence, which are computed from Quality Measures by using the spatial and temporal collapsing functions.

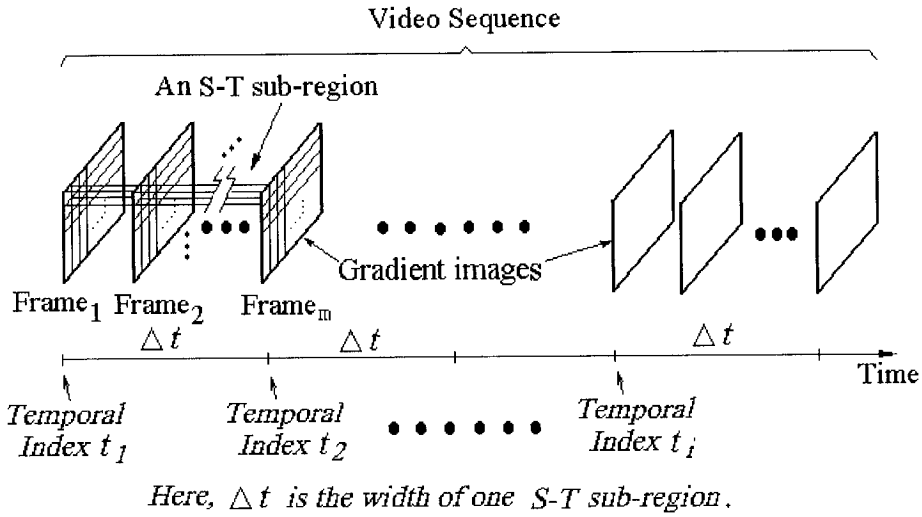


Figure 2.12 Illustration of the temporal indexes

2.6.5 Spatial and temporal collapsing functions

In the Wolf-Pinson metric (see Figure 2.4), the spatial collapsing function is computed as the average of the worst 5% of the measured distortions over all S-T regions in the temporal index t_j (j =1, 2,..., i) as shown in Figure 2.12 [28]. That means,

- The average of the maximal 5% of all f_1_gains in the temporal index t_j is calculated, which is denoted by $f_1\_gain\_t_j$.
- The average of the maximal 5% of all f_2_gains in the temporal index t_j is calculated, which is denoted by $f_2\_gain\_t_j$.

- The average of the minimal 5% of all $f_1_loss$es in the temporal index t_j is calculated, which is denoted by $f_1\_loss\_t_j$.
- The average of the minimal 5% of all $f_2_loss$es in the temporal index t_j is calculated, which is denoted by $f_2\_loss\_t_j$.

Thus, the four averages (i.e., $f_1\_gain\_t_j$, $f_2\_gain\_t_j$, $f_1\_loss\_t_j$ and $f_2\_loss\_t_j$) are obtained from each temporal index t_j.

The temporal collapsing function is computed as the average of the measured distortions (i.e., artifacts) over all temporal index t_js (j =1, 2,..., i) [28]. That means

- The average of all $f_1_gain_t_j$s (j =1, 2,..., i) is calculated, which is denoted by Quality Primitive $Qp\_f_1\_gain$.
- The average of all $f_1_loss_t_j$s (j =1, 2,..., i) is calculated, which is denoted by Quality Primitive $Qp\_f_1\_loss$.
- The average of all $f_2_gain_t_j$s (j =1, 2,..., i) is calculated, which is denoted by Quality Primitive $Qp\_f_2\_gain$.
- The average of all $f_2_loss_t_j$s (j =1, 2,..., i) is calculated, which is denoted by Quality Primitive $Qp\_f_2\_loss$.

Because the four Quality Primitives $Qp\_f_1\_gain$, $Qp\_f_1\_loss$, $Qp\_f_2\_gain$ and $Qp\_f_2\_loss$ are computed from all temporal index t_js, they reflect degree of artifacts in the whole compressed video sequence (i.e., test video sequence). Expected relationship between artifacts and four Quality Primitives is listed in Table 2.2

Quality Primitives	Value Range	Expected relationship Between Artifacts And Quality Primitives
$Qp\_f_1\_gain$	≥ 0	$Qp\_f_1\_gain$ reflects Ringing degree of the compressed video. $Qp\_f_1\_gain$ increases, if Ringing increases.
$Qp\_f_1\_loss$	[0, -1]	$Qp\_f_1\_loss$ reflects Blurring degree of the compressed video. $Qp\_f_1\_loss$ decreases, if Blurring increases.
$Qp\_f_2\_gain$	≥ 0	$Qp\_f_2\_gain$ reflects Blocking degree of the compressed video. $Qp\_f_2\_gain$ increases, if Blocking increases.
$Qp\_f_2\_loss$	[0, -1]	$Qp\_f_2\_loss$ does not have the monotonous increase or decrease relationships with the three artifacts.

Table 2.2 Expected relationships between artifacts and four Quality Primitives

In Table 2.2, the expected relationship comes from the following analysis.

- As mentioned in Section 2.6.3, if the Ringing increases the amount of spatial activity in compressed video sequence, then the f_1 extracted from the compressed video sequence is getting bigger than the f_1 extracted from the original video sequence. Thus, the f_1_gain computed from Eqn.2.9 reflects the Ringing degree of the compressed video sequence. Because $Qp\_f_1\_gain$ is calculated from the f_1_gain, $Qp\_f_1\_gain$ reflects the Ringing degree of the compressed video sequence.
- As mentioned in Section 2.6.3, if the Blurring reduces the amount of spatial activity in compressed video sequence, then the f_1 extracted from the compressed video sequence is getting smaller than the f_1 extracted from the original video sequence. Thus, the f_1_loss computed from Eqn.2.10 reflects the Burring degree of the compressed video sequence. Because $Qp\_f_1\_loss$ is

calculated from the f_1_loss, $Qp\_f_1\_loss$ reflects the Blurring degree of the compressed video sequence.

- As mentioned in Section 2.6.3, if erroneous horizontal or vertical edges are introduced to a compressed video sequence, such as the Blocking artifact, then the f_2 of the compressed video sequence is getting bigger than the f_2 of the original video sequence [28]. Thus, the f_2_gain computed from Eqn.2.11 reflects the Blocking degree of the compressed video sequence. Because $Qp\_f_2\_gain$ is calculated from the f_2_gain, $Qp\_f_2\_gain$ reflects the Blocking degree of the compressed video sequence.

After the four Quality Primitives (i.e. $Qp\_f_1\_gain$, $Qp\_f_1\_loss$, $Qp\_f_2\_gain$ and $Qp\_f_2\_loss$) of a compressed video sequence are calculated, an automatic quality score (i.e., a quality score computed by using an automatic quality metric) is produced by using a Quality Primitive equation given by Wolf and Pinson [28].

2.6.6 Calculation of an automatic quality score

An automatic quality score is a predicted value of a non-automatic quality score (i.e., a quality score produced by Human rating) for a compressed video sequence. In the Wolf-Pinson metric, an automatic quality score is computed by Eqn. 2.13 [28].

$$\textit{Automatic quality score} = 0.38 \times Qp\_f_1\_loss + 0.39 \times Qp\_f_2\_loss - 0.23 \times Qp\_f_2\_gain \quad (2.13)$$

In Eqn. 2.13, there is no $Qp\_f_1\_gain$ since it did not produce consistent correlation

results across all test data sets used by Wolf and Pinson [28]. The automatic quality score calculated by Eqn. 2.13 ranges from –1 to 0. Smaller automatic quality scores value indicate worse quality, and a value of zero means no distortions in the compressed video sequence.

2.6.7 Discussion of the Wolf-Pinson Metric

As one of the proponents, the Wolf-Pinson metric was taken in Phase I test of the VQEG. The test result showed this metric is successful in quality measurement for most of test video sequences.[12] However, this metric can not well measure quality degradations in several of video sequences of more spatial details.

In order to improve the predicted performance of the automatic metrics, two new metrics (i.e., Two Context-Artifact-based Metric (2BCAM) and Three Context-Artifact-based Metric (3BCAM)) are designed and simulated in Chapter 4, which take advantage of the Masking property of the HVS to measure the Blocking and Ringing.

Chapter 3 Impairment Detection Method of Image Quality

3.1 Introduction

The impairment Detection Method (IDM) [37] of image quality by Bishtawi is a gray image quality metric based on the properties of the HVS, which may detect and measure three common artifacts, Blurring, Blocking, and Ringing. This metric is a full-reference method that requires original (reference) and compressed (test) images. In this chapter, the IDM is introduced briefly.

The rest of this chapter is organized as follows. The block classification algorithm is introduced in Section 3.2. Then, Section 3.3 describes Bisthawi's Quality Primitives. Afterwards, discussion is presented in Section 3.4.

3.2 Block Classification Algorithm

In the IDM, in order to take advantage of Masking property of the HVS to measure the Blocking and Ringing, 8×8 blocks of images are classified into

1. flat blocks, which have less spatial details (i.e. low spatial frequency),
2. sharp edge blocks, which have sharp edges,
3. texture blocks, which do not belong to either flat blocks or sharp edge blocks.

Here, the sharp edge blocks and texture blocks belong to non-flat blocks, which have more spatial details (i.e. higher spatial frequency) than those of the flat blocks.

The above block classification method is produced from the following analysis. As mentioned in Chapter 1, there is Masking property in the HVS, which is the reduction in the visibility of one image component because of the presence of another [32]. Thus, artifacts like Blocking may be masked in the blocks with more spatial details (i.e., sharp edge blocks and texture blocks), while in the blocks with less spatial details (i.e., flat blocks) they may be easily perceivable [32]. Ringing which may be masked in texture blocks, while it may be easily observed near sharp edges in sharp edge blocks as well as near weak edges in flat blocks [34].

In the IDM, in order to classify the 8×8 blocks of an image into three classes, a block classification algorithm is utilized, which includes three major steps as follows:

1. Extraction of gradient image,
2. Detection of flat blocks,
3. Detection of sharp edge blocks and texture blocks.

In the following two sections, these three steps are introduced. Here, all the materials are referenced Bishtawi's thesis [37].

3.2.1 Extraction of gradient image

For the purpose of extracting and enhancing spatial activity information (i.e., spatial edge information), in the block classification algorithm, the original image needs to be processed with 3×3 Sobel filters first, which is described in Section 2.6.1.

3.2.2 Detection of flat blocks

As mentioned in Section 2.6.1, a gradient image is an image processed with 3×3 Sobel filters, which is composed of gradient magnitude R(i, j)s calculated with Eqn. 2.3. In the detection of flat blocks, after a gradient image is extracted from an original image, a threshold T_i is used to select the flat blocks by using the following method.

- If the gradient magnitude R(i, j) of each pixel in a block is less than or equal to T_i (i.e., flat block threshold), then this block is labeled as a flat block; otherwise, this block is labeled as a non-flat block.

In [37], the recommended value of T_i is 20.

After all flat blocks of the original image are identified, the remaining blocks are called non-flat blocks. These will later be further classified into sharp edge blocks and texture blocks (see Section 3.2.3). Figure 3.1, Figure 3.2 and Figure 3.3 show the original "Susie" image, its gradient image and the flat blocks in the original "Susie" image, respectively.

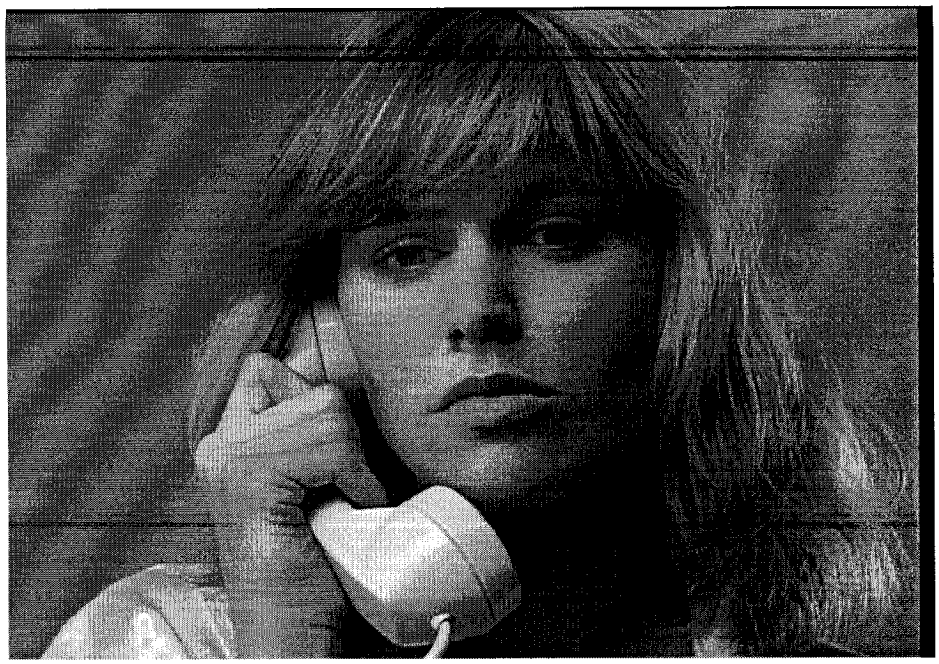


Figure 3.1 The original "Susie" image

Figure 3.2 The gradient image of the original "Susie" image

Figure 3.3 The flat blocks (unblackened blocks) of the original "Susie" image

3.2.3 Detection of sharp edge blocks and texture blocks

Prior to the sharp edge block detection, the sharp edges need to be extracted from the non-flat blocks. The first step is to label individual pixels in the non-flat block as either sharp edge pixels or not. If the gradient magnitude R(i, j)s of pixels in the non-flat blocks are greater than T_s (the sharp edge threshold), these pixels are marked as sharp edges; otherwise, they are labeled as non sharp edges (i.e., background).

Next, sharp edge pixels that are adjacent are labeled as belonging to the same sharp edge region. Likewise non sharp edge pixels that are adjacent are labeled as belonging to the same non sharp edge region.

Two conditions must be met in order for the block to be labeled as a sharp edge block:

1) It must contain exactly one sharp edge region and either one or two

non sharp edge regions

2) The contour surrounding the sharp edge region must have a length less than T_c.

Figure 3.4 (a) and (b) give examples of sharp edge blocks

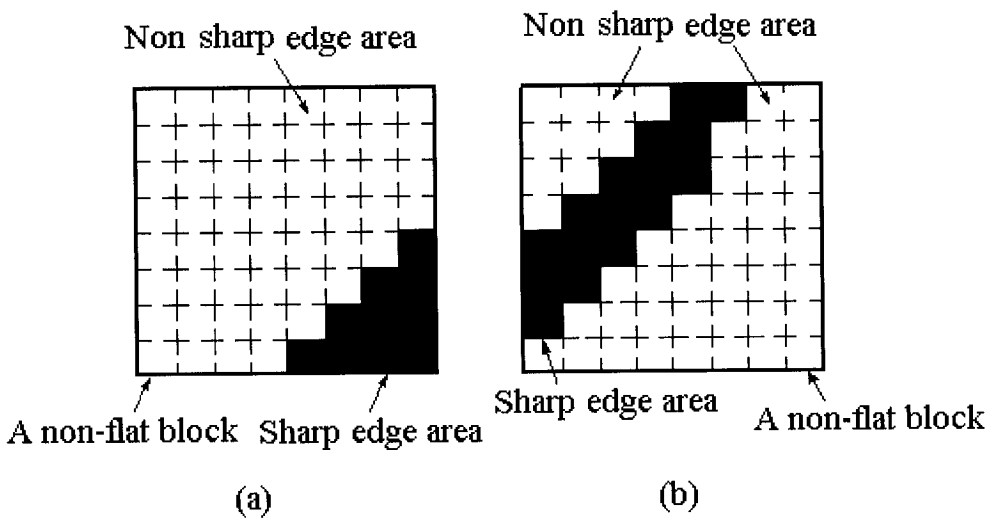


Figure 3.4 Illustration of sharp edge blocks

(a) A sharp edge block that has one edge area and one non-edge area
(b) A sharp edge block that has one edge area and two non-edge areas

In [37], the recommended values of T_s and T_c are 200 and 30, respectively.

After that, non-flat blocks that do not belong to sharp edge blocks are labeled as texture blocks.

Figure 3.5 and Figure 3.6 show the sharp edge blocks and texture blocks in the original "Susie" image, respectively.

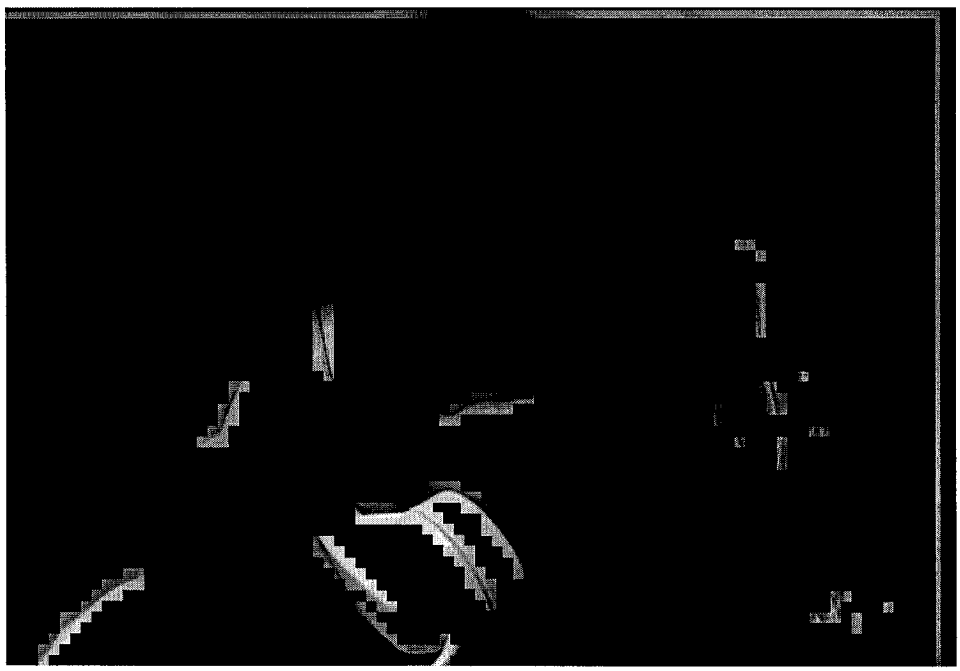


Figure 3.5 The sharp edge blocks (unblackened blocks) of the original "Susie" image

Figure 3.6 The texture blocks (unblackened blocks) of the original "Susie" image

3.3 Bisthawi's Quality Primitives

In the IDM, after the blocks of an original image are classified into three classes, the Quality Features are extracted from the original and compressed images. Afterwards, by comparing these Quality Features, fourteen Quality Primitives are calculated. Table 3.1 lists the fourteen Quality Primitives and their measured locations.

Measured Locations			Quality Primitives
Flat blocks	Block frame	Horizontal and vertical edges	$Qp\_fbf\_P_{hv}$
		Non-horizontal and vertical edges	$Qp\_fbf\_P_{nhv}$
	Block interior	Horizontal and vertical edges	$Qp\_fbi\_P_{hv}$
		Non-horizontal and vertical edges	$Qp\_fbi\_P_{nhv}$
Texture blocks	Block Frame	Horizontal and vertical edges	$Qp\_tbf\_P_{hv}$
		Non-horizontal and vertical edges	$Qp\_tbf\_P_{nhv}$
	Block Interior	Horizontal and vertical edges	$Qp\_tbi\_P_{hv}$
		Non-horizontal and vertical edges	$Qp\_tbi\_P_{nhv}$
Sharp edge blocks	Block frame	Horizontal and vertical edges	$Qp\_sef\_P_{hv}$
		Non-horizontal and vertical edges	$Qp\_sef\_P_{nhv}$
	Sharp edges	Horizontal and vertical edges	$Qp\_sei\_P_{hv}$
		Non-horizontal and vertical edges	$Qp\_sei\_P_{nhv}$
	Background	Horizontal and vertical edges	$Qp\_seb\_P_{hv}$
		Non-horizontal and vertical edges	$Qp\_seb\_P_{nhv}$

Table 3.1 The fourteen Quality Primitives and their measured locations

In Table 3.1, the block frame and block interior are the regions as shown in Figure 3.7.

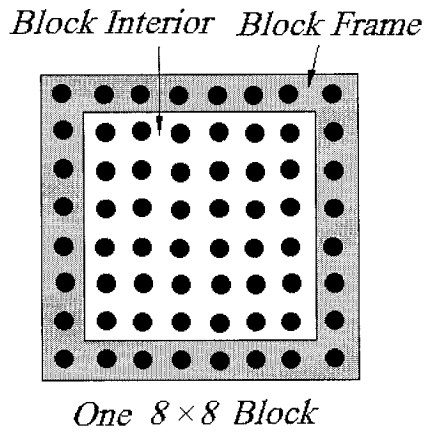


Figure 3.7 The block frame and block interior

Quality Primitives	Blurring Artifact Increasing	Blocking Artifact Increasing	Ringing Artifact Increasing
$Qp\_fbf\_P_{hv}$		↑	
$Qp\_fbi\_P_{nhv}$	↓		
$Qp\_tbf\_P_{hv}$		↑	
$Qp\_tbi\_P_{nhv}$	↓		
$Qp\_sef\_P_{hv}$		↑	
$Qp\_sei\_P_{nhv}$	↓		
$Qp\_seb\_P_{hv}$			↑
Note: The other seven Quality Primitives do not have the monotonous increase or decrease relationships with the three classes of artifacts.			

Table 3.2 Relationships between Quality Primitives and three artifacts

Table 3.2 presents the relationships between these Quality Primitives and artifacts [37]. In the table, ↑ and ↓ represent an increase and a decrease in the amount of a Quality Primitive, respectively.

The above Quality Primitives and their calculation methods [37] are not discussed in detail, since they are out of the scope of this thesis.

3.4 Discussion

In this chapter, the IDM by Bishtawi is described briefly. It is concluded that there are two major advantages of this metric.

1. The IDM takes advantage of the Masking property of the HVS to detect the Blocking and Ringing.
2. The seven of the fourteen Quality Primitives are better measures since they are monotonic and have one-to-one relationship with the artifacts.

However, Bishtawi never used quality estimates by human beings to create a quality metric. In the following chapter, by means of linear regression analysis, a Quality Primitive combination is presented. Furthermore, the performance of the IDM is evaluated.

Chapter 4 Context-Artifact-Based Metric

4.1 Introduction

In this chapter, five automatic metrics (i.e., Wolf-Pinson Metric, Bishtawi Metric, Extensive Wolf-Pinson Metric, Two Context-Artifact-Based Metric and Three Context-Artifact-Based Metric) are described, and then their performance is evaluated. Meanwhile, Two Context-Artifact-Based Metric (2CABM) and Three Context-Artifact-Based Metric (3CABM) are two new metrics. In these two metrics, in order to take advantage of the Masking property of the HVS to measure the Blocking and Ringing (as mentioned in Section 3.2), the block classification algorithm illustrated in Section 3.2 is used.

In the 2CABM and 3CABM, using the Wolf-Pinson method, four Quality Measures (i.e., f_1_gain, f_1_loss, f_2_gain and f_2_loss) are calculated from each of the 8×8 blocks of video frames, and then these Quality Measures are classified into two classes (i.e., flat blocks and non-flat blocks) or three classes (i.e., flat blocks, texture blocks and sharp edge blocks). Afterwards, by applying the spatial and temporal collapsing functions mentioned in Section 2.6.5, eight or twelve Quality Primitives are produced. Finally, an automatic quality score is computed by using a Quality Primitive combination obtained by linear regression analysis.

The rest of this chapter is organized as follows. Five simulated automatic metrics are described in Section 4.2. Then, Section 4.3 presents how to combine Quality Primitives to an automatic quality score. Afterwards, figure of merit are stated in Section 4.4. Finally, simulation results are illustrated in Section 4.5.

4.2 The simulated metrics

In this section, five metrics are described, which are the Wolf-Pinson Metric, Bishtawi Metric, Extensive Wolf-Pinson Metric (EWPM), 2CABM and 3CABM.

4.2.1 The Wolf-Pinson Metric

This metric is the Wolf-Pinson metric [28] described in Section 2.6. The block diagram of this metric is shown as Figure 4.1. In this metric, the luminance components (i.e. luminance images) of the original and compressed video sequences are extracted and filtered with the Sobel filters, and then the Quality Features are calculated from 8×8 blocks of video frames (i.e., S-T regions). Afterwards, a perceptibility threshold is applied to these Quality Features. Then, four sets of the Quality Measures (i.e., a set of f_1_gains, a set of $f_1_loss$es, a set of f_2_gains and a set of $f_2_loss$es) are computed from each video frame. As the spatial and temporal collapsing functions are applied to the above four collections of the Quality Measures, four Quality Primitives (i.e. $Qp\_f_1\_gain$, $Qp\_f_1\_loss$, $Qp\_f_2\_gain$ and $Qp\_f_2\_loss$) are produced. Finally, with Eqn. 2.13 given by Wolf and Pinson, an automatic quality score is computed.

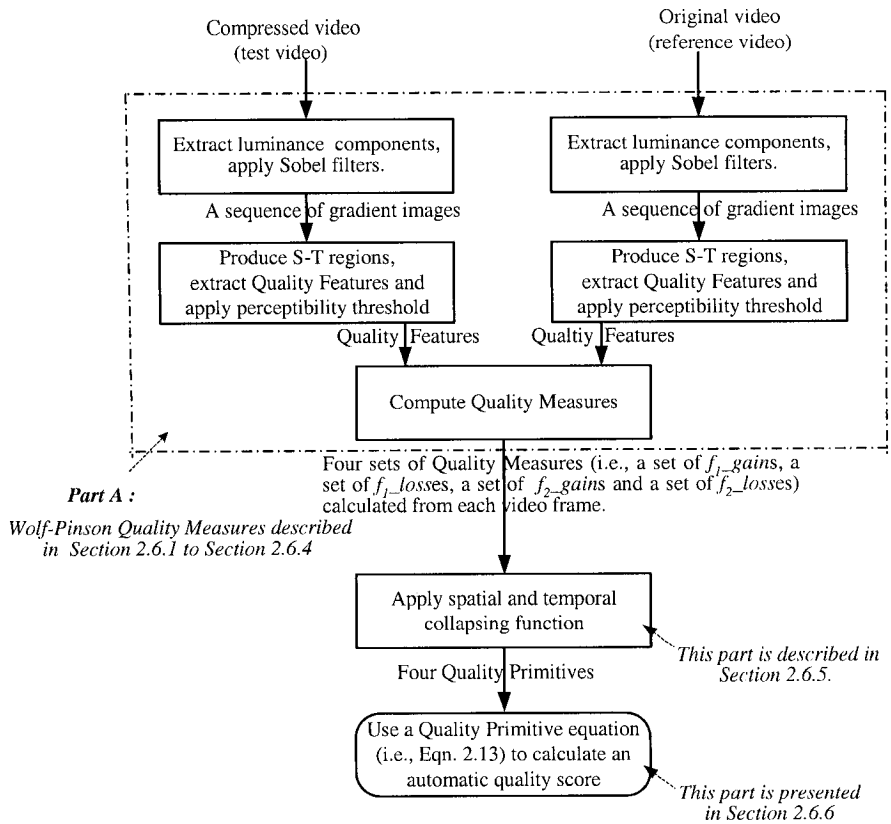


Figure 4.1 The block diagram of Wolf-Pinson Metric [28]

Note that, in our simulation, because all of the test video sequences are MPEG-2 video clips, size of the S-T region of this metric is the value recommended by Wolf and Pinson, i.e., 8 horizontal pixels × 8 vertical pixels × 1 video frame. In this case, each 8×8 block of video frames is an S-T region. This recommended value is also used for the following simulated metrics, EWPM, 2CABM and 3CABM.

4.2.2 The Bishtawi Metric

The core of this metric is the IDM of image quality [37] described in Chapter 3. In this metric, the luminance components (i.e. luminance images) are extracted from the original and compressed video sequences first, then with the IDM, fourteen Quality

Primitives of each video frame (i.e., each luminance image) are computed. Afterwards, the average of each of these fourteen Quality Primitives over all video frames is calculated, which is the Quality Primitive of the whole compressed video sequence. Finally, an automatic quality score is calculated by using a Quality Primitive combination obtained through linear regression analysis.

4.2.3 Extensive Wolf-Pinson Metric (EWPM)

The block diagram of this metric is shown in Figure 4.2. In the EWPM, most parts are the same as those of the Wolf-Pinson metric, while the only difference between them is that the EWPM utilizes a Quality Primitive combination instead of Eqn. 2.13 to calculate an automatic quality score. Here, this Quality Primitive combination is obtained by linear regression analysis.

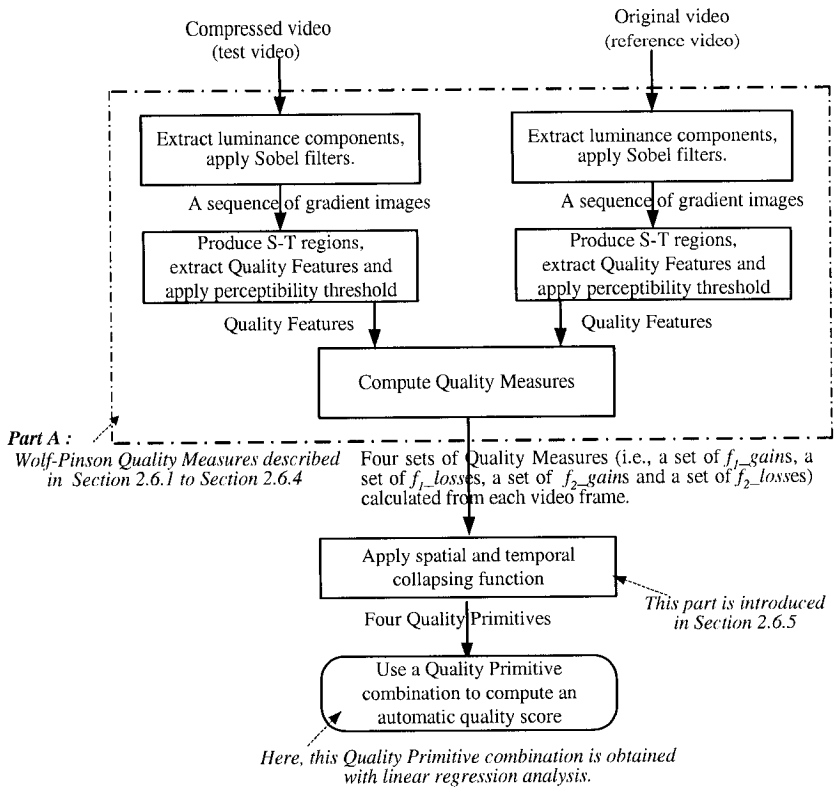


Figure 4.2 The block diagram of the EWPM

4.2.4 Two Context-Artifact-Based Metric (2CABM)

The block diagram of this metric is shown as Figure 4.3. Part A of the diagram is the Wolf-Pinson Quality Measures described in Section 2.6.1 to Section 2.6.4. Part B of the diagram is the block classification algorithm illustrated in Section 3.2.

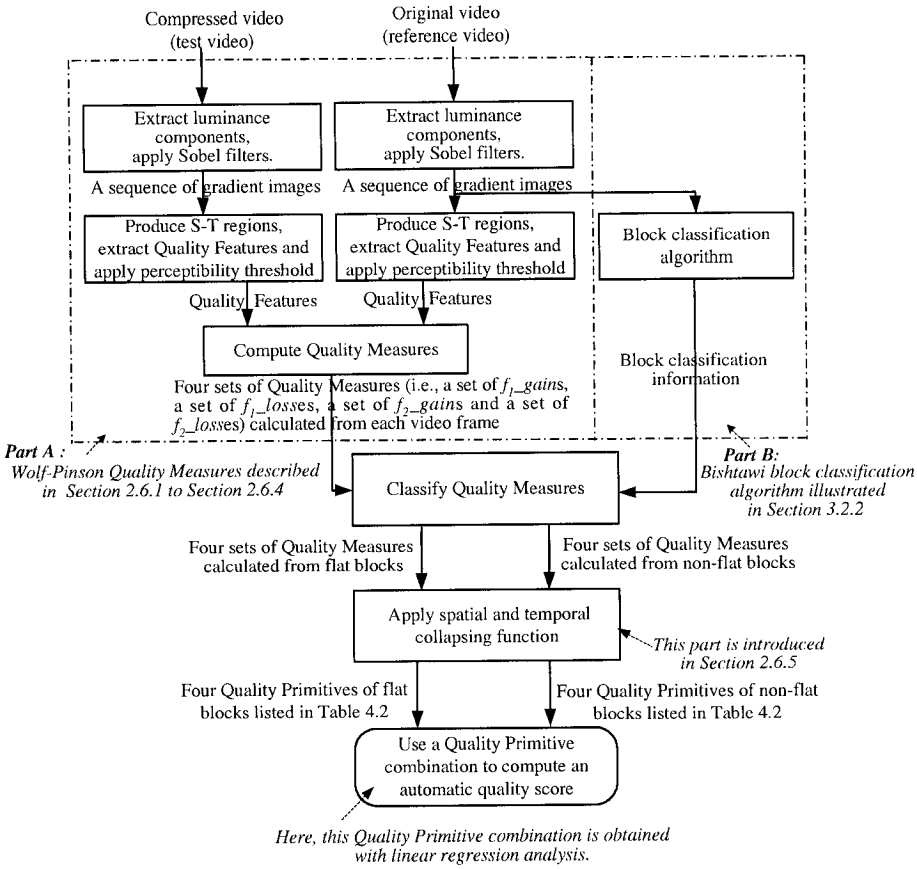


Figure 4.3 The block diagram of the 2CABM

In this metric, the luminance components (i.e. luminance images) are extracted from the original and compressed video sequences first. With the block classification algorithm introduced in Section 3.22 (see Part B in Figure 4.3), all 8×8 blocks of each frame (i.e., each luminance image) of the original (reference) video sequence are classified into flat blocks (blocks with less spatial details) and non-flat blocks (blocks with more spatial details) as shown in Figure 4.4.

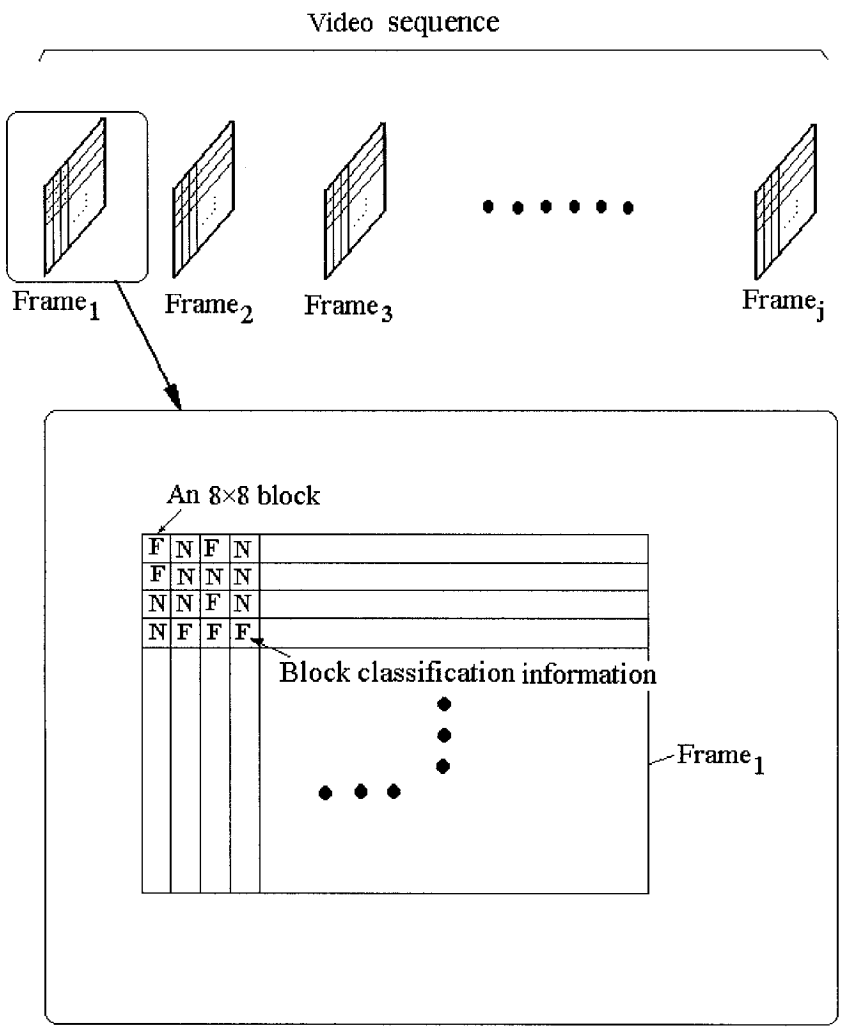


Figure 4.4 Illustration of the block classification

Figure 4.4 shows that all 8×8 blocks of each video frame are classified into flat blocks and non-flat blocks. In Figure 4.4, F represents a flat block, while N represents a non-flat block. Note that, the block classification information in Figure 4.4 is hypothetical.

After that, by using the Wolf-Pinson measure method (see Part A in Figure 4.3), four sets of Quality Measures (i.e. a set of *f_1_gains*, a set of *f_1_losses*, a set of *f_2_gains* and a set of *f_2_losses*) are calculated from each video frame. Then, according to the block classification information as shown in Figure 4.4, these four sets of Quality Measures are classified into eight sets listed in Table 4.1.

Order	Quality Measure Set	Description of Quality Measure
1	A set of *fb_f_1_gains*	The Quality Measure *f_1_gain* calculated from the flat blocks is denoted as *fb_f_1_gain*.
2	A set of *fb_f_1_losses*	The Quality Measure *f_1_loss* calculated from the flat blocks is denoted as *fb_f_1_loss*.
3	A set of *fb_f_2_gains*	The Quality Measure *f_2_gain* calculated from the flat blocks is denoted as *fb_f_2_gain*.
4	A set of *fb_f_2_losses*	The Quality Measure *f_2_loss* calculated from the flat blocks is denoted as *fb_f_2_loss*.
5	A set of *nfb_f_1_gains*	The Quality Measure *f_1_gain* calculated from the non-flat blocks is denoted as *nfb_f_1_gain*.
6	A set of *nfb_f_1_losses*	The Quality Measure *f_1_loss* calculated from the non-flat blocks is denoted as *nfb_f_1_loss*.
7	A set of *nfb_f_2_gains*	The Quality Measure *f_2_gain* calculated from the non-flat blocks is denoted as *nfb_f_2_gain*.
8	A set of *nfb_f_2_losses*	The Quality Measure *f_2_loss* calculated from the non-flat blocks is denoted as *nfb_f_2_loss*.

Table 4.1 Eight sets of Quality Measures

After the above eight sets of Quality Measures are achieved from each video frame, by applying the spatial and temporal collapsing functions introduced in Section 2.6.5, eight Quality Primitives are calculated, which represent the measured distortions of the whole compressed video sequence. Table 4.2 lists the expected relationship between artifacts and these Quality Primitives. Finally, an automatic quality score is calculated with a Quality Primitive combination achieved by the use of linear regression analysis.

Quality Primitive Class	Quality Primitive	Expected Relationship Between Artifacts And Quality Primitives
Quality Primitives of Flat Blocks	*Qp_fb_f1_gain*	*Qp_fb_f1_gain* reflects Ringing degree of flat blocks of the compressed video. *Qp_fb_f1_gain* increases, if Ringing increases.
	Qp_fb_f1_loss	*Qp_fb_f1_loss* reflects Blurring degree of flat blocks of the compressed video. *Qp_fb_f1_loss* decreases, if Blurring increases.
	Qp_fb_f2_gain	*Qp_fb_f2_gain* reflects Blocking degree of flat blocks of the compressed video. *Qp_fb_f2_gain* increases, if Blocking increases.
	Qp_fb_f2_loss	*Qp_fb_f2_loss* does not have the monotonous increase or decrease relationships with the three classes of artifacts.
Quality Primitives of Non-flat Blocks	*Qp_nfb_f1_gain*	*Qp_nfb_f1_gain* reflects Ringing degree of non-flat blocks of the compressed video. *Qp_nfb_f1_gain* increases, if Ringing increases.
	Qp_nfb_f1_loss	*Qp_nfb_f1_loss* reflects Blurring degree of non-flat blocks of the compressed video. *Qp_nfb_f1_loss* decreases, if Blurring increases.
	Qp_nfb_f2_gain	*Qp_nfb_f2_gain* reflects Blocking degree of non-flat blocks of the compressed video. *Qp_nfb_f2_gain* increases, if Blocking increases.
	Qp_nfb_f2_loss	*Qp_nfb_f2_loss* does not have the monotonous increase or decrease relationships with the three artifacts.

Table 4.2 Expected Relationship between artifacts and Quality Primitives

In Table 4.2, there are two classes of Quality Primitives, Quality Primitives of flat block and Quality Primitives of non-flat Block. Here, the expected relationship between artifacts and Quality Primitives is extrapolated from those listed in Table 2.2.

4.2.5 Three Context-Artifact-Based Metric (3CABM)

The block diagram of this metric is shown as Figure 4.5. Part A of the diagram is the Wolf-Pinson Quality Measures described in Section 2.6.1 to Section 2.6.4. Part B of the diagram is the block classification algorithm illustrated in Section 3.2.2 and Section 3.2.3.

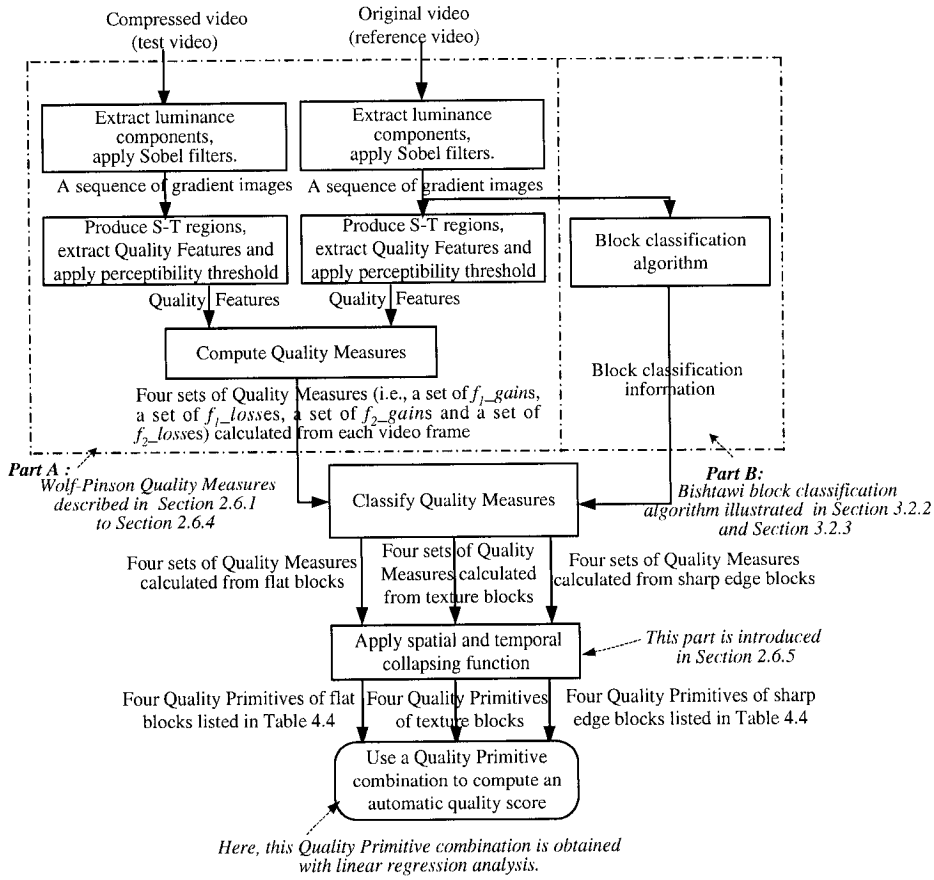


Figure 4.5 The block diagram of the 3CABM

In this metric, the luminance components (i.e. luminance images) are extracted from the original and compressed video sequences first. With the block classification algorithm (see Part B in Figure 4.5) introduced in Section 3.22 and Section 3.23, all 8 $\times$8 blocks of each frame (i.e., each luminance image) of the original (reference) video sequence are classified into flat blocks, texture blocks, and sharp edge blocks as shown in Figure 4.6.

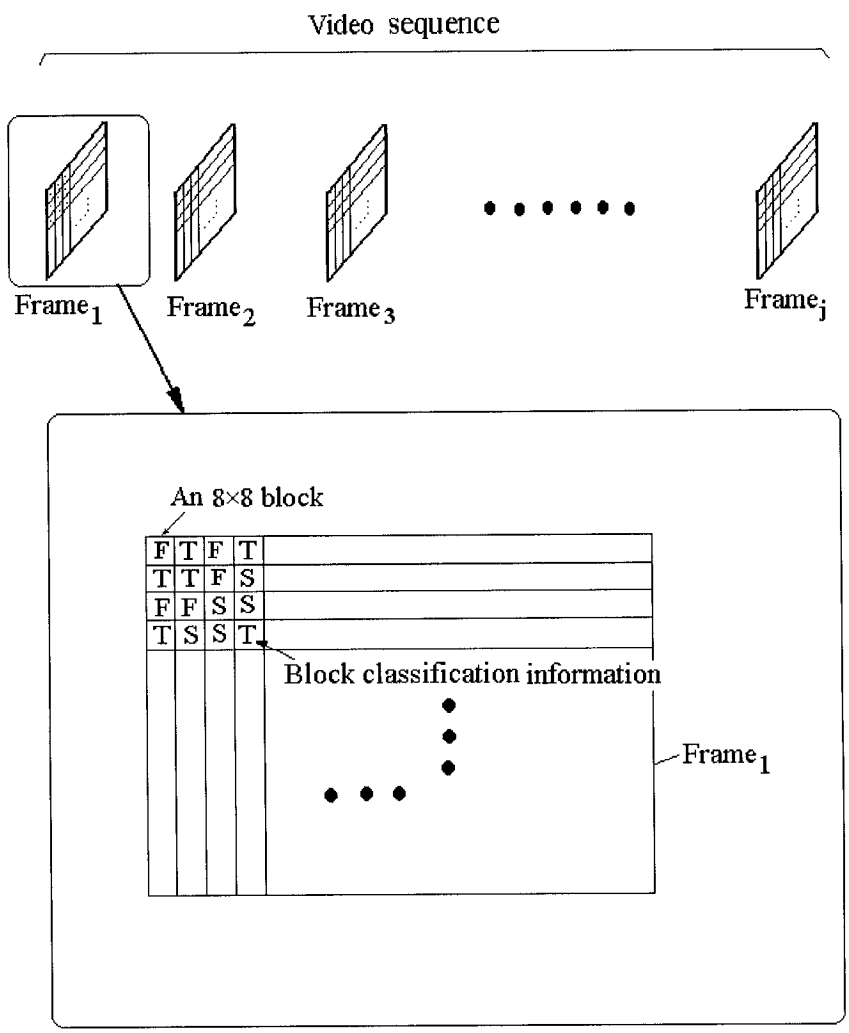


Figure 4.6 Illustration of the block classification

Figure 4.6 shows that all 8×8 blocks of each video frame are classified into flat blocks, texture blocks, and sharp edge blocks. In Figure 4.6, F represents a flat block, T represents a texture block, and S represents a sharp edge block. Note that, the block classification information in Figure 4.6 is hypothetical.

After that, by using the Wolf-Pinson measure method (see Part A in Figure 4.5), four sets of Quality Measures (i.e. a set of *f1_gains*, a set of *f1_losses*, a set of *f2_gains* and a set of *f2_losses*) are calculated from each video frame. Afterward, according to the block classification information as shown in Figure 4.6, these four sets of Quality Measures are classified into twelve sets listed in Table 4.3.

Order	Quality Measure Sets	Description of Quality Measure
1	A set of *fb_f1_gains*	The Quality Measure *f1_gain* calculated from the flat blocks is denoted as *fb_f1_gain*.
2	A set of *fb_f1_losses*	The Quality Measure *f1_loss* calculated from the flat blocks is denoted as *fb_f1_loss*.
3	A set of *fb_f2_gains*	The Quality Measure *f2_gain* calculated from the flat blocks is denoted as *fb_f2_gain*.
4	A set of *fb_f2_losses*	The Quality Measure *f2_loss* calculated from the flat blocks is denoted as *fb_f2_loss*.
5	A set of *tb_f1_gains*	The Quality Measure *f1_gain* calculated from the texture blocks is denoted as *tb_f1_gain*.
6	A set of *tb_f1_losses*	The Quality Measure *f1_loss* calculated from the texture blocks is denoted as *tb_f1_loss*.
7	A set of *tb_f2_gains*	The Quality Measure *f2_gain* calculated from the texture blocks is denoted as *tb_f2_gain*.
8	A set of *tb_f2_losses*	The Quality Measure *f2_loss* calculated from the texture blocks is denoted as *tb_f2_loss*.
9	A set of *seb_f1_gains*	The Quality Measure *f1_gain* calculated from the sharp edge blocks is denoted as *seb_f1_gain*.
10	A set of *seb_f1_losses*	The Quality Measure *f1_loss* calculated from the sharp edge blocks is denoted as *seb_f1_loss*.
11	A set of *seb_f2_gains*	The Quality Measure *f2_gain* calculated from the sharp edge blocks is denoted as *seb_f2_gain*.
12	A set of *seb_f2_losses*	The Quality Measure *f2_loss* calculated from the sharp edge blocks is denoted as *seb_f2_loss*.

Table 4.3 Twelve sets of Quality Measures

After the above twelve sets of Quality Measures are achieved from each video frame, by employing the spatial and temporal collapsing functions introduced in Section 2.6.5, twelve Quality Primitives are calculated, which represent the measured distortions of the whole compressed video sequence. The expected relationship between these Quality Primitives and artifacts are listed in Table 4.4.

Finally, an automatic quality score is calculated with a Quality Primitive combination acquired by means of linear regression.

Quality Primitive Class	Quality Primitive	Expected Relationship Between Artifacts And Quality Primitives
Quality Primitives of Flat Blocks	*Qp_fb_f1_gain*	*Qp_fb_f1_gain* reflects Ringing degree of flat blocks of the compressed video. *Qp_fb_f1_gain* increases, if Ringing increases.
	Qp_fb_f1_loss	*Qp_fb_f1_loss* reflects Blurring degree of flat blocks of the compressed video. *Qp_fb_f1_loss* decreases, if Blurring increases.
	Qp_fb_f2_gain	*Qp_fb_f2_gain* reflects Blocking degree of flat blocks of the compressed video. *Qp_fb_f2_gain* increases, if Blocking increases.
	Qp_fb_f2_loss	*Qp_fb_f2_loss* does not have the monotonous increase or decrease relationships with the three artifacts.
Quality Primitives of Texture Blocks	*Qp_tb_f1_gain*	*Qp_tb_f1_gain* reflects Ringing degree of texture blocks of the compressed video. *Qp_tb_f1_gain* increases, if Ringing increases.
	Qp_tb_f1_loss	*Qp_tb_f1_loss* reflects Blurring degree of texture blocks of the compressed video. *Qp_tb_f1_loss* decreases, if Blurring increases.
	Qp_tb_f2_gain	*Qp_tb_f2_gain* reflects Blocking degree of texture blocks of the compressed video. *Qp_tb_f2_gain* increases, if Blocking increases.
	Qp_tb_f2_loss	*Qp_tb_f2_loss* does not have the monotonous increase or decrease relationships with the three artifacts.
Quality Primitives of Sharp Edge Blocks	*Qp_seb_f1_gain*	*Qp_seb_f1_gain* reflects Ringing degree of sharp edge blocks of the compressed video. *Qp_seb_f1_gain* increases, if Ringing increases.
	Qp_seb_f1_loss	*Qp_seb_f1_loss* reflects Blurring degree of sharp edge blocks of the compressed video. *Qp_seb_f1_loss* decreases, if Blurring increases.
	Qp_seb_f2_gain	*Qp_seb_f2_gain* reflects Blocking degree of sharp edge blocks of the compressed video. *Qp_seb_f2_gain* increases, if Blocking increases.
	Qp_seb_f2_loss	*Qp_seb_f2_loss* does not have the monotonous increase or decrease relationships with the three artifacts.

Table 4.4 Expected relationship between artifacts and Quality Primitives

In Table 4.4, there are three classes of Quality Primitives: Quality Primitives of flat blocks, Quality Primitives of texture blocks and Quality Primitives of sharp edge blocks. Here, the expected relationship between artifacts and Quality Primitives is

extrapolated from those listed in Table 2.2.

4.3 Combining Quality Primitives to form an automatic quality score

In each of the four simulated metrics (i.e., Bishtawi Metric, EWPM, 2CABM and 3CABM), there are two major steps. The first step is to calculate Quality Primitives, and the second step is to find a Quality Primitive combination and use it to calculate the automatic quality scores. The HVS is complex and many of its properties are not well understood [24], therefore it is difficult to find a Quality Primitive combination that matches the properties of the HVS well. Currently, several methods (such as Minkowski metric[22], linear regression and non-linear regression analysis) are used to achieve the Quality Primitive combination. Linear regression analysis is a commonly used method and will be used in this thesis.

In Section 4.2, the first step is described. In this section, the second step is illustrated. Note that, in the Wolf-Pinson metric, the automatic quality scores are calculated with a Quality Primitive combination (i.e., Eqn.2.13) given by Wolf and Pinson.

4.3.1 Achieving a Quality Primitive combination

In each of the four simulated metrics (i.e., Bishtawi Metric, EWPM, 2CABM and 3CABM), prior to producing a Quality Primitive combination, the Quality Primitives of the training video clips listed in Table 4.5 need to be computed. Then, a Quality Primitive combination is obtained by using the following steps.

1. By making linear regression between the Quality Primitive data and the non-automatic quality data, a Quality Primitive combination is achieved, which is denoted as QPC_1. Here, knowledge of linear regression is described in Appendix A.1.
2. By using hypothesis tests introduced in Appendix A.2, the Quality Primitives that have significant effect on the automatic quality scores (i.e., predicted values of non-automatic quality scores) are selected from Quality Primitives in the QPC_1.
3. By making linear regression between the Quality Primitives selected from the step 2 and the non-automatic quality data, a new Quality Primitive combination is achieved, which is denoted as QPC_2.
4. Test if each Quality Primitive in the QPC_2 has a significant effect on the automatic quality scores by using the hypothesis tests. If all the Quality Primitives in the QPC_2 have significant effect on the automatic quality scores, then the QPC_2 is taken as a Quality Primitive combination of the current metric; otherwise, Step 3 and Step 4 are repeated until all the Quality Primitives in the QPC_i (i= 3,4,...) have significant effect on the automatic quality scores.

4.3.2 Description of non-automatic quality data set

In our non-automatic quality data (i.e., quality score data produced by human rating), there are the non-automatic quality scores (see Table 4.5) of eighty-four compressed

video clips. These compressed video clips are generated by twelve ten-second original video clips (see Table 4.6) by using different Quantization factors (Q-factors) of MPEG-2 compression.

Currently, the most primitive kind of constant quality compression is constant Q-factor compression. If the scene characteristic of a compressed video sequence does not change much, constant Q-factor may lead to almost constant quality video. The video clips in table 4.5 belong to constant quality compressed video.

Compressed Video Clips \ Quantization factors (Q-factors)		12	16	24	32	36	40	48
Training Video Clips	Autumn Leaves	2.2	12	17.8	39.7	51.3	54.3	57.2
	Sailboat	5.3	9.3	6.5	10.8	14.6	26.1	25.8
	Flower Garden	8.5	8	10.7	22.7	24.1	30.7	35.6
	Mobile & Calendar	0	0.5	11.7	9.2	29.8	29.5	33.6
	Table Tennis	1.5	7.2	10.2	19.6	31.1	38.7	36.9
	Bette Pas Bette	1.3	8.7	18.6	22.2	41.3	49.2	46.6
	Ferries Wheel	4.6	7	9.3	21.5	35.5	37.3	41.6
	Susie	7.2	15.1	11.1	31.8	53	68	70.4
Non-training Video Clips	Birches	4.5	6.5	7.5	8.3	8.4	21.3	22.6
	Football	1.5	10.2	9.4	37.3	36.8	56.3	46.6
	Horse Riding	3.2	10.6	19.4	28.8	50	52.3	55.5
	Tempete	0.6	2.5	12.1	11.1	23.3	23.7	29.8

Table 4.5 Non-automatic quality scores of the eighty-four compressed video clips

In these eighty-four compressed video clips listed in Table 4.5, totally fifty-six video clips are used as a fixed training set for the linear regression analysis, while twenty-eight video clips are used as a fixed non-training set for testing the performance of the simulated metrics.

In Table 4.5, the data of each row is the non-automatic quality scores of the video clips compressed by using MPEG-2 Encoders with the different Q-factors for an original video. In the above non-automatic quality data, the values of the non-automatic quality scores lie between 0 and 100. Values near 0 indicate a good

video quality, while values near 100 indicate a bad video quality. Value of zero indicates no quality impairment. Figure 4.7 shows non-automatic data curves of twelve groups of the compressed video clips.

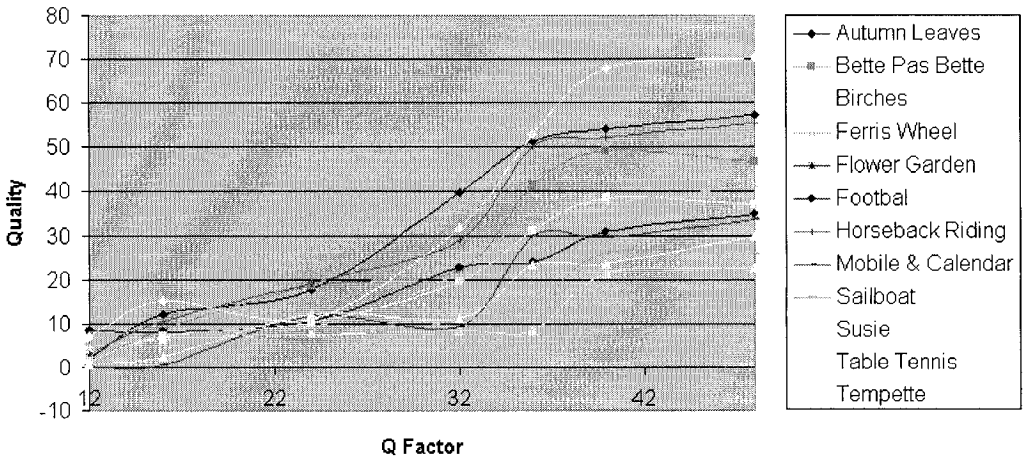


Figure 4.7 Non-automatic data curves of twelve groups of compressed video clips

In Figure 4.7, the horizontal coordinate is the Q-factor, and the vertical coordinate is the non-automatic quality scores. Each curve represents the non-automatic (i.e. those given by human observers) quality scores of the video clips compressed using MPEG-2 Encoders with different Q-factors. Here higher quality scores are worse. Not surprisingly these correspond to sequences compressed with higher Q-factors.

Video Clip Class	Order	Original Video Clips	Characteristics
Training Video Clips	1	Autumn_leaves	Color, landscape, zooming, water fall movement
	2	Sailboat	Almost still
	3	Flower garden	Flower garden
	4	Mobile&calendar	Available in both formats, color, movement
	5	Table Tennis	Table Tennis
	6	Betes_pas_betes	Color, synthetic, movement, scene cut
	7	Ferris_Wheel	Color, movement
	8	Susie	Skin color, almost still
Non-training Video Clips	9	Birches	Almost still
	10	Football	Color, movement
	11	Horsing	Color, movement
	12	Tempete	Color, movement

Table 4.6 The twelve original video clips

In Table 4.6, there are twelve original video clips; the first eight clips are training video clips, the rest ones are non-training video clips. Each of the original video clips consists of 300 frames, 10 seconds in duration with resolution of 720×480.

4.3.3 Quality Primitive Combinations of the four simulated metrics

The Quality Primitive data calculated from the training video clips with the four simulated metrics (i.e., Bishtawi Metric, EWPM, 2CABM and 3CABM) are put in Appendix B.1, which includes the data as follows:

- Quality Primitive data calculated from the training video clips with the Bishtawi Metric are listed in Table B.2.

- Quality Primitive data calculated from the training video clips with the EWPM are listed in Table B.3.
- Quality Primitive data calculated from the training video clips with the 2CABM are listed in Table B.4.
- Quality Primitive data calculated from the training video clips with the 3CABM are listed in Table B.5.

After the above Quality Primitive data are computed, Quality Primitive combinations of the four simulated metrics (i.e., Bishtawi Metric, EWPM, 2CABM and 3CABM) are obtained by using the procedures described in Section 4.3.1. The processes of producing these Quality Primitive combinations are listed in Section B.2.1, Section B.2.2, Section B.2.3 and Section B.2.4 of Appendix B.2.

4.3.4 Automatic quality score calculation

The Quality Primitive data calculated from all the test video clips with the five simulated metrics (i.e., Wolf-Pinson metric, Bishtawi Metric, EWPM, 2CABM and 3CABM) are put in Appendix B.1. After the Quality Primitive data is obtained, the automatic quality scores of the test video clips are computed.

Table 4.7 and Table 4.8 list the automatic quality scores computed from the training and non-training video set with the five simulated metrics. Figure 4.8, Figure 4.9, Figure 4.10 and Figure 4.11 show the scatter plots of non-automatic quality scores versus the automatic quality scores calculated by using the five simulated metrics.

Note that, as mentioned in Section 2.6.6, in the Wolf-Pinson metric, the automatic quality scores range from -1 to 0, while the automatic quality scores calculated by other simulated metrics range from 0 to 100. Table B.13 in Appendix B lists the automatic quality scores calculated from the test video clips with the Wolf-Pinson metric.

For the purpose of comparing the performance of the Wolf-Pinson metric with that of other metrics easily, the data in Table B.13 (which ranges from 1 to 0) will be transformed to the range from 0 to 100 by using Eqn. 4.1.

$$Z = X\beta \tag{4.1}$$

where, $\mathbf{X}=(x_1, x_2 \ldots x_n)'$ is a vector that consists of the automatic quality scores in Table B.13 (in Appendix B.1). In the vector $\mathbf{X}$, x_i (i=1,2...n) ranges from –1 to 0.

Least squares [40] is used to calculate an appropriate β (scaler) so that the components of $\mathbf{Z}$ match the components of $\mathbf{X}$ as closely as possible in a least squares sense. Thus [40],

$$\beta = (X^T X)^{-1} X^T Y \tag{4.2}$$

where, $\mathbf{Y}=(y_1, y_2 \ldots y_n)'$ is a vector that consists of the non-automatic quality scores in Table 4.7 and Table 4.8.

By using the above method, the data ranging from –1 to 0 (i.e., the data in Table B.13) are transformed to the data ranging from 0 to 100 (i.e., the data of the Wolf-Pinson Metric in Table 4.7 and Table 4.8).

Video Clips	Q-Factor	Non-automatic Quality Scores	Wolf-Pinson Metric	Bishtawi Metric	EWPM	2CABM	3CABM
Autumn_Leaves	Q12	2.2	15.1013	3.7071	0.0000	6.9487	12.0524
	Q16	12	19.2724	7.0934	7.1046	13.2869	17.2135
	Q24	17.8	26.7145	21.3136	22.2164	24.015	27.6826
	Q32	39.7	29.9799	34.5774	32.6445	34.0261	38.3412
	Q36	51.3	30.7717	45.1415	37.1124	39.5786	43.8871
	Q40	54.3	32.0903	52.2543	42.4354	45.1047	49.0674
	Q48	57.2	33.136	68.8892	49.5953	54.7694	57.9555
Sailboat	Q12	5.3	16.0676	2.902	0.0000	0.0000	0.2085
	Q16	9.3	19.5471	5.9088	3.5204	0.4281	3.8299
	Q24	6.5	23.9742	13.1652	12.3337	9.0676	10.0536
	Q32	10.8	26.0926	17.9676	20.1629	17.1965	16.8576
	Q36	14.6	26.6777	22.9991	24.441	21.2999	21.2146
	Q40	26.1	27.5687	24.6603	29.1374	25.5427	25.766
	Q48	25.8	28.5564	29.2362	36.58	33.56	34.0683
Flower_Garden	Q12	8.5	13.9043	4.4942	0.0000	0.0000	0.0000
	Q16	8	16.2185	7.4293	0.0000	0.0000	2.262
	Q24	10.7	20.4837	13.7548	4.3018	6.5354	8.583
	Q32	22.7	23.7542	17.2412	11.149	14.5245	14.5488
	Q36	24.1	25.0689	20.8069	15.1628	17.8454	18.0956
	Q40	30.7	26.9224	21.4966	19.7616	21.0997	20.8384
	Q48	35.6	29.1868	26.4997	27.7842	27.5024	26.5795
Mobile_Celendar	Q12	0	16.9575	3.4157	0.1942	0.0000	0.0000
	Q16	0.5	18.9401	7.6717	4.6814	0.0000	0.7304
	Q24	11.7	21.3102	15.9596	14.5561	12.4262	9.7349
	Q32	9.2	22.742	21.4475	25.3462	25.1795	19.5981
	Q36	29.8	23.366	23.3142	30.401	31.0831	25.2842
	Q40	29.5	24.2814	25.1207	36.2193	36.7649	30.7698
	Q48	33.6	25.5121	28.0745	46.1197	47.0039	40.0879
Table_Tennis	Q12	1.5	19.7484	0.9242	6.9219	7.2585	9.5534
	Q16	7.2	25.9806	5.1862	17.0009	14.2737	14.2228
	Q24	10.2	32.0346	16.4279	26.5034	23.138	20.7603
	Q32	19.6	34.1806	24.0202	32.1612	30.2108	26.7262
	Q36	31.1	34.5993	29.3278	34.5958	33.3337	29.3672
	Q40	38.7	35.4653	32.6581	38.1547	36.0728	31.9373
	Q48	36.9	35.8744	35.4364	43.3406	42.8223	38.7588

Table 4.7 (1) The automatic quality scores calculated from the training video clips with the five simulated metrics

Video Clips	Q-factor	Non-automatic Quality Scores	Wolf-Pinson Metric	Bishtawi Metric	EWPM	2CABM	3CABM
Betes_Pas_Betes	Q12	1.3	16.1419	6.9218	1.2501	4.3319	3.8392
	Q16	8.7	19.6831	8.9653	7.614	10.7111	9.0282
	Q24	18.6	24.1715	19.1018	18.7379	22.3606	17.1212
	Q32	22.2	26.5722	27.6335	28.4948	32.6294	24.4022
	Q36	41.3	27.655	35.5266	33.7974	38.0027	29.553
	Q40	49.2	28.863	40.2292	40.0531	42.6927	35.8683
	Q48	46.6	30.0712	49.6398	49.1729	51.9348	45.608
Ferris_Wheel	Q12	4.6	13.7843	3.2904	0.0000	0.0000	0.0000
	Q16	7	16.1219	5.8749	4.967	3.0323	3.05
	Q24	9.3	19.861	14.5647	15.1265	14.0428	12.682
	Q32	21.5	21.9679	20.8861	24.7409	25.3347	23.5726
	Q36	35.5	22.6954	28.9252	30.7973	31.6727	33.0108
	Q40	37.3	24.0374	30.7842	35.2624	36.1422	39.008
	Q48	41.6	25.7663	41.9386	43.8219	45.8467	50.4501
Susie	Q12	7.2	15.874	3.5388	9.425	9.4607	10.906
	Q16	15.1	20.1861	8.621	17.0778	15.2888	16.1439
	Q24	11.1	25.5325	24.1537	28.4229	24.6058	25.5965
	Q32	31.8	27.2969	39.4354	37.3923	35.1116	37.4501
	Q36	53	27.5804	54.6278	43.4755	40.6646	47.541
	Q40	68	28.5244	58.9924	49.3488	45.5638	55.5504
	Q48	70.4	29.0351	75.4944	59.2827	54.6194	71.9411

Table 4.7 (2) The automatic quality scores calculated from the training video clips with the five simulated metrics

Video Clips	Q-factor	Non-automatic Quality Scores	Wolf-Pinson Metric	Bishtawi Metric	EWPM	2CABM	3CABM
Birches	Q12	4.5	13.593	7.1577	0.0000	0.0000	0.0000
	Q16	6.5	15.3148	8.4517	0.0000	0.0000	0.7678
	Q24	7.5	19.074	12.7923	0.0000	5.3849	6.6821
	Q32	8.3	22.8593	15.7218	6.2911	12.5421	11.9672
	Q36	8.4	24.7637	17.273	9.9767	15.944	14.2565
	Q40	21.3	26.9046	17.7451	14.6188	19.6532	16.749
	Q48	22.6	30.0625	18.1969	22.2032	25.7668	20.7287
Football	Q12	1.5	19.3569	27	15.8711	8.0763	9.5695
	Q16	10.2	22.1938	0.2151	22.8846	15.6765	17.655
	Q24	9.4	26.6887	18.9855	33.9971	29.1459	32.2085
	Q32	37.3	28.5068	30.7242	41.092	40.8563	44.2016
	Q36	36.8	28.7862	41.9645	44.3227	45.6638	49.8326
	Q40	56.3	29.5842	46.9519	48.0539	50.5142	55.7812
	Q48	46.6	29.9772	63.0624	54.8183	59.4922	67.3204
Horse_Ridings	Q12	3.2	16.8458	0.0000	3.5087	3.1386	4.9054
	Q16	10.6	21.9134	0.0000	12.1347	9.4484	9.8621
	Q24	19.4	28.651	7.9716	23.372	19.2818	18.4438
	Q32	28.8	31.8853	15.2377	30.7472	27.864	24.5125
	Q36	50	32.8095	19.2048	33.9776	31.5176	37.3125
	Q40	52.3	33.9505	24.1753	38.3515	35.6626	41.5936
	Q48	55.5	35.0154	35.8906	43.6944	42.2299	47.7433
Tempete	Q12	0.6	15.6486	7.1021	9	0.0000	0.0000
	Q16	2.5	18.3111	8.9303	3.7524	3.5269	3.7709
	Q24	12.1	22.5021	14.7623	12.6307	13.7973	11.5982
	Q32	11.1	25.6602	17.3625	21.5335	23.0935	18.6231
	Q36	23.3	26.8103	20.6188	26.3159	28.6659	22.0122
	Q40	23.7	28.1395	20.8417	30.7254	32.5088	25.8422
	Q48	29.8	29.6285	25.635	38.6027	41.8979	33.2559

Table 4.8 The automatic quality scores calculated from the non-training video clips with the five simulated metrics

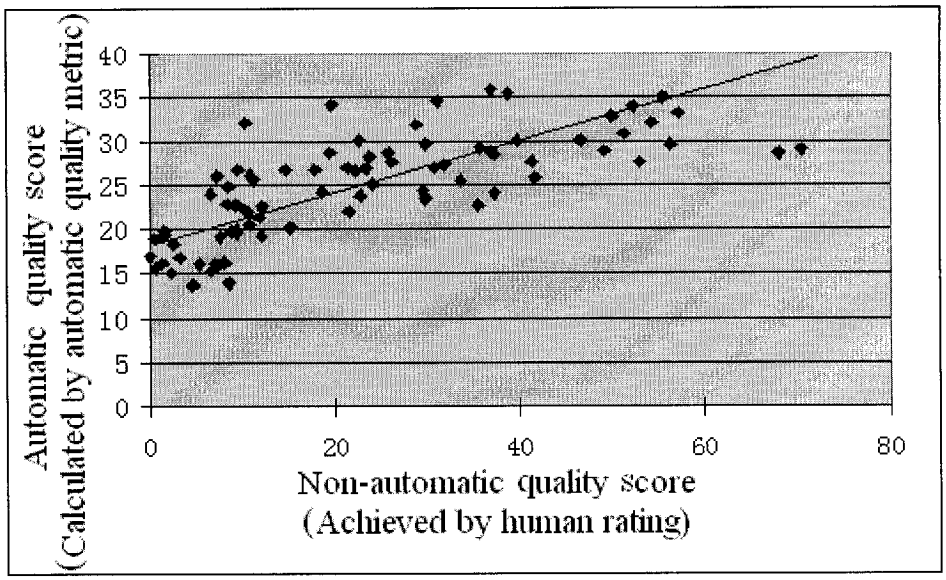


Figure 4.8 Scatter plot of the non-automatic quality scores versus the automatic quality scores calculated by using the Wolf-Pinson Metric

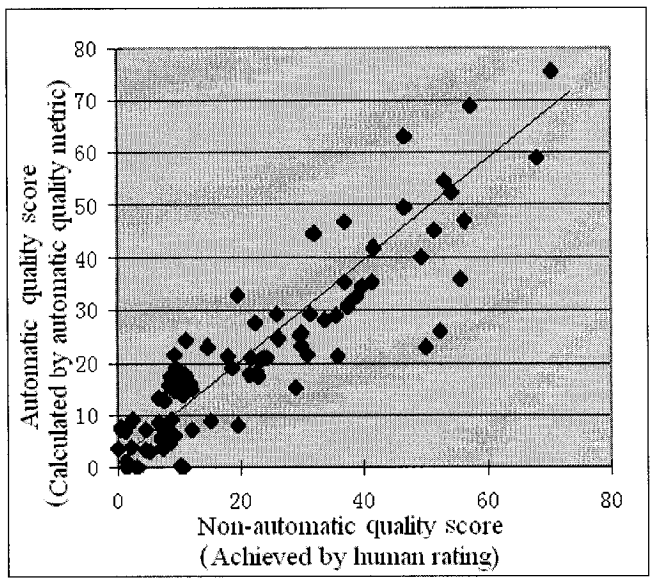


Figure 4.9 Scatter plot of the non-automatic quality scores versus the automatic quality scores calculated by using the Bishtawi Metric

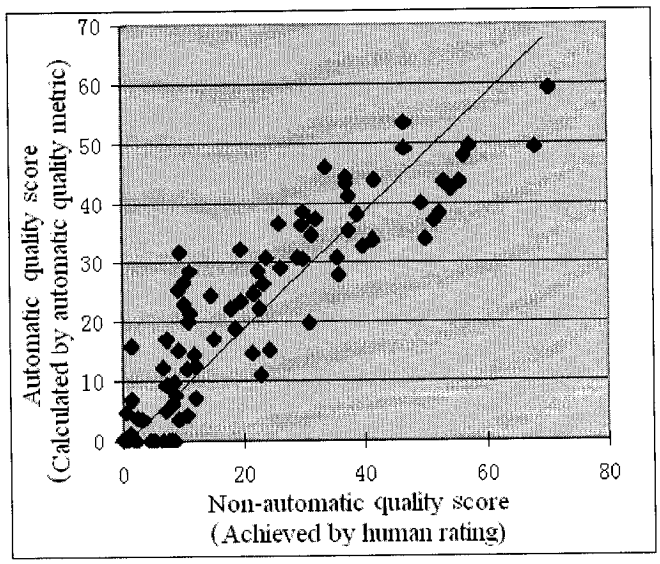


Figure 4.10 Scatter plot of the non-automatic quality scores versus the automatic quality scores calculated by using the EWPM

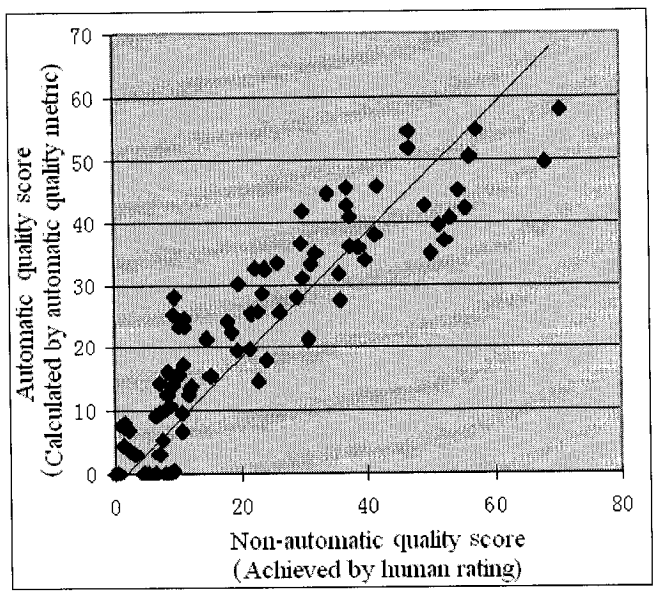


Figure 4.11 Scatter plot of the non-automatic quality scores versus the automatic quality scores calculated by using the 2CABM

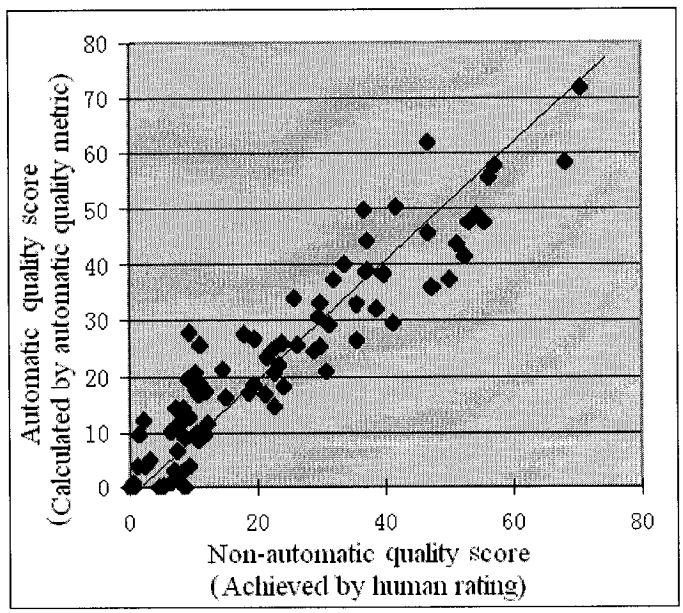


Figure 4.12 Scatter plot of the non-automatic quality scores versus the automatic quality scores calculated by using the 3CABM

The above five scatter plots suggest that the predicted accuracy of the EWPM, 2CABM and 3CABM are better than those of the Wolf-Pinson metric and the Bishtawi metric because the scatter points in Figure 4.10, Figure 4.11 and Figure 4.12 are located closer to the plot diagonals than the scatter points in Figure 4.8 and Figure 4.9. In the next section, the performance of the five simulated metrics is precisely evaluated by using mathematical methods.

4.4 Figures of merit

The performance of the simulated metrics is evaluated with respect prediction accuracy and prediction monotonicity in relation to the human observations [14].

.

4.4.1 Prediction accuracy

The prediction accuracy is the ability of the metric to predict the non-automatic quality score (i.e., quality score obtained by human visual rating) in a minimum average error. The Pearson linear correlation coefficient (denoted as ρ1) and Mean Square Error (MSE) are two common methods used to measure the average error [14].

Here, ρ_1 represents a Pearson linear correlation coefficient between the non-automatic quality scores and the automatic quality scores computed by the simulated metrics. ρ_1 is defined as follows [35]:

$$\rho_1 = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{(n(\sum x^2) - (\sum x)^2)(n(\sum y^2) - (\sum y)^2)}} \tag{4.3}$$

where, n is the sample size. x represents the non-automatic quality scores, while y represents the automatic quality scores. $\sum x$ is the sum of the individual values in the non-automatic quality scores. $(\sum x)^2$ is the square of summation of the individual values in the non-automatic quality scores. $(\sum x^2)$ is the summation of square of the individual values in the non-automatic quality scores. $(\sum xy)$ is the sum of the product of each pair of the non-automatic quality scores and the automatic quality scores.

The value of ρ_1 lies between +1 and –1. A value near +1 indicates a strong positive linear correlation, while a value near -1 indicates a strong negative linear correlation. A value near zero indicates no linear correlation.

Here, the MSE represents Mean Squared Error between the non-automatic quality scores and the automatic quality scores. MSE is defined as follows [35].

$$MSE = \sum (D)^2 / n \qquad (4.4)$$

Where, n is the sample size. D is the differences between a non-automatic quality score and an automatic quality score. $\sum(D^2)$ is calculated over all n pairs.

Smaller MSE implies better predication accuracy of the metric.

4.4.2 Prediction Monotonicity

The relationship between automatic quality scores (i.e., quality scores calculated by using an automatic quality metric) and non-automatic quality scores (i.e., quality scores produced by human rating) should ideally be monotonic[14]. That is a bigger automatic quality score for test video sequences imply a bigger non-automatic quality score. This attribute is called prediction monotonicity. Figure 4.13 and Figure 4.14 below show hypothetical relationships between automatic quality scores and non-automatic quality scores for two automatic metrics. Of these two metrics, the metric of Figure 4.13 has better monotonicity. The metric in Figure 4.14 is less monotonic because in Figure 4.14, the automatic quality score of Point A is bigger than that of Point B, while the non-automatic quality score of Point A is smaller than that of Point B.

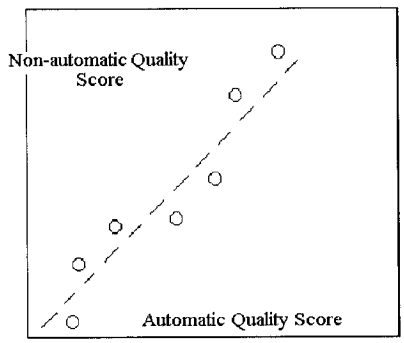


Figure 4.13 A model with more monotonicity [14]

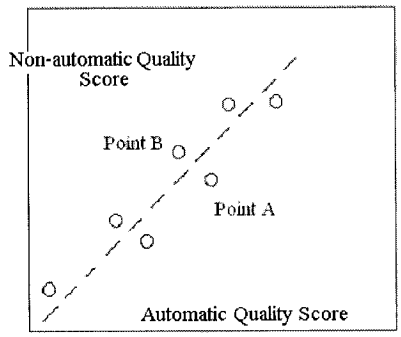


Figure 4.14 A metric with less monotonicity [14]

As mentioned in chapter 1, the Spearman rank-order correlation coefficient (denoted as ρ_2) is a measure of monotonicity. ρ_2 is defined as follows [35].

$$\rho_2 = 1 - \frac{6 \times \Sigma(D^2)}{n(n^2 - 1)} \qquad (4.5)$$

where, n is the sample size. D is the differences of the non-automatic quality scores and the automatic quality scores. In this equation, $\Sigma(D^2)$ is calculated over all n pairs.

Values of ρ_2 lay between +1 and −1. A value near +1 indicates a strong positive monotonicity, while a value near -1 indicates a strong negative monotonicity. A value near zero indicates no relationship.

4.5 Simulation results

Table 4.9 lists ρ_1 (the Pearson linear correlation coefficient), ρ_2 (the Spearman rank-order correlation coefficient) and **MSE** (Mean Squared Error) between the non-automatic quality data and the automatic quality data calculated by using the five simulated metrics.

Video clip Classes	Statistical Parameter	Wolf-Pinson Metric	Bishtawi Metric	EWPM	2CABM	3CABM
Training Video Clip Set (56 video clips)	ρ_1	0.6959	0.8723	0.8843	0.8940	0.9319
	ρ_2	0.5946	0.8611	0.8657	0.8759	0.9192
	MSE	211.79	71.91	70.15	64.79	42.20
Non-training Video Clip Set (28 video clips)	ρ_1	0.8496	0.7831	0.8563	0.8642	0.8973
	ρ_2	-0.3804	-0.0005	0.3433	0.3704	0.4855
	MSE	201.24	130.56	85.69	82.16	67.13
All video clips (84 video clips)	ρ_1	0.7408	0.8551	0.8745	0.8834	0.9184
	ρ_2	0.8288	0.9148	0.9359	0.9399	0.9570
	MSE	207.67	91.27	75.33	70.58	50.51

Table 4.9 ρ_1, ρ_2 and **MSE** between the non-automatic and automatic quality scores computed with the five simulated metrics

The statistical data in Table 4.9 indicate that

- Among the five simulated metrics, the performance of the 3CABM is the best because the ρ_1's and ρ_2's of the 3CABM are the closet to one, as well as the **MSE**s of the 3CABM are the smallest.

- The performance of the EWPM and 2CABM is close because the ρ_1's, ρ_2's and **MSE**s of the EWPM and 2CABM are close.
- The performance of the Wolf-Pinson Metric and Bishtawi Metric is worse than those of other metrics because the ρ_1's and ρ_2's of the Wolf-Pinson metric and Bishtawi metric are smaller than those of the other metrics, as well as their **MSE**s are more than those of other metrics.

4.5.1 Discussion

The simulation results suggest that in all the simulated metrics, the 3CABM has the best performance (prediction accuracy and prediction monotonicity). As mentioned in Section B.2.4 of Appendix B.2, by using statistical analysis, the following results are obtained,

- In the twelve Quality Primitives of the 3CABM, six Quality Primitives (i.e., *Qp_fb_f1_gain, Qp_fb_f2_gain, Qp_tb_f1_loss, Qp_seb_f1_gain, Qp_tb_f2_loss* and *Qp_seb_f2_loss*) have significant effect on the automatic quality scores, and then they take calculation of the automatic quality scores (see Regression equation B.8 in Table B.28 of Appendix B.2.4).
- However, the left six Quality Primitives (i.e., *Qp_fb_f1_loss, Qp_fb_f2_loss, Qp_tb_f1_gain, Qp_tb_f2_gain, Qp_seb_f2_gain* and *Qp_seb_f1_loss*) appear to have little effect on the automatic quality scores.

For the purpose of analyzing the above Quality Primitives further, Table 4.10 lists the relationship between the above Quality Primitives and artifacts.

Quality Primitive Class	Order	Quality Primitive	Relationship Between Artifacts And Quality Primitives
The Quality Primitives that have significant effect on automatic quality scores	1	*Qp_fb_f1_gain*	*Qp_fb_f1_gain* reflects Ringing degree of flat blocks of the compressed video.
	2	*Qp_fb_f2_gain*	*Qp_fb_f2_gain* reflects Blocking degree of flat blocks of the compressed video.
	3	*Qp_tb_f1_loss*	*Qp_tb_f1_loss* reflects Blurring degree of texture blocks of the compressed video.
	4	*Qp_tb_f2_loss*	*Qp_tb_f2_loss* does not have the monotonous increase or decrease relationships with the three artifacts.
	5	*Qp_seb_f1_gain*	*Qp_seb_f1_gain* reflects Ringing degree of sharp edge blocks of the compressed video.
	6	*Qp_seb_f2_loss*	*Qp_seb_f2_loss* does not have the monotonous increase or decrease relationships with the three artifacts.
The Quality Primitives that appear to have little effect on automatic quality scores	7	*Qp_fb_f1_loss*	*Qp_fb_f1_loss* reflects Blurring degree of flat blocks of the compressed video.
	8	*Qp_fb_f2_loss*	*Qp_fb_f2_loss* does not have the monotonous increase or decrease relationships with the three artifacts.
	9	*Qp_tb_f1_gain*	*Qp_tb_f1_gain* reflects Ringing degree of texture blocks of the compressed video.
	10	*Qp_tb_f2_gain*	*Qp_tb_f2_gain* reflects Blocking degree of texture blocks of the compressed video.
	11	*Qp_seb_f1_loss*	*Qp_seb_f1_loss* reflects Blurring degree of sharp edge blocks of the compressed video.
	12	*Qp_seb_f2_gain*	*Qp_seb_f2_gain* reflects Blocking degree of sharp edge blocks of the compressed video.

Table 4.10 Relationship between artifacts and Quality Primitives

Table 4.10 suggests that in the 3CABM, the Blocking of flat blocks has more effect on automatic quality scores than the Blocking of texture blocks and of sharp edge blocks, and the Ringing of flat blocks and sharp edge blocks has more effect on automatic quality scores than the Ringing of texture blocks.

From the above analysis, it may be seen that the results achieved by the statistical analysis is consistent with the Masking property of the HVS as mentioned in Section 3.2; that is,

- the Blocking in the blocks with less spatial details (i.e., flat blocks) is easily perceivable by the HVS, while the Blocking may be masked in the blocks with more spatial details (i.e., texture blocks and sharp edge blocks)[32].
- the Ringing may be masked in texture blocks, while it may be easily observed near sharp edges in sharp edge blocks, as well as near weak edges of flat blocks [34].

The simulation results suggest that the 3CABM has significantly better performance than that of the other simulated metrics.

In this chapter, five automatic metrics based on the measures of three spatial artifacts (i.e., Blurring, Blocking and Ringing) are described, and then their performance is evaluated. In the next chapter, an artifact, Block Flashing, is defined and measured, which differs from these spatial artifacts; it is a temporal and spatial artifact.

Chapter 5 Blocking Flashing Primitives

5.1 Introduction

In this chapter, an artifact Block Flashing is introduced, and then a new primitive based on Block Flashing is developed. Afterwards, an Artifact-Measured-based Automatic Metric (AMAM) is proposed, which may detect Blurring, Blocking, Ringing and Block Flashing.

This chapter is organized as follows. Block Flashing is described in Section 5.2. Estimation of degree of Block Flashing is presented in Section 5.3. Then, an Artifacts-Measured-based Automatic Metric (AMAM) is introduced in Section 5.4. Section 5.5 gives a Quality Primitive combination of the AMAM. Afterwards, the simulation result of the AMAM is stated in Section 5.6.

5.2Block Flashing

As mentioned in Chapter 1, the Block Flashing artifact is that some blocks flash when a compressed video sequence is played. How is Block Flashing artifact produced? An illustrative example follows.

Figure 5.1 and Figure 5.2 show the first frame images of an original video clip "Sailboat" and of a compressed video clip "Sailboat.24.mpeg2", respectively. Here, "Sailboat" is an original video sequence with the resolution of 720×480 pixels;

"Sailboat.24.mpeg2" is a compressed video sequence produced by encoding "Sailboat" with the MPEG-2 encoder (the quantization factor is 48). Each of these is 10 seconds in duration, with 30 frames per second (300 frames in 10 seconds). In these two video sequences, each video frame consists of 5400 blocks, of size 8×8 pixels.

In Figure 5.1, an 8×8 block located at (21, 57) is marked with a white square. Here, this block is denoted as block(21,57), 21 is the horizontal coordinate, which means the current block is the 21^{st} block counted from the left to the right in the horizontal direction, while 57 is the vertical coordinate, which means the current block is the 57^{th} block counted from up to down in the vertical direction. In Figure 5.2, two blocks block(21,57) and block(9,3) are also marked with the white squares, respectively.

By our observation, the block(21,57) of the "Sailboat.24.mpeg2" is a typical flashing block while the block(9, 3) is not, when "Sailboat.24.mpeg2" is played.

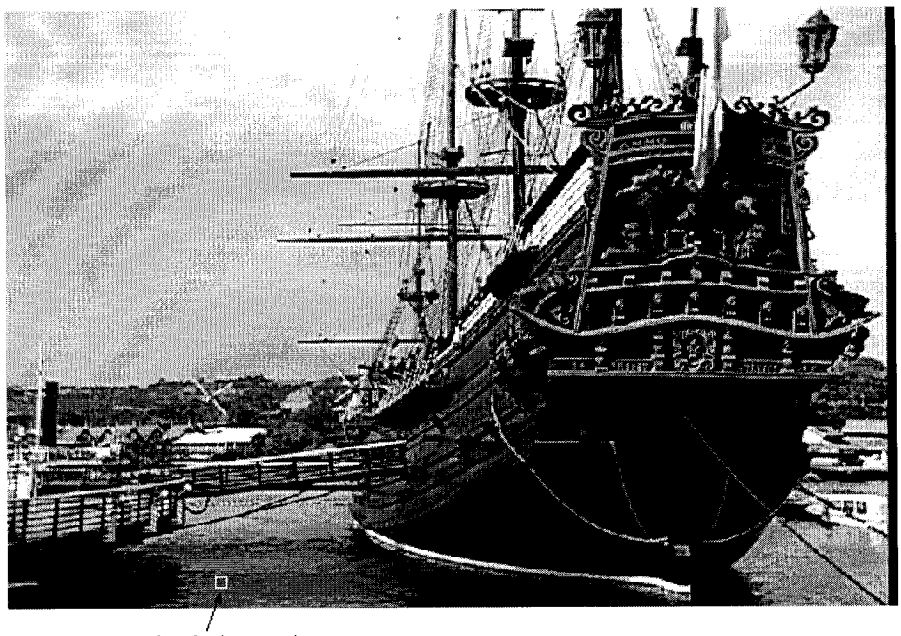


Figure 5.1 The first frame image of the original video clip "Sailboat"

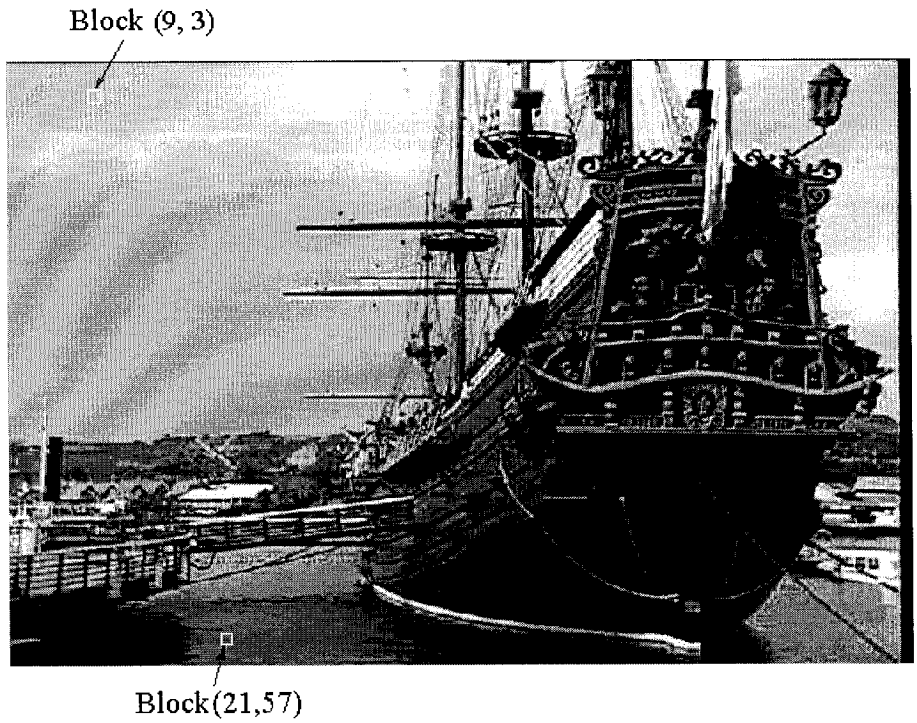


Figure 5.2 The first frame image of the compressed video clip "Sailboat.24.mpeg2"

Here, "Sailboat.24.mpeg2" is produced by encoding "Sailboat" with MPEG-2 encoder.

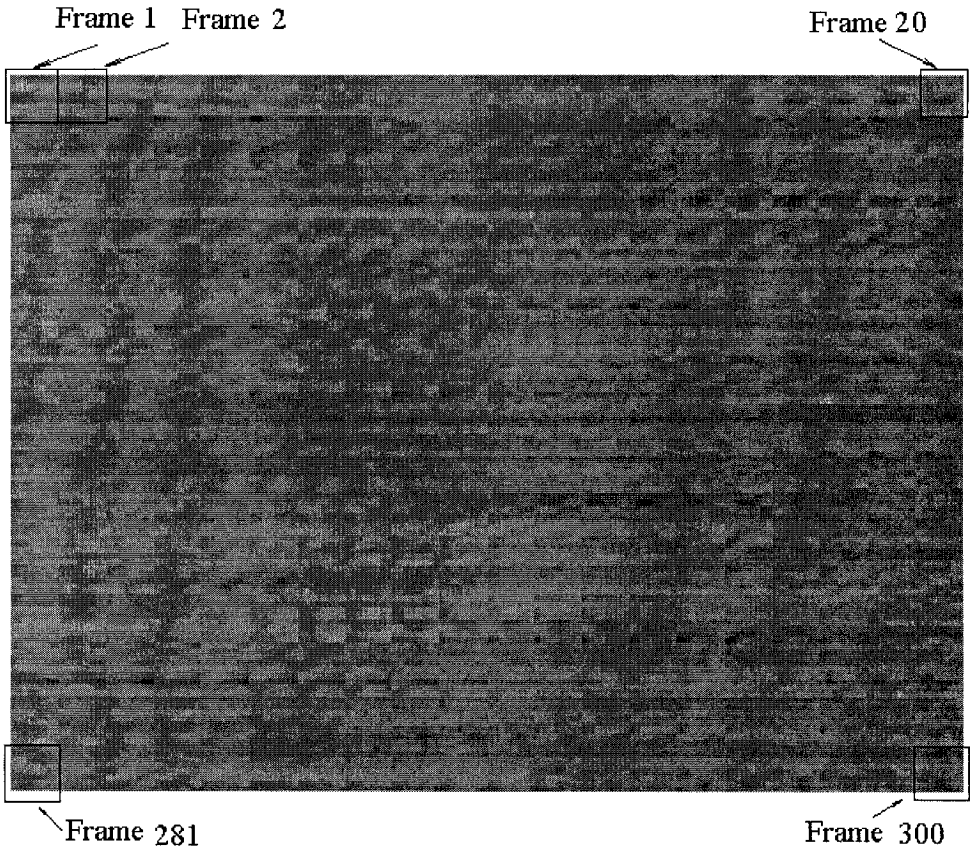


Figure 5.3 Block (21, 57) from the 1st to the 300th frame in the original video "Sailboat"

In Figure 5.3, there are 300 8×8 blocks, the order of which is $block_{f1}$, $block_{f2}$, ... $block_{f300}$ according to the direction from the left to the right and from up to down. Here, $block_{fi}$ represents the block of the i^{th} frame.

957.75	1.44	2.18	-3.66	2.75	-3.86	-7.13	-0.55
-2.55	6.59	1.82	-8.49	-5.93	-1.75	-0.40	-2.62
-7.91	14.71	-2.89	1.68	8.06	3.25	-2.90	-1.98
-10.27	19.81	3.88	-2.35	2.84	-4.27	1.98	1.52
-52.50	-19.99	-6.77	0.83	-3.50	0.97	0.63	1.78
7.89	2.60	2.51	9.63	3.87	3.91	-0.98	-2.54
5.71	7.70	1.84	-0.92	-1.44	5.03	0.64	-1.42
-2.05	0.78	-3.89	-1.87	3.01	5.88	0.08	-1.14

Table 5.1 2D DCT coefficients of the block of the first frame in Figure 5.3

In Table 5.1, there are 64 coefficients; the coefficient 957.75 located in the first row and the first column is a Direct Current (DC) coefficient, while the other coefficients are Alternating Current (AC) coefficients.

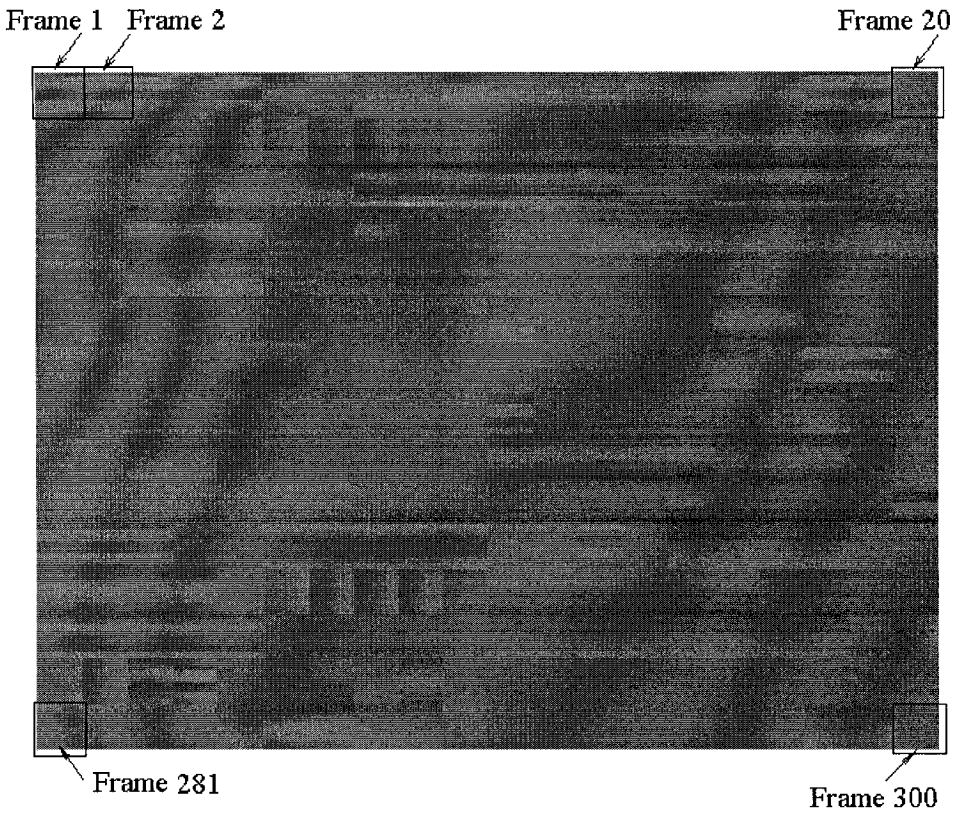


Figure 5.4 Block (21, 57) from the 1st to the 300th frame in the compressed video "Sailboat.24.mpeg2"

In Figure 5.4, there are 300 8×8 blocks, the order of which is $block_{f1}$, $block_{f2}$, … $block_{f300}$ according to the direction from left to right and from up to down.

960	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
-64	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

Table 5.2 2D DCT coefficients of the block in the first frame in Figure 5.4

In Table 5.2, there are 64 coefficients; the coefficient 960 located in the first row and the first column is a DC coefficient, while the other coefficients are AC coefficients.

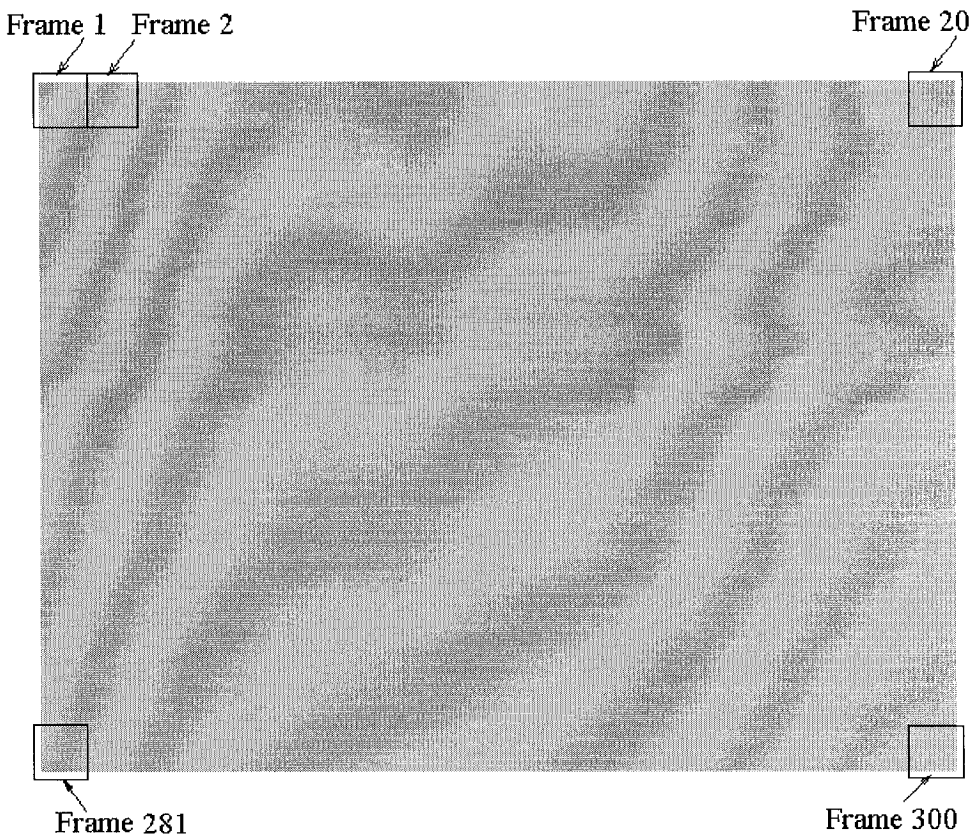


Figure 5.5 Block (9, 3) from the 1st to the 300th frame in the compressed video "Sailboat.24.mpeg2"

In Figure 5.5, there are 300 8×8 blocks, the order of which is $block_{f1}$, $block_{f2}$, … $block_{f300}$ according to the direction from left to right and from up to down.

1480	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	0	0	0	0	0	0	0

Table 5.3 2D DCT coefficients of the block in the first frame in Figure 5.5

In Table 5.3, there are 64 coefficients; the coefficient 1480 located in the first row and the first column is a DC coefficient, while the other coefficients are AC coefficients.

In examining Figure 5.3, Figure 5.4 and Figure 5.5, two attributes of these figures are interesting. First note that in terms of spatial detail, the most detailed is Figure 5.3 (original block (21,57) observed as "not flashing") followed by Figure 5.4 (compressed block (21,57) observed as "flashing"), while Figure 5.5 showed the least spatial detail (compressed block (9,3) observed as "not flashing"). Second note that Figure 5.5 shows much less variation in the time direction than the other two. The conclusion that is drawn from these observations is that "flashing" blocks are those with most spatial detail wiped out, but that still exhibit variation over time. Note that variation over time generally implies that not all spatial detail be wiped out. But as Figure 5.3 shows, a good deal of spatial detail can mask the "flashing" artifact.

Table 5.1, Table 5.2 and Table 5.3 list the Two Dimension Discrete Cosine Transform (2D DCT) coefficients of the first blocks in Figure 5.3, Figure 5.4 and Figure 5.5, respectively. From these three tables, it is obvious that there are more values of zero in Table 5.2 and Table 5.3 than in Table 5.1; suggesting that block(21,57) and block(9,3) of the first frame in the compressed video sequence (see Figure 5.4 and Figure 5.5) have less spatial detail than the block(21,57) of the first frame in the original video sequence (see Figure 5.3). In fact, the block(21,57) and block(9, 3) not only of the first frame but also of the other frames in the compressed video sequence have less spatial detail than those in the original video sequence.

The observations on the DCT coefficients are consistent with the conclusions of the visual inspection of Figure 5.3, Figure 5.4 and Figure 5.5. These conclusions lead

to the proposed block flashing primitive presented below.

5.3 Estimation of Degree of Block Flashing

The block diagram of Block Flashing primitive is shown in Figure 5.6. Each step in Figure 5.6 will be described in the following subsections.

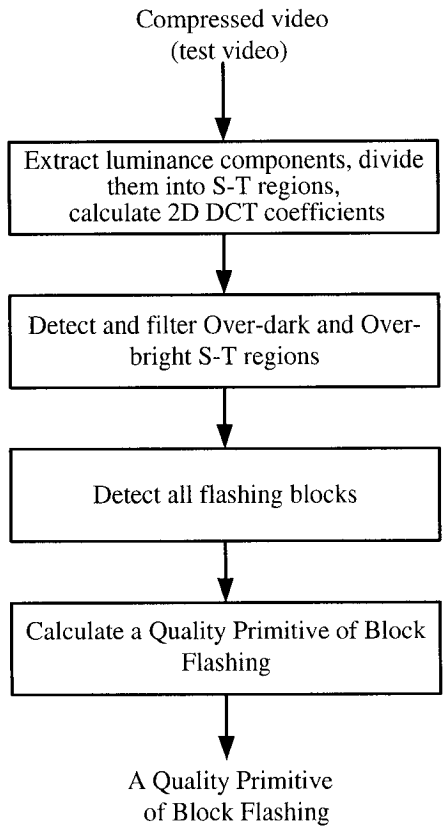


Figure 5.6 The block diagram of Block Flashing primitive

5.3.1 Extract luminance components and compute 2D DCT coefficients

In measurement of Block Flashing, luminance components (i.e., luminance images) of a compressed video sequence are extracted first because only these components are required. Then, these luminance images need to be divided into S-T regions with the

size of 8 horizontal pixels × 8 vertical pixels × the temporal width △t (5 seconds), as shown in Figure 5.7.

Note that, in the simulated metrics described in Section 4.2 (i.e., Wolf-Pinson Metric, EWPM, 2CABM and 3CABM), the temporal width △t of an S-T region is 1/30 second (i.e., a frame), while in measurement of Block Flashing, the temporal width △t of an S-T region is 5 seconds (i.e., 150 frames) because the Block Flashing happens in a duration.

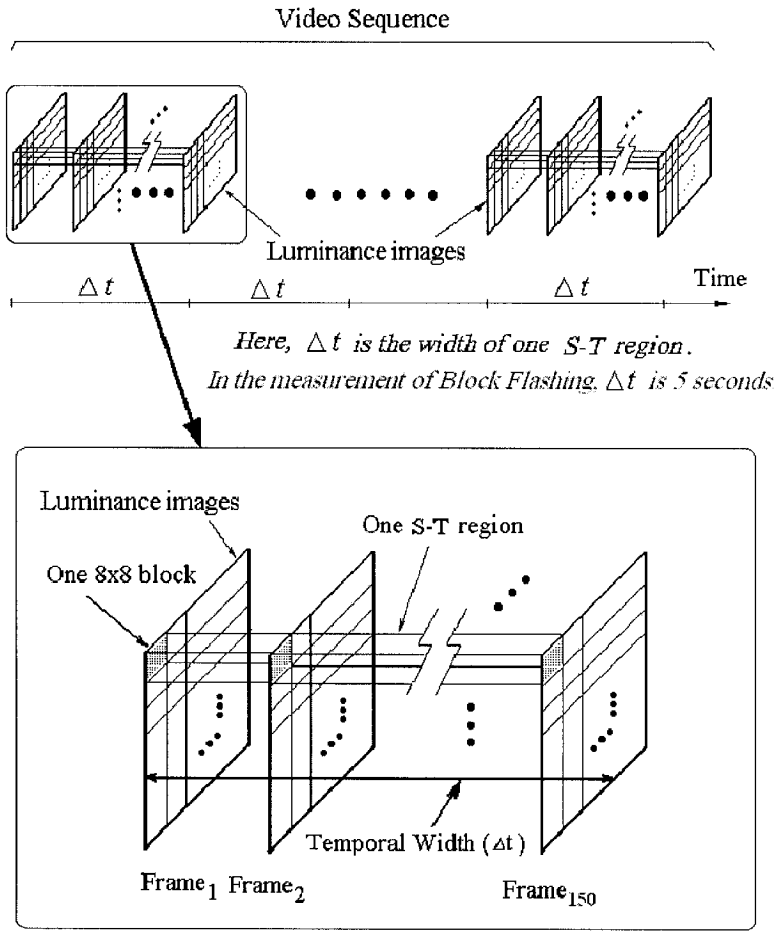


Figure 5.7 Illustration of the S-T regions

Afterwards, the 2D DCT of each block in the S-T regions is computed. The 2D

DCT is given by [2] as:

$$F(u,v)=\frac{2}{N}C(u)C(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}f(x,y)\cos\frac{(2x+1)u\pi}{2N}\cos\frac{(2y+1)v\pi}{2N} \qquad (5.1)$$

$$C(u),C(v)=\begin{cases}\frac{1}{\sqrt{2}} & for\ u,v=0\\ 1 & otherwise\end{cases}$$

Where, N is the block size, which equals to 8.

5.3.2 Over-bright and Over-dark S-T regions

In [38], it is shown that the HVS is sensitive to artifacts in mid-luminance range of brightness while is not sensitive to artifacts in Over-dark and Over-bright areas. Thus Block Flashing, may not be observed in Over-dark and Over-bright regions. Figure 5.8 shows an Over-dark block and an Over-bright block.

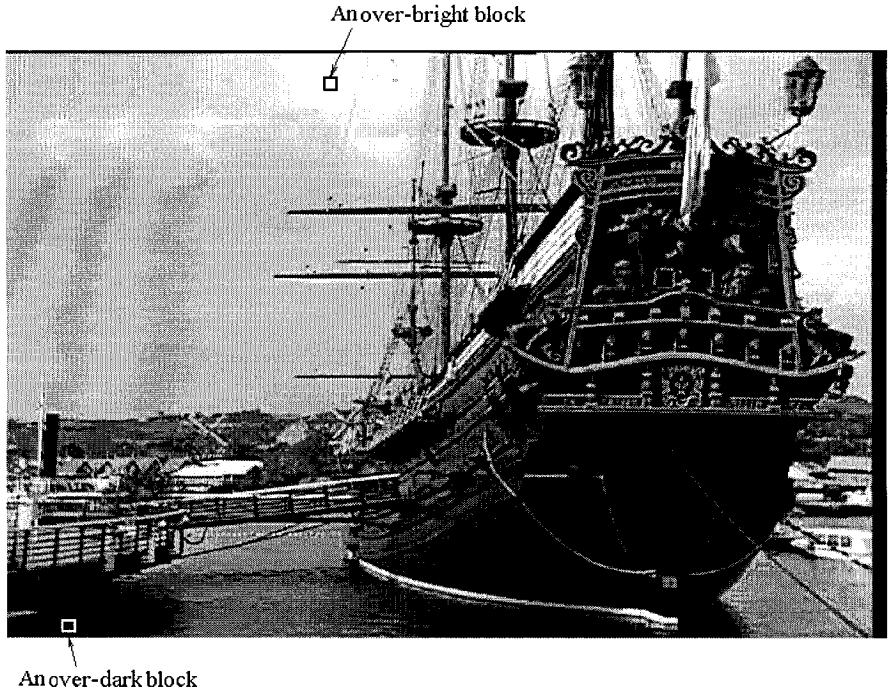


Figure 5.8 An Over-dark block and an Over-bright block

In the detection of the Over-dark and Over-bright S-T regions, the average of the DC coefficients of all blocks in an S-T region is calculated. Then, two thresholds, T_1 (i.e., Over-bright threshold) and T_2 (i.e., Over-dark threshold) are employed. The condition of detection is listed in Table 5.4.

S-T Region Class	Detection condition
An Over-bright S-T region	• Average of DC coefficients of all blocks in an S-T region $\geqslant T_1$
An Over-dark S-T region	• Average of DC coefficients of all blocks in an S-T region $\leqslant T_2$

Table 5.4 The detection condition of the Over–bright and Over-dark S-T regions

Here, the selections of T_1 and T_2 values are discussed in Section 5.3.5.

Note that blocks that are labeled as Over-dark or Over-bright are considered to not have any Block Flashing and no processing is done on these blocks in subsequent steps.

5.3.3 Detection of Block Flashing

Based on the conclusions in Section 5.2, the two following conditions for Block Flashing to be present in an S-T region are formulated:

1. Number of non-zero AC coefficients of all blocks in an S-T region $\leq T_3$, and
2. Variance of DC coefficients of all blocks in this S-T region$\geq T_4$.

S-T regions that satisfy these two conditions (as well as not being either Over-dark or Over-bright as outlined in the section above) are labeled as having Block Flashing.

In the detection method, two thresholds T_3 (i.e., frequency threshold) and T_4 (i.e., variance threshold) are employed. Their value selection is discussed in Section 5.3.5.

5.3.4 Calculation of Block Flashing Quality Primitive

After all the S-T regions of Block Flashing are detected from a compressed video sequence by using the method described in the above section, a Quality Primitive of Block Flashing is calculated with the equation as follows:

$$Qp\_bf = \frac{Number\ of\ S-T\ regions\ of\ Block\ Flashing}{Total\ number\ of\ S-T\ regions\ of\ compressed\ video\ sequence} \quad (5.2)$$

5.3.5 Selection of thresholds

In the Block Flashing detection algorithm described in the above section, the four thresholds T_1, T_2, T_3 and T_4, are used. In this section, these thresholds are set so as to minimize misses and false alarms. (A miss is an S-T region that truly has Block Flashing but that the proposed algorithm does not label as having Block Flashing. A false alarm is an S-T region that does not have Block Flashing but that the proposed

algorithm labels as having Block Flashing. These definitions are the standard ones in detection theory [39].)

In order to find values of four thresholds, it is necessary to have video where S-T regions are labeled accurately as either having Block Flashing or not. To this end, the sequence "Sailboat.24.mpeg" was observed by the author (Mr. Wei Dai) and another thesis student Mr. Zhen Cai who independently labeled each S-T region as either having Block Flashing or not. Finally, the S-T regions where only if both observers agreed that the Block Flashing occurred were label as having Block Flashing.

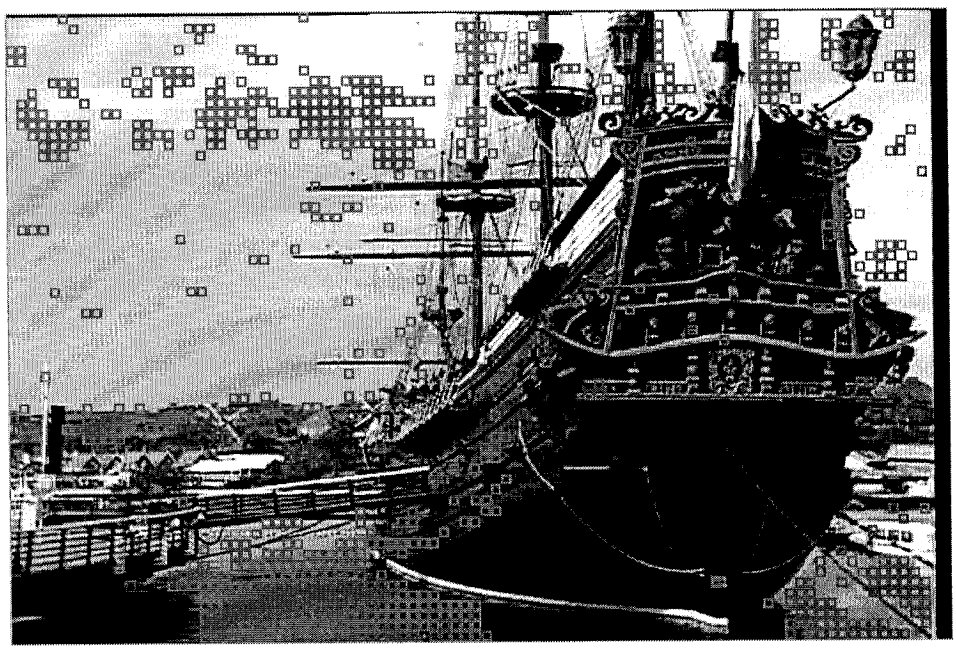


Figure 5.9 All flashing blocks selected manually from the compressed video clip "Sailboat.24.mpeg2"

In the process of selecting values of four thresholds, the miss rate and false alarm rate are used, which are defined as follows [39];

$$Miss\ rate = \frac{Number\ of\ flashing\ blocks\ missed}{Number\ of\ flashing\ blocks} \tag{5.3}$$

$$False-alarm\ rate = \frac{Number\ of\ flashing\ blocks\ false-alarmed}{Total\ number\ of\ blocks\ -\ Number\ of\ flashing\ blocks} \tag{5.4}$$

The miss rate and false alarm rate range from 0 to 1. A higher value means a worse measurement, while a smaller value means a better measurement.

In order to find a suitable threshold combination, the four thresholds are initialized as T_1=1800, T_2=60, T_3=320 and T_4=100 first. Then each of the four thresholds is selected, respectively.

In the selection of T_1 value, the three thresholds T_2 (=60), T_3 (=320) and T_4 (=100) are frozen. By using different T_1 values, the false alarm rates and the miss rates are computed, which are listed in Table 5.5; the relationship between them is represented in a curve as shown in Figure 5.10. From Table 5.5 and Figure 5.10, it may be seen that when the miss rate is getting smaller, the false alarm rate is getting bigger; therefore, there exists no T_1 value that can make the miss rate and false alarm rate be the smallest, concurrently. Thus, the selection of T_1 value depends on which rate between the miss rate and false alarm rate is more important for Block Flashing detection. In this thesis, there is no special requirement for the false alarm rate and the miss rate; therefore, T_1 value is selected by using the following method.

- From the relationship curve in Figure 5.10, a point (i.e., Point A) closer to the

original point (i.e., (0, 0)) is selected, then the miss rate and the false alarm rate corresponding to Point A are obtained from Figure 5.10. By looking up Table 5.5, a T_1 value corresponding to Point A (i.e., T_1= 1780) is selected.

	T_1	Number of flashing blocks selected by program	Number of blocks false-alarmed	Number of blocks missed	False alarm rate	Miss rate
1	1700	587	61	215	1.3%	29.0%
2	1740	645	76	172	1.6%	23.2%
3	1780	690	88	139	1.9%	18.7%
4	1820	720	93	131	1.99%	17.3%
5	1860	744	107	104	2.3%	14.0%
6	1900	761	119	99	2.5%	13.3%
7	1940	774	130	97	2.8%	13.0%
8	1980	777	133	97	2.85%	13.0%
9	2020	777	133	97	2.85%	13.0%
10	2060	777	133	97	2.85%	13.0%

Table 5.5 The false alarm rates and the miss rates calculated by using the different T_1 s

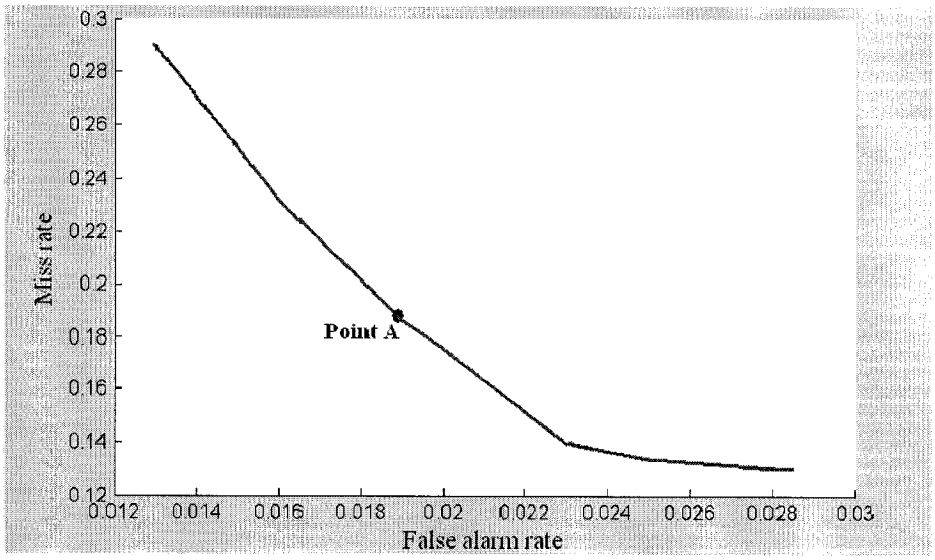


Figure 5.10 A relationship curve between the false alarm rates and the miss rates calculated by using the different T_1s (see Table 5.5)

In the selection of T_2 value, the three thresholds T_1 (=1780), T_3 (=320) and T_4 (=100) are fixed. By using the different T_2 values, the false alarm rates and the miss rates are

computed, which are listed in Table 5.6; Figure 5.11 shows the relationship curve between them. Then, a T_2 value is selected by using the following method.

- From the relationship curve in Figure 5.11, a point (i.e., Point B) closer to the original point (i.e., (0,0)) is selected; then, the miss rate and the false alarm rate corresponding to Point B are obtained from Figure 5.11. By looking up Table 5.6, a T_2value corresponding to Point B (i.e., T_2= 30) is selected.

	T_2	Number of flashing blocks selected by program	Number of blocks false-alarmed	Number of blocks missed	False alarm rate	Miss rate
1	4	1553	848	36	18.2%	4.86%
2	7	1331	630	40	13.5%	5.39%
3	10	1176	483	48	10.4%	6.4%
4	20	1024	340	57	7.3%	7.6%
5	30	967	289	63	6.2%	8.5%
6	40	902	242	81	5.2%	10.9%
7	50	821	185	105	3.9%	14.2%
8	60	726	116	114	2.5%	15.4%
9	70	661	63	143	1.4%	19.2%
10	80	630	46	157	0.9%	21.2%
11	90	600	31	172	0.6%	23.2%
12	100	567	11	185	0.2%	24.9%

Table 5.6 The false alarm rates and the miss rates calculated by using the different T_2s

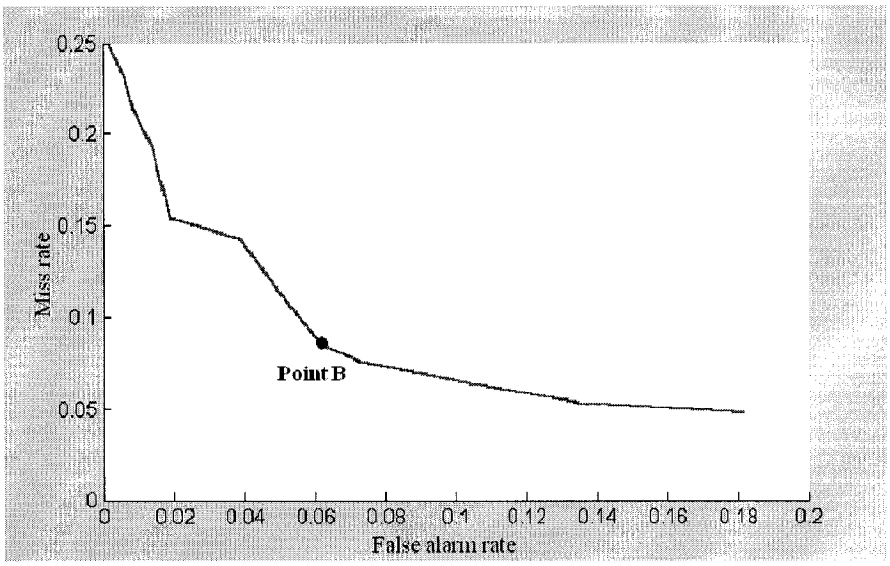


Figure 5.11 A relationship curve between the false alarm rates and the miss rates calculated by using the different T_2s (see Table 5.6)

In the selection of T_3 value, the three thresholds T_1(=1780), T_2 (=30) and T_4 (=100) are frozen. With the different T_3 values, the false alarm rates and the miss rates are computed, which are listed in Table 5.7. Figure 5.12 shows the relationship curve between the false alarm rates and the miss rates calculated by using the different T_3s. Then, a T_3 value is selected by using the following method.

- From the relationship curve in Figure 5.12, a point (i.e., Point C) closer to the original point (i.e., (0,0)) is selected, then the miss rate and the false alarm rate corresponding to Point C are obtained from Figure 5.12. By looking up Table 5.7, a T_3 value corresponding to Point C (i.e., T_3=400) is selected.

	T_3	Number of flashing blocks selected by program	Number of blocks false-alarmed	Number of blocks missed	False alarm rate	Miss rate
1	300	1493	797	45	17.1%	6.1%
2	320	1404	717	54	15.3%	7.2%
3	340	1267	586	60	12.5%	8.1%
4	360	1012	348	77	7.5%	10.3%
5	380	812	166	95	3.5%	12.8%
6	400	778	139	102	2.9%	13.8%
7	420	724	121	114	2.6%	15.7%
8	440	546	89	284	1.91%	38.3%

Table 5.7 The false alarm rates and the miss rates calculated by using the different T_3s

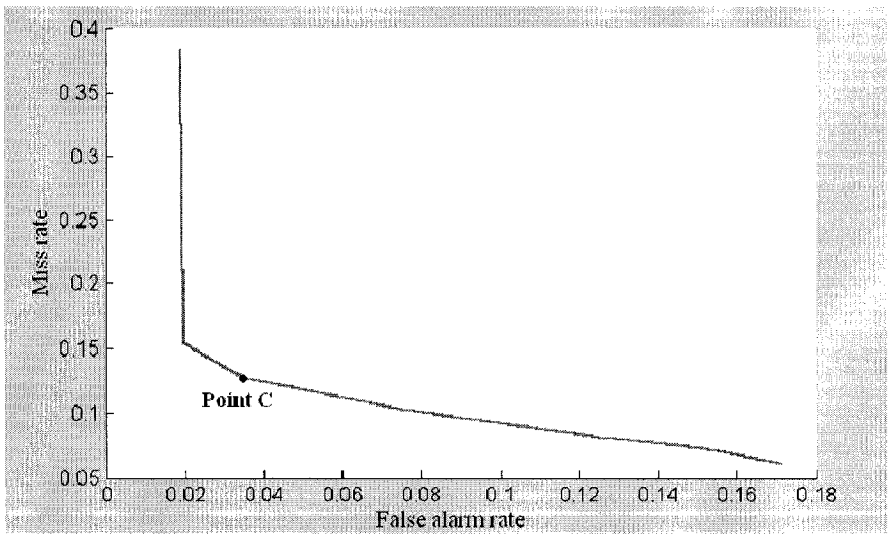


Figure 5.12 A relationship curve between the false alarm rates and the miss rates calculated by using the different T_3s (see Table 5.7)

In the selection of T_4 value, the three thresholds T_1 (=1780), T_2 (=30) and T_3 (=400) are fixed. With the different T_4 values, the false alarm rates and the miss rates are computed and listed in Table 5.8. Figure 5.13 shows the relationship curve between them. Then, a T_4 value is selected by using the following method.

- From the relationship curve in Figure 5.13, a point (i.e., Point D) closer to the original point (i.e., (0,0)) is selected, then the miss rate and false alarm rate corresponding to Point D are obtained from Figure 5.13. By looking up Table

5.8, a T_4 value corresponding to Point D (i.e., T_2= 73.1) is selected.

	T_4	Number of flashing blocks selected by program	Number of blocks false-alarmed	Number of blocks missed	False alarm rate	Miss rate
1	244.32	481	17	277	0.36%	37.38%
2	234.15	491	19	269	0.40%	36.31%
3	221.06	501	23	263	0.49%	35.49%
4	205.76	511	24	254	0.51%	34.27%
5	192.75	521	27	247	0.57%	33.33%
6	186.74	531	30	240	0.64%	32.38%
7	178.77	541	31	231	0.66%	31.17%
8	173.29	551	32	222	0.69%	29.95%
9	164.62	561	34	214	0.73%	28.87%
10	159.60	571	35	205	0.75%	27.66%
11	149.52	581	36	196	0.77%	26.45%
12	144.66	591	38	188	0.82%	25.37%
13	138.27	601	39	179	0.84%	24.15%
14	133.06	611	43	173	0.92%	23.34%
15	127.60	621	45	165	0.96%	22.26%
16	125.84	631	50	160	1.07%	21.59%
17	123.08	641	56	156	1.20%	21.05%
18	121.32	651	62	152	1.33%	20.51%
19	118.85	661	70	150	1.50%	20.24%
20	114.11	671	73	143	1.57%	19.29%
21	109.99	681	79	139	1.69%	18.75%
22	106.99	691	80	130	1.71%	17.54%
23	103.80	701	84	124	1.80%	16.73%
24	98.00	721	93	113	1.99%	15.24%
25	95.31	731	101	111	2.16%	14.97%
26	89.21	741	108	108	2.31%	14.57%
27	85.97	751	114	104	2.45%	14.04%
28	83.06	761	121	101	2.59%	13.63%
29	80.16	771	127	97	2.73%	13.09%
30	77.16	781	134	94	2.88%	12.69%
31	74.96	791	142	92	3.05%	12.42%
32	73.10	801	150	90	3.2%	12.15%
33	57.96	869	209	81	4.5%	10.87%
34	38.2	998	320	63	6.87%	8.50%
35	26.25	1112	424	53	9.10%	7.12%
36	19.89	1182	484	43	10.4%	5.80%
37	15.71	1350	631	22	13.5%	3.00%
38	14.32	1471	741	11	15.9%	1.48%

Table 5.8 The false alarm rates and the miss rates calculated by using the different T_4s

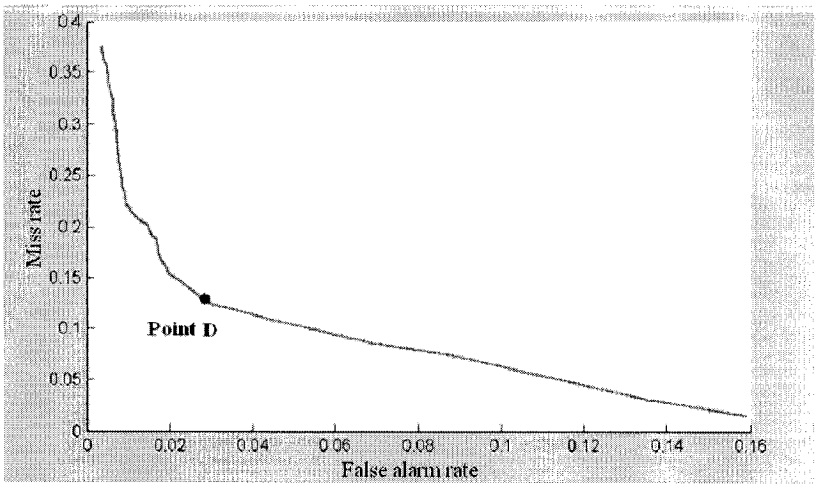


Figure 5.13 A relationship curve between the false alarm rates and the miss rates calculated by using different T_4s (see Table 5.8)

5.4 An Artifacts-Measured-based Automatic Metric (AMAM)

In Chapter 4, the four automatic metrics (i.e. Wolf-Pinson Metric, Bishtawi Metric, EWPM, 2CABM, 3CABM) are simulated. The simulation results suggest that the 3CABM has the best performance. In this section, based on the 3CABM and Block Flashing detection technique, an Artifacts-Measured-based Automatic Metric (AMAM) is proposed. In this metric, after Quality Primitives of the four artifacts Blocking, Blurring, Ringing and Block Flashing are calculated, they are formed into an overall measure by using a Quality Primitive combination.

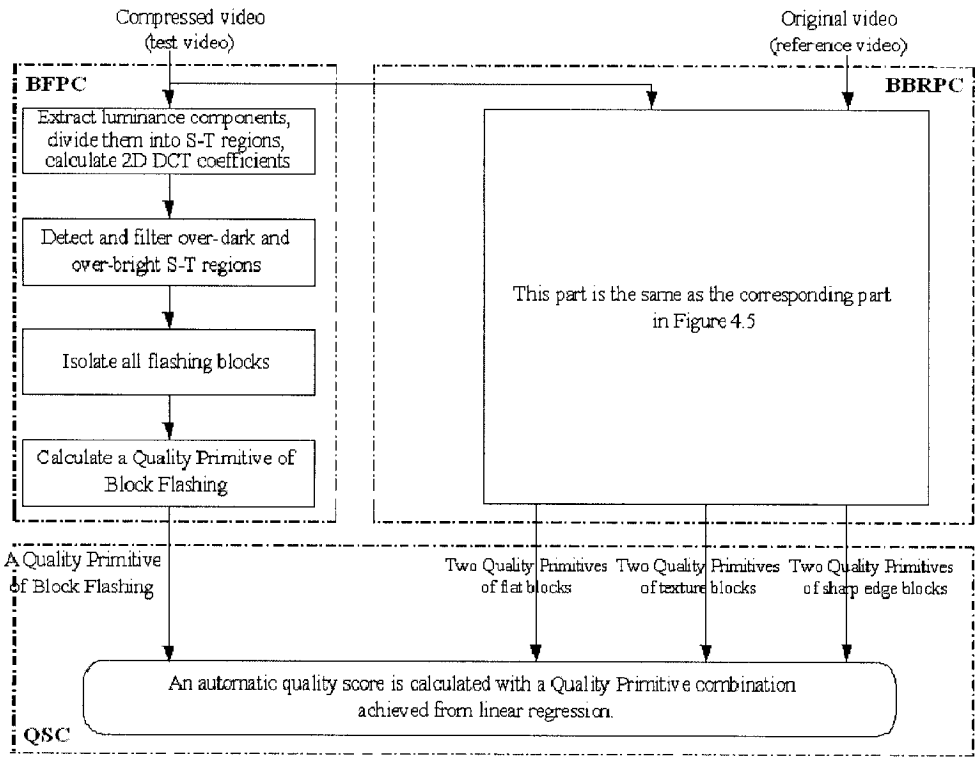


Figure 5.14 The block diagram of the AMAM

The block diagram of AMAM is shown in Figure 5.14, which is composed of three parts, Block Flashing Primitive Calculation (BFPC), Blurring, Blocking, Ringing Primitive Calculation (BBRPC) and Quality Score Calculation (QSC).

5.5 Quality Primitive Combination of the AMAM

In Section 5.3, the BFPC has been described. In Section 4.2, the BBRPC has been illustrated. In the following, the QSC is introduced.

In order to obtain a Quality Primitive combination utilized to calculate quality score, the Quality Primitive data of test video clips listed in Table 4.3 need to be calculated

by using the AMAM. These calculated data are put in Appendix B, which includes the following data:

- Quality Primitive data calculated from the training video clips with the AMAM are listed in Table B.11.
- Quality Primitive data calculated from the non-training video clips with the AMAM are listed in Table B.12.

After the above Quality Primitive data are achieved, a Quality Primitive combination of the AMAM is obtained by using the method described in Section 4.3.1. This process is listed Section B.2.5 of Appendix B.2.

5.6 Simulation results

Table 5.9 and Table 5.10 list the automatic quality scores calculated from training and non-training video clips with the AMAM. Figure 5.15 shows the scatter plot of non-automatic quality scores versus automatic quality scores calculated by using the AMAM. Table 5.11 lists simulation results of the 3CABM and AMAM.

Video Clips	Q-Factor	Non-automatic Quality Scores	AMAM	Video Clips	Q-Factor	Non-automatic Quality Scores	AMAM
Autumn_Leaves	Q12	2.2	8.6091	Table_Tennis	Q12	1.5	8.2811
	Q16	12	13.05		Q16	7.2	14.4868
	Q24	17.8	24.4078		Q24	10.2	22.8407
	Q32	39.7	36.77		Q32	19.6	28.4821
	Q36	51.3	44.0217		Q36	31.1	31.7116
	Q40	54.3	48.6784		Q40	38.7	34.1627
	Q48	57.2	55.9118		Q48	36.9	40.0068
Sailboat	Q12	5.3	0.0000	Betes_Pas_Betes	Q12	1.3	7.5993
	Q16	9.3	1.9698		Q16	8.7	12.4823
	Q24	6.5	7.4202		Q24	18.6	18.0484
	Q32	10.8	13.5442		Q32	22.2	23.1143
	Q36	14.6	18.3801		Q36	41.3	28.2113
	Q40	26.1	23.1216		Q40	49.2	35.4999
	Q48	25.8	31.508		Q48	46.6	45.1797
Flower_Garden	Q12	8.5	1.9687	Ferris_Wheel	Q12	4.6	0.0000
	Q16	8	4.8777		Q16	7	3.6993
	Q24	10.7	10.1375		Q24	9.3	12.1525
	Q32	22.7	15.1041		Q32	21.5	22.0779
	Q36	24.1	19.1472		Q36	35.5	34.1048
	Q40	30.7	21.8262		Q40	37.3	40.9585
	Q48	35.6	27.6773		Q48	41.6	53.3379
Mobile_Celendar	Q12	0	0.0000	Susie	Q12	7.2	9.6693
	Q16	0.5	3.0272		Q16	15.1	16.6766
	Q24	11.7	10.0585		Q24	11.1	26.0615
	Q32	9.2	18.4218		Q32	31.8	36.745
	Q36	29.8	24.5481		Q36	53	47.9021
	Q40	29.5	30.017		Q40	68	57.0701
	Q48	33.6	38.9234		Q48	70.4	74.6319

Table 5.9 The automatic quality scores calculated from the training video clips with the AMAM

Video Clips	Q-Factor	Non-automatic Quality Scores	AMAM	Video Clips	Q-Factor	Non-automatic Quality Scores	AMAM
Birches	Q12	4.5	0.0000	Horse_Ridings	Q12	3.2	6.3855
	Q16	6.5	0.0000		Q16	10.6	10.9721
	Q24	7.5	3.9164		Q24	19.4	20.0429
	Q32	8.3	7.895		Q32	28.8	27.1817
	Q36	8.4	10.00		Q36	50	38.3035
	Q40	21.3	12.4201		Q40	52.3	42.8476
	Q48	22.6	16.236		Q48	55.5	48.2129
Football	Q12	1.5	4.8552	Tempete	Q12	0.6	1.3402
	Q16	10.2	18.4229		Q16	2.5	4.5949
	Q24	9.4	32.9615		Q24	12.1	10.5762
	Q32	37.3	43.5197		Q32	11.1	15.9926
	Q36	36.8	49.2199		Q36	23.3	19.6071
	Q40	56.3	55.3022		Q40	23.7	23.6503
	Q48	46.6	67.5856		Q48	29.8	31.1301

Table 5.10 The automatic quality scores calculated from the non-training video clips with the AMAM

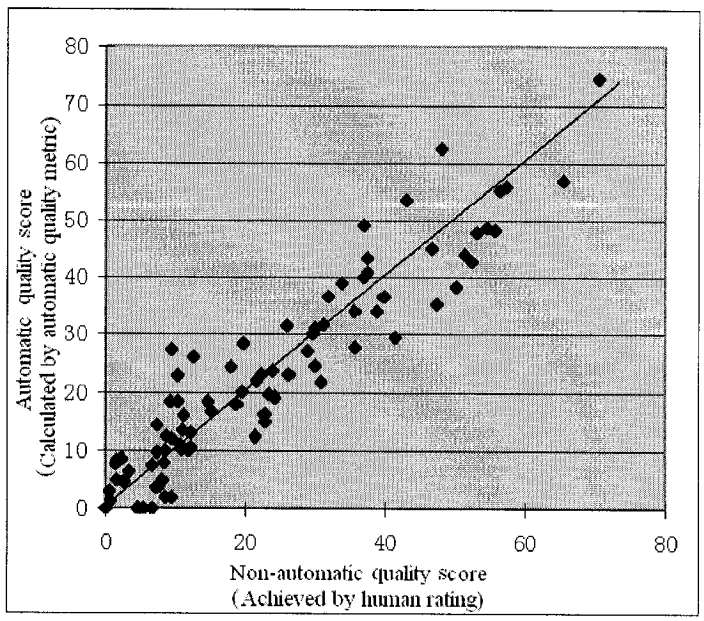


Figure 5.15 Scatter plot of non-automatic quality scores versus automatic quality scores calculated using the AMAM

Figure 5.15 suggests that the AMAM has better predicted accuracy than those of the Wolf-Pinson metric and Bishtawi metric since the scatter points in Figure 5.15 are located closer to the plot diagonal than the scatter points in Figure 4.8 and Figure 4.9.

The performance of the AMAM is further evaluated using ρ_1, ρ_2 and **MSE**.

Video clip Classes	Statistical Parameter	3CABM	AMAM
Training Video Clip Set (56 video clips)	ρ_1	0.9319	0.9382
	ρ_2	0.9192	0.9263
	MSE	42.20	38.48
Non-training Video Clip Set (28 video clips)	ρ_1	0.8973	0.9001
	ρ_2	0.4855	0.4973
	MSE	67.13	65.59
All video clips (84 video clips)	ρ_1	0.9184	0.9235
	ρ_2	0.9570	0.9595
	MSE	50.51	47.52

Table 5.11 Simulation results of the 3CABM and AMAM

From the statistical data in Table 5.11, it is obvious that the ρ_1's and ρ_2's of the AMAM are closer to one than those of the 3CABM, as well as the **MSE**s of the AMAM are smaller than those of the 3CABM. The above analysis indicates that the predicted performance of the AMAM is better than that of the 3CABM. As mentioned in Section 5.4, Quality Primitives of the AMAM consist of Quality Primitives of 3CABM and a Quality Primitive of Block Flashing. Thus, the above statement indicates that Blocking Flashing measurement is successful in the AMAM, and it may improve the performance of the automatic metrics.

Chapter 6 Conclusions and Future Work

6.1 Contribution

This thesis presents an Artifact-Measure-based Automatic Metric (AMAM), which is based on measures of four artifacts Blocking, Ringing, Blurring and Block Flashing. The main contributions of this work are as follows:

- **Taking advantage of the Masking phenomenon of the HVS to measure the Blocking and Ringing**

  In the AMAM, the 8×8 blocks of images in video sequences are classified into flat blocks, sharp edge blocks and texture blocks. Then, Quality Primitives of Blocking, Blurring and Ringing are computed from each class of blocks, respectively. Afterwards, by using the statistical analysis methods (i.e., linear regression and hypothesis test described in Appendix A), the Quality Primitives that have significant effect on automatic quality scores are selected, which take part in the automatic quality score calculation.

- **Defining Block Flashing**

  The Block Flashing is an artifact that differs from the other three artifacts (i.e., Blocking, Ringing and Blurring). It is only observed in a video sequence instead of in a single frame. In some cases, where a block stands out against the rest of the video it will be more objectionable because it is varying over time (i.e. flashing) than if it was not. Thus this thesis distinguishes between blocking (spatial artifact only), and block flashing (spatial temporal artifact). Block

flashing may merit more weight than blocking when it is used to construct a Quality Score.

- **Developing an algorithm of Block Flashing measurement**

  An algorithm of Block Flashing measurement is designed and developed; the simulation results indicate that this algorithm is successful.

6.2 Conclusions

The following conclusions are drawn from this thesis:

- The performance of the automatic quality metrics may be improved by taking advantage of the Masking phenomenon of the HVS to measure the Blocking and Ringing.
- The performance of the automatic quality metrics may be improved by detecting Block Flashing.

6.3 Future works

In order to improve the performance of the AMAM further, the following directions are suggested in the future work.

- The performance of the AMAM should be evaluated by using more test video sequences. The test sequences should not be limited on MPEG-2, but rather be extended to MPEG-1 and MPEG-4 video sequences. Here, the AMAM may directly measure MPEG-1 and MPEG-4 video without any modification.

- The block classification algorithm used in AMAM could be further refined in order to improve the block classification accuracy.
- The HVS weights artifacts of compressed video sequences according to a non-linear and a complex way based on the artifact types and their locations. Thus, non-linear regression analysis may be applied to produce Quality Primitive combinations of simulated metrics.
- In order to improve the accuracy of Block Flashing measurement, the four threshold values (i.e., T_1 =1780, T_2 =30 and T_3 =400 and T_4= 73.1) should be adjusted depending on the various scenes in video sequences. The relationship between the four thresholds and the changes of scenes needs to be investigated in order to make the thresholds adaptive to different scenes.
- In the AMAM, because all Quality Primitives are computed just from luminance images of video sequence, they do not reflect the influences of color artifacts for video quality. Meanwhile, color bleeding and color fading are two main color artifacts. In the future work, these two artifacts should be detected.

Bibliography

[1] W. Effelsberg, "Video compression techniques," Dpunkt-Verlag, Heidelberg, Germany, pp. 23, 1998.

[2] Y.Wang, J. Ostermann and Y. Q. Zhang, "Video Processing and Communications," Prentice Hall, pp. 406, 2001.

[3] "American National Standard for Telecommunications, Digital Transport of Video Teleconferencing/Video Telephony Signals Performance Terms, Definitions, and Examples," ANSI T1.801.02-1995, American National Standards Institute, New York, 1995.

[4]K.Taylor, C.Carey-Smith, and I. Goodwin, "Objective Assessment of video Coding Techniques," *Proceeding SPIE*, Auckland, New Zealand, vol. 1524, pp. 252, October 1993.

[5] Hans Marmolin, "Subjective MSE Measures," *IEEE Transactions on Systems*, Man, and Cybernetics, vol. SMC16, pp. 486-489, June 1986.

[6] Ahmet M. Eskicioglu, Paul S. Fisher, and Siyuan Chen, "Image Quality Measures and Their Performance," *Space and Earth Science Data Compression Workshop, Proceeding SPIE*, Conference Publication 3255, Salt Lake City, Utah, pp. 55-67, April 1994.

[7] A.M.Eskicioglu, "An Improved Graphical Quality Measure for Monochrome Compressed Images," *Optical Engineering Midwest '95, Proceeding SPIE*, Chicago, vol. 1913, pp. 164, May 1995.

[8] A. B. Watson, "DCT quantization matrices visually optimized for individual

images," *Proceeding SPIE*, vol. 1913, pp. 202-216, 1993.

[9] B. Girod, "What's wrong with mean-squared error," *Digital Images and Human Vision* (A. B. Watson, ed.), pp. 207-220, the MIT press, 1993.

[10] P. C. Teo and D. J. Heeger, "Perceptual image distortion," *Proceeding SPIE*, vol. 2179, pp. 127-141, 1994.

[11] K.T. Tan, M. Ghanbari and D.E. Pearson, "A video distortion meter," *Picture Coding Symposium*, pp.119-122, 1997.

[12] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment," pp. 28-41, March 2000

[13] VQEG home page: http://www.its.bldrdoc.gov/vqeg/

[14] VQEG, "FR-TV Full Reference Television Phase II Subjective Test Plan," pp 28-29, September 2002. http://www.its.bldrdoc.gov/vqeg/

[15] VQEG, "RRNR-TV Group Test Plan Draft Version 1.4," September 2002. http://www.its.bldrdoc.gov/vqeg/

[16] VQEG, "Draft Final report from the video quality experts group on the validation of objective models of video quality assessment," pp.7-8, August 2003. http://www.its.bldrdoc.gov/vqeg/

[17] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Processing Letters*, vol. 9, pp. 81-84, March 2002.

[18] Z. Wang, H. R. Sheikh, and A. C. Bovik, "Objective video quality assessment," The Handbook of Video Databases: *Design and Applications* (B. Furht and O. Marques, eds.), CRC Press, 2003.

[19] H.peterson, A.J.Ahumada, Jr. and A.Watson, "An Improved Detection Model for DCT Coefficient Quantization," *Proceeding SPIE*, vol. 1913, pp. 191-120, 1983

[20] S. A. Karunasekera and N. G. Kingsbury, "A Distortion Measure for Blocking Artifacts in Images Based on Human Visual Sensitivity" *IEEE Transaction on Image Processing*, vol. 4, No. 6, pp. 713-724, 1995.

[21] A.B. Watson, "Toward a perceptual video quality metric," *Human Vision, Visual Processing and Digital Display*, vol. 3299, pp. 139-147, 1998.

[22] A. B. Watson, "Design and performance of a digital video quality metric," *Human Vision and Electronic Imaging IV*, Vol. 3644, B.E. Rogowitz and T.N. Pappas eds., SPIE, Bellingham, WA, 1999.

[23] S. Winkler, "Issues in vision modeling for perceptual video quality assessment," *Signal Processing*, vol. 78, no. 2, pp. 231-252, 1999.

[24] S. Winkler, "A perceptual distortion metric for digital color video," *Human Vision and Electronic Imaging, Proceeding SPIE*, vol. 3644, pp. 175-184, 1999.

[25] A. A. Webster, C. T. Jones, M.H. Pinson, S. D. Voran, and S. Wolf, "An Objective Video Quality Assessment System Based on Human Perception," *Human Vision, Visual Processing and Digital Display, Proceeding SPIE*, vol. 1913, San Jose, California, February 1993.

[26] S. Wolf, M. Pinson, C. Jones, and A. Webster, "A Summary of Methods of Measurement for Objective Video Quality Parameters Based on the Sobel Filtered Image and Motion Difference Image," T1A1.93/152, NTIA/ITS, November 1993.

[27] S. Wolf and M. Pinson, "In-service performance metrics for MPEG-2 video

systems," Made to Measure 98 - Measurement Techniques of the Digital Age Technical Seminar," technical conference jointly sponsored by the International Academy of Broadcasting (IAB), the ITU, and the Technical University of Braunschweig (TUB), Montreux, Switzerland, November 1998.

[28] S. Wolf, and M. Pinson, "Spatial-temporal distortion Metric for in-service quality monitoring of any digital video system," *Proceeding SPIE*, Boston, MA, September 1999.

[29] S.Wolf and M.Pinson: "Video Quality Measurement Techniques," NTIA Report 02-392, 2002.

[30] EePing Ong, Weisi Lin, Zhongkang Lu, Susu Yao, Xiaokang Yang, Lijun Jiang, "No-reference jpeg-2000 image quality metric,", IEEE International Conference on Multimedia. March 2003

[31] ITU-R Recommendation BT.500-8: "Methodology for the subjective assessment of the quality of television pictures." ITU, Geneva, Switzerland, 1998.

[32] S. A. Karunasekera and N. G. Kingsbury, "A Distortion Measure for Blocking Artifacts in Images Based on Human Visual Sensitivity," *IEEE Transactions on image Processing*, Vol.4, NO.6, June 1995.

[33] K. T. Tan, M. Ghanbari, and D. E. Pearson, "An objective measurement tool for MPEG video quality," *Signal Processing*, vol.70, no.3, pp. 279-294, November 1998.

[34] T. N. Pappas, and R. J. Safranek, "Perceptual criteria for image quality evaluation," Handbook of *Image and Video Processing*, A. C. Bovik, ed., Academic Press, may 2000.

[35] L. Sachs "Applied Statistics," Allyn and Bacon, Boston, pp.396, pp.413-414, 1984.

[36] R. C. Gonzalez and R. E. Woods, "Digital Image Processing," AddisonWesley, 1992.

[37]Wajih W. Bishtawi, "Objective measurement of subjective image quality," Master's thesis, Department of Electrical and Computer Engineering, Concordia University, 1996,

[38] C. H. Chou and Y.C.Li, "A perceptually tuned subband image coder based on the measure of Just-Noticeable-Distortion profile," *IEEE Trans. Circuits & System for Video Technology*, vol. 5, pp. 467-476, December 1995.

[39] D. Kazakos, "Detection and estimation," *Computer Science Press*, pp.78, 1990.

[40]B. Abraham and J. Ledolter, "Statistical Methods for Forecasting," Wiley, pp. 8-27, 1998

Appendix A

A.1 Linear regression

The purpose of linear regression is to establish a best linear relationship between an independent variable y and a group of independent variables x_i (i=1, 2...p). The linear regression model is as follows [40]:

$$y = \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p + \varepsilon \qquad (A.1)$$

where, y is an independent variable, $x_1, x_2, ... x_p$ are a group of independent variables. p is the number of independent variables x_i. $\beta_1, \beta_2, ... \beta_p$ are a group of regression parameters, and ε is a random disturbance, which follows a Gaussian distribution (or nearly so).

The linear regression model may also be written as Eqn. A.2 by using matrix notation [40]:

$$\boldsymbol{y}=\boldsymbol{X\beta}+\boldsymbol{\varepsilon} \qquad (A.2)$$

where, $\boldsymbol{y}=(y_1,y_2...y_n)'$, which is an observation vector of the independent variable y, n is the number of samples taking the regression; $\boldsymbol{X}$ is an n-by-p matrix of independent variables x_i (i=1,2...p), which may be expressed as follows:

$\boldsymbol{\beta}$ is a vector of regression parameters, $\boldsymbol{\beta} = (\beta_1, \beta_2... \beta_p)'$. ε is a vector of random

$$X = \begin{bmatrix} x_{11} & x_{12} & ..... & x_{1p} \\ x_{21} & x_{22} & ..... & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n21} & ..... & x_{np} \end{bmatrix}$$

disturbances, $\varepsilon = (\varepsilon_1, \varepsilon_2 \ldots \varepsilon_n)'$, which follows a Gaussian distribution (or nearly so).

In the linear regression, a vector $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2 \ldots \hat{\beta}_p)'$ is computed, which is an estimate of the regression parameter vector $\boldsymbol{\beta}$. The least squares estimate $\hat{\beta}$ of the vector $\boldsymbol{\beta}$ may be calculated by using Eqn. A.3 [40].

$$\hat{\beta} = (X^T X)^{-1} X^T y \qquad (A.3)$$

After the vector $\hat{\beta}$ is achieved, the predicted values of the independent variable y in Eqn. A.1 may be computed by using Eqn.A.4 [40].

$$\hat{y} = \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_p x_p \qquad (A.4)$$

where, $\hat{y}$ is the predicted values of the independent variable y.

A.2 Hypothesis Test for Individual Parameters

According to statistical theory, after the regression equation A.4 is achieved, each independent variable x_i (i=1, 2, 3...p) in Eqn. A.4 needs to be tested [40]. The hypothesis test target is to check if each x_i (i=1,2...p) has a significant effect on the $\hat{y}$. The main test steps are given as follows [40].

Hypothesis test for individual parameters β_i is done by testing the null hypothesis H_0: β_i=0 against the alternative H_1: $\beta_i \neq 0$ [40].

1. If $|t_i| > t_{\alpha/2}(n-p)$, then H_0 is rejected and H_1 is accepted at a significance level α. In this case, x_i has a significant effect on the $\hat{y}$.
2. If $|t_i| \leq t_{\alpha/2}(n-p)$, then there is not enough evidence to reject H_0 and accept H_1 at a significance level α. Thus, the H_0 is accepted. In this case, x_i has little effect (or no effect) on the $\hat{y}$.

Here, t_i is a statistical quantity defined as follows: [40]

$$t_i = \frac{\hat{\beta}_i - 0}{s_{\hat{\beta}_i}} \qquad (A.5)$$

where, β_i is the i[th] regression parameter in Eqn. A.1, while $\hat{\beta}_i$ is the estimate of β_i. $s_{\hat{\beta}_i}$ is the estimated standard error of $\hat{\beta}_i$, which is calculated by Eqn. A.6. [40]

$$s_{\hat{\beta}_i} = \sqrt{s^2 \cdot c_{ii}} \qquad (A.6)$$

where, c_{ii} is the i[th] diagonal element in matrix $(X'X)^{-1}$. s^2 is the mean square error, which is defined as follows [40].

$$s^2 = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n - p} \qquad (A.7)$$

where, y_i is the i[th] observation value of the independent variable y, while $\hat{y}_i$ is the predicted value of the y_i. n is the number of samples taking the regression. p is the number of the independent variables x_i.

$t_{\alpha/2}(n\text{-}p)$ is the 100(1-α/2) percentage point of a t distribution with n-p degrees of freedom. α is a given significance level. In this thesis, α is given to be 0.1. In this case, the confidence interval is 90% (i.e., 100(1-α) percentage).

Appendix B

B.1 Quality Primitive data calculated by the simulated metrics

Training Video Clips	Q-factor	$Qp\_f_1\_gain$	$Qp\_f_1\_loss$	$Qp\_f_2\_gain$	$Qp\_f_2\_loss$
Autumn_Leaves	Q12	0.0441	-0.2908	0.5674	-0.6252
	Q16	0.0548	-0.3448	0.7651	-0.6391
	Q24	0.0810	-0.4323	1.1445	-0.6512
	Q32	0.1155	-0.4892	1.3194	-0.6603
	Q36	0.1342	-0.5098	1.3681	-0.6601
	Q40	0.1519	-0.5277	1.4535	-0.6621
	Q48	0.1817	-0.5498	1.5349	-0.6637
Sailboat	Q12	0.0217	-0.2506	0.6581	-0.6502
	Q16	0.0324	-0.3002	0.8185	-0.6804
	Q24	0.0590	-0.3779	1.0205	-0.7196
	Q32	0.0951	-0.4345	1.1247	-0.7478
	Q36	0.1159	-0.4550	1.1638	-0.7538
	Q40	0.1357	-0.4720	1.2267	-0.7633
	Q48	0.1705	-0.5003	1.3026	-0.7741
Flower_Garden	Q12	0.0237	-0.2164	0.5786	-0.6852
	Q16	0.0332	-0.2626	0.6675	-0.7142
	Q24	0.0496	-0.3450	0.8340	-0.7438
	Q32	0.0717	-0.4131	0.9697	-0.7637
	Q36	0.0859	-0.4387	1.0362	-0.7704
	Q40	0.0973	-0.4621	1.1363	-0.7764
	Q48	0.1243	-0.4949	1.2738	-0.7876
Mobile_Celendar	Q12	0.0346	-0.2405	0.7540	-0.7257
	Q16	0.0535	-0.2842	0.8416	-0.7561
	Q24	0.1050	-0.3726	0.9356	-0.7874
	Q32	0.1617	-0.4393	1.0138	-0.8043
	Q36	0.1865	-0.4627	1.0574	-0.8103
	Q40	0.2114	-0.4831	1.1249	-0.8155
	Q48	0.2557	-0.5138	1.2286	-0.8237
Table_Tennis	Q12	0.0404	-0.3159	0.8191	-0.6259
	Q16	0.0507	-0.3816	1.1296	-0.6488
	Q24	0.0685	-0.4797	1.3878	-0.6772
	Q32	0.0932	-0.5399	1.4685	-0.6960
	Q36	0.1063	-0.5586	1.4868	-0.7016
	Q40	0.1207	-0.5758	1.5386	-0.7095
	Q48	0.1477	-0.5981	1.5739	-0.7148
Betes_Pas_Betes	Q12	0.0338	-0.2917	0.6155	-0.5872
	Q16	0.0488	-0.3503	0.7723	-0.6065
	Q24	0.0863	-0.4328	0.9888	-0.6424
	Q32	0.1272	-0.4850	1.1266	-0.6667
	Q36	0.1476	-0.5061	1.1961	-0.6681
	Q40	0.1679	-0.5223	1.2816	-0.6568
	Q48	0.2050	-0.5437	1.3870	-0.6566
Ferris_Wheel	Q12	0.0389	-0.2068	0.6125	-0.6414
	Q16	0.0559	-0.2469	0.7257	-0.6685
	Q24	0.0930	-0.3195	0.9096	-0.7024
	Q32	0.1368	-0.3772	1.0243	-0.7245
	Q36	0.1654	-0.4000	1.0820	-0.7313
	Q40	0.1802	-0.4205	1.1597	-0.7375
	Q48	0.2137	-0.4521	1.2757	-0.7451
Susie	Q12	0.0429	-0.2616	0.6634	-0.4502
	Q16	0.0546	-0.3125	0.8770	-0.4665
	Q24	0.0823	-0.3948	1.1325	-0.4862
	Q32	0.1223	-0.4491	1.2243	-0.4914
	Q36	0.1503	-0.4665	1.2613	-0.4789
	Q40	0.1686	-0.4810	1.3292	-0.4555
	Q48	0.2129	-0.4999	1.4060	-0.4419

Table B.1 Quality Primitive data calculated from training video clips with the Wolf-Pinson metric

Training Video Clips	Q-factor	$Qp\_fbf\_P_{hv}$	$Qp\_fbf\_P_{nhv}$	$Qp\_fbi\_P_{hv}$	$Qp\_fbi\_P_{nhv}$	$Qp\_tbf\_P_{hv}$	$Qp\_tbf\_P_{nhv}$	$Qp\_tbi\_P_{hv}$
Autumn_Leaves	Q12	0.6860	16.099	1.2411	12.388	0.0115	-0.0453	0.2376
	Q16	0.9231	15.417	2.5038	9.4357	0.0546	-0.0532	0.7263
	Q24	2.0004	13.916	3.1749	5.7490	0.3996	-0.0835	2.8647
	Q32	4.2438	11.836	4.4981	2.3479	1.5537	-0.1187	6.5447
	Q36	5.6013	10.182	3.9931	1.3132	2.5916	-0.1443	8.2115
	Q40	6.6561	9.4754	3.6451	1.3563	3.7781	-0.1636	10.672
	Q48	8.7580	7.4348	2.3756	1.2178	6.9144	-0.2140	12.221
Sailboat	Q12	3.1017	5.1269	2.4152	2.2587	0.0350	-0.0128	0.7151
	Q16	3.8783	4.6065	2.2863	1.6570	0.2254	-0.0140	1.5197
	Q24	5.4598	3.4664	1.5084	0.9702	0.9080	-0.0255	3.6600
	Q32	6.8842	2.5660	1.4561	0.1302	1.9305	-0.0420	5.5940
	Q36	7.4785	2.1439	0.9477	0.0735	2.6419	-0.0512	6.9300
	Q40	7.8048	2.0327	0.7599	0.1451	3.2996	-0.0570	7.9577
	Q48	8.4284	1.7243	0.3850	0.1541	4.4855	-0.0665	9.3785
Flower_Garden	Q12	2.5139	4.0576	1.5836	1.8619	0.0365	-0.0128	0.2887
	Q16	3.1373	3.7481	1.4261	1.4459	0.1099	-0.0132	0.6020
	Q24	4.3712	3.0275	1.0401	0.9243	0.3752	-0.0215	1.3812
	Q32	5.4171	2.3533	1.1394	0.2349	0.8629	-0.0328	2.5534
	Q36	6.0012	2.0398	0.8546	0.1517	1.1851	-0.0399	3.2025
	Q40	6.1694	1.9257	0.7235	0.1991	1.5734	-0.0436	4.0800
	Q48	6.7985	1.5838	0.4179	0.2014	2.4761	-0.0533	5.3233
Mobile_Celendar	Q12	2.3243	7.9372	2.7896	2.6406	0.2692	-0.0036	1.4746
	Q16	2.7885	7.4563	2.4838	1.5910	0.4041	0.0014	2.1582
	Q24	3.4772	6.9255	1.3523	0.7637	0.8618	0.0011	3.5366
	Q32	4.1287	6.6038	0.8293	0.1743	1.4313	-0.0040	4.5731
	Q36	4.2856	6.4950	0.5893	0.0786	1.8263	-0.0102	5.6692
	Q40	4.6092	6.5819	0.4957	0.1178	2.1781	-0.0107	6.6780
	Q48	5.0279	6.7586	0.2812	0.1304	2.9125	-0.0156	8.1620
Table_Tennis	Q12	2.1748	9.0921	2.4849	5.5148	0.1249	-0.0547	0.8373
	Q16	3.2157	7.4886	2.8826	3.2601	0.4570	-0.0792	2.7438
	Q24	5.1364	4.8533	1.8854	1.5591	2.5182	-0.1667	7.0499
	Q32	6.4639	3.1540	1.7437	0.3706	5.5270	-0.2596	8.5876
	Q36	6.8598	2.5673	1.2991	0.2135	6.8469	-0.3087	9.3553
	Q40	7.1446	2.3834	1.1046	0.2696	8.0682	-0.3406	9.8398
	Q48	7.6233	1.9012	0.6344	0.2746	9.9631	-0.4009	10.4429
Betes_Pas_Betes	Q12	2.7682	4.4315	1.1522	2.5213	0.1311	0.0297	1.2494
	Q16	3.1341	4.3272	1.3735	2.0139	0.4372	0.0308	2.1530
	Q24	3.9178	3.9812	1.2638	1.4742	1.2162	0.0172	4.0393
	Q32	4.8040	3.4684	1.6094	0.6340	2.2446	-0.0069	5.7619
	Q36	5.2567	3.1624	1.3902	0.4615	3.0749	-0.0317	6.7781
	Q40	5.5329	3.0517	1.3329	0.4777	3.9009	-0.0527	7.9849
	Q48	6.1702	2.6857	0.8829	0.5293	5.4307	-0.0876	8.9006
Ferris_Wheel	Q12	2.3800	5.9118	1.4325	3.4696	0.2940	-0.0066	1.3485
	Q16	2.8382	5.8178	1.6669	2.8976	0.5578	-0.0040	2.2532
	Q24	4.0311	5.2145	1.5432	2.1217	1.2658	-0.0081	4.5507
	Q32	5.3479	4.4309	2.2096	0.9266	2.1956	-0.0148	6.3794
	Q36	6.1613	3.9400	1.8899	0.6546	2.8447	-0.0235	7.8927
	Q40	6.4550	3.7462	1.7470	0.6945	3.4245	-0.0265	9.3693
	Q48	7.4887	3.1560	1.1391	0.6809	4.6934	-0.0353	11.2642
Susie	Q12	2.1908	7.1694	1.4635	4.7352	0.0984	-0.0386	0.7706
	Q16	2.8323	6.6826	1.7302	3.7773	0.2594	-0.0467	2.0427
	Q24	4.2169	5.6791	1.7845	2.5769	1.4849	-0.0869	5.5268
	Q32	5.8034	4.6136	2.4879	1.1455	3.8828	-0.1407	9.6786
	Q36	6.7692	3.9505	2.1389	0.7944	5.9556	-0.1816	11.3707
	Q40	7.2879	3.6465	1.9925	0.8153	7.7627	-0.2120	13.3760
	Q48	8.5540	2.9022	1.3504	0.7893	11.9197	-0.2841	13.2116

Table B.2 (1) Quality Primitive data calculated from training video clips with the Bishtawi Metric

Training Video Clips	Q-factor	$Qp\_tbi\_P_{nh}$ v	$Qp\_sef\_P_{hv}$	$Qp\_sef\_P_{n}$ hv	$Qp\_sei\_P_{hv}$	$Qp\_sei\_P_{n}$ hv	$Qp\_seb\_P_{h}$ v	$Qp\_seb\_P_{n}$ hv
Autumn_Leaves	Q12	-0.0793	-0.179	-0.010	2.4341	-0.0872	0.5871	-0.0466
	Q16	-0.1094	-0.105	-0.009	4.9295	-0.1614	1.3020	-0.0706
	Q24	-0.2074	0.1200	-0.013	10.341	-0.3118	3.1496	-0.1477
	Q32	-0.3187	0.5567	-0.027	12.359	-0.3804	4.3754	-0.2038
	Q36	-0.3785	0.9620	-0.040	16.257	-0.4926	6.2796	-0.2749
	Q40	-0.4358	1.4760	-0.049	17.453	-0.5369	7.3993	-0.3136
	Q48	-0.5120	2.4705	-0.055	18.334	-0.5832	8.3007	-0.3353
Sailboat	Q12	-0.0243	-0.0386	-0.000	2.8626	-0.0460	1.2080	0.0044
	Q16	-0.0348	0.0396	0.0021	3.2201	-0.0533	1.8827	0.0149
	Q24	-0.0703	0.2584	0.0038	4.1464	-0.0745	3.3619	0.0259
	Q32	-0.1068	0.8521	0.0029	4.4840	-0.0891	4.7487	0.0369
	Q36	-0.1304	1.1445	0.0000	5.9957	-0.1146	6.2979	0.0153
	Q40	-0.1451	1.3851	0.0004	6.3395	-0.1237	7.4590	0.0151
	Q48	-0.1717	1.9371	-0.001	7.3562	-0.1477	9.0286	0.0162
Flower_Garden	Q12	-0.0200	0.3196	-0.007	3.1710	-0.0485	0.8485	-0.0032
	Q16	-0.0248	0.4292	-0.006	3.1631	-0.0520	1.0582	0.0079
	Q24	-0.0468	0.7873	-0.011	4.4327	-0.0790	1.8875	0.0041
	Q32	-0.0726	0.7705	-0.011	3.9804	-0.0904	2.8478	0.0040
	Q36	-0.0889	1.1530	-0.014	4.9434	-0.1096	4.0895	-0.0065
	Q40	-0.1023	1.4276	-0.014	5.7484	-0.1255	6.5202	-0.0109
	Q48	-0.1278	1.3397	-0.017	5.3442	-0.1355	7.3410	-0.0213
Mobile_Celendar	Q12	-0.0185	0.0277	0.0073	2.1421	-0.0245	0.8418	0.0342
	Q16	-0.0215	0.2227	0.0130	3.3613	-0.0349	1.5687	0.0509
	Q24	-0.0391	0.6068	0.0187	4.8865	-0.0553	2.8651	0.0747
	Q32	-0.0606	1.1291	0.0224	5.9337	-0.0758	3.8157	0.0976
	Q36	-0.0806	1.5546	0.0191	7.8963	-0.0954	4.8574	0.0904
	Q40	-0.0928	1.8674	0.0219	8.5264	-0.1057	5.8477	0.0886
	Q48	-0.1183	2.5029	0.0237	10.576	-0.1306	7.1734	0.0982
Table_Tennis	Q12	-0.0791	-0.0500	0.0017	2.2025	-0.0444	0.9992	-0.0113
	Q16	-0.1345	0.0325	0.0056	3.5989	-0.0696	1.8912	-0.0192
	Q24	-0.2841	0.4734	0.0025	6.8736	-0.1279	3.7421	-0.0468
	Q32	-0.3938	1.1039	-0.000	8.2203	-0.1547	5.1935	-0.0537
	Q36	-0.4507	1.5844	-0.007	11.084	-0.2018	7.1058	-0.1029
	Q40	-0.4873	1.9398	-0.010	11.849	-0.2171	7.8764	-0.1148
	Q48	-0.5490	2.8663	-0.016	13.692	-0.2516	9.6044	-0.1139
Betes_Pas_Betes	Q12	-0.0316	-0.1880	0.0774	1.1798	-0.0328	0.9761	0.1311
	Q16	-0.0620	-0.0662	0.0850	1.5632	-0.0706	1.5928	0.1598
	Q24	-0.1519	0.3140	0.1069	2.9496	-0.1293	2.9034	0.1898
	Q32	-0.2357	0.9006	0.1151	3.8039	-0.1596	4.4686	0.2045
	Q36	-0.2950	1.4595	0.1132	5.1862	-0.2135	6.4442	0.1680
	Q40	-0.3449	1.8105	0.1135	5.5656	-0.2470	7.4163	0.1437
	Q48	-0.4040	2.5838	0.1149	6.8394	-0.2619	8.9478	0.1242
Ferris_Wheel	Q12	-0.0320	0.9538	-0.024	3.4743	-0.1993	2.8063	-0.0538
	Q16	-0.0432	1.1383	-0.025	3.6436	-0.2142	3.4751	-0.0565
	Q24	-0.0853	1.5677	-0.028	4.1904	-0.2610	5.6576	-0.0867
	Q32	-0.1241	2.2972	-0.041	4.2277	-0.2816	7.1652	-0.0994
	Q36	-0.1564	2.7022	-0.049	4.4701	-0.3065	9.5850	-0.1421
	Q40	-0.1787	3.1891	-0.054	5.0706	-0.3452	10.8631	-0.1625
	Q48	-0.2180	4.0875	-0.075	5.4154	-0.3827	12.8940	-0.2094
Susie	Q12	-0.0889	0.4965	-0.024	4.2024	-0.4858	2.4930	-0.1884
	Q16	-0.1363	0.8585	-0.026	4.7040	-0.5452	2.9479	-0.2032
	Q24	-0.2663	1.9250	-0.085	5.6183	-0.6531	4.5424	-0.2795
	Q32	-0.4044	2.3296	-0.107	4.8165	-0.5710	4.2595	-0.2732
	Q36	-0.4767	2.3269	-0.108	5.6508	-0.6778	5.4259	-0.4000
	Q40	-0.5356	2.5435	-0.107	6.1635	-0.7312	5.8179	-0.4192
	Q48	-0.5987	2.9219	-0.069	6.0798	-0.7314	7.6289	-0.4428

Table B.2 (2) Quality Primitive data calculated from training video clips with the Bishtawi Metric

Training Video Clips	Q-factor	$Qp\_f_1\_gain$	$Qp\_f_1\_loss$	$Qp\_f_2\_gain$	$Qp\_f_2\_loss$
Autumn_Leaves	Q12	0.0441	-0.2908	0.5674	-0.6252
	Q16	0.0548	-0.3448	0.7651	-0.6391
	Q24	0.0810	-0.4323	1.1445	-0.6512
	Q32	0.1155	-0.4892	1.3194	-0.6603
	Q36	0.1342	-0.5098	1.3681	-0.6601
	Q40	0.1519	-0.5277	1.4535	-0.6621
	Q48	0.1817	-0.5498	1.5349	-0.6637
Sailboat	Q12	0.0217	-0.2506	0.6581	-0.6502
	Q16	0.0324	-0.3002	0.8185	-0.6804
	Q24	0.0590	-0.3779	1.0205	-0.7196
	Q32	0.0951	-0.4345	1.1247	-0.7478
	Q36	0.1159	-0.4550	1.1638	-0.7538
	Q40	0.1357	-0.4720	1.2267	-0.7633
	Q48	0.1705	-0.5003	1.3026	-0.7741
Flower_Garden	Q12	0.0237	-0.2164	0.5786	-0.6852
	Q16	0.0332	-0.2626	0.6675	-0.7142
	Q24	0.0496	-0.3450	0.8340	-0.7438
	Q32	0.0717	-0.4131	0.9697	-0.7637
	Q36	0.0859	-0.4387	1.0362	-0.7704
	Q40	0.0973	-0.4621	1.1363	-0.7764
	Q48	0.1243	-0.4949	1.2738	-0.7876
Mobile_Celendar	Q12	0.0346	-0.2405	0.7540	-0.7257
	Q16	0.0535	-0.2842	0.8416	-0.7561
	Q24	0.1050	-0.3726	0.9356	-0.7874
	Q32	0.1617	-0.4393	1.0138	-0.8043
	Q36	0.1865	-0.4627	1.0574	-0.8103
	Q40	0.2114	-0.4831	1.1249	-0.8155
	Q48	0.2557	-0.5138	1.2286	-0.8237
Table_Tennis	Q12	0.0404	-0.3159	0.8191	-0.6259
	Q16	0.0507	-0.3816	1.1296	-0.6488
	Q24	0.0685	-0.4797	1.3878	-0.6772
	Q32	0.0932	-0.5399	1.4685	-0.6960
	Q36	0.1063	-0.5586	1.4868	-0.7016
	Q40	0.1207	-0.5758	1.5386	-0.7095
	Q48	0.1477	-0.5981	1.5739	-0.7148
Betes_Pas_Betes	Q12	0.0338	-0.2917	0.6155	-0.5872
	Q16	0.0488	-0.3503	0.7723	-0.6065
	Q24	0.0863	-0.4328	0.9888	-0.6424
	Q32	0.1272	-0.4850	1.1266	-0.6667
	Q36	0.1476	-0.5061	1.1961	-0.6681
	Q40	0.1679	-0.5223	1.2816	-0.6568
	Q48	0.2050	-0.5437	1.3870	-0.6566
Ferris_Wheel	Q12	0.0389	-0.2068	0.6125	-0.6414
	Q16	0.0559	-0.2469	0.7257	-0.6685
	Q24	0.0930	-0.3195	0.9096	-0.7024
	Q32	0.1368	-0.3772	1.0243	-0.7245
	Q36	0.1654	-0.4000	1.0820	-0.7313
	Q40	0.1802	-0.4205	1.1597	-0.7375
	Q48	0.2137	-0.4521	1.2757	-0.7451
Susie	Q12	0.0429	-0.2616	0.6634	-0.4502
	Q16	0.0546	-0.3125	0.8770	-0.4665
	Q24	0.0823	-0.3948	1.1325	-0.4862
	Q32	0.1223	-0.4491	1.2243	-0.4914
	Q36	0.1503	-0.4665	1.2613	-0.4789
	Q40	0.1686	-0.4810	1.3292	-0.4555
	Q48	0.2129	-0.4999	1.4060	-0.4419

Table B.3 Quality Primitive data calculated from training video clips with the EWPM

Training Video Clips	Q-factor	$Qp\_fb\_f_1\_gain$	$Qp\_fb\_f_1\_loss$	$Qp\_fb\_f_2\_gain$	$Qp\_fb\_f_2\_loss$
Autumn_Leaves	Q12	0.0070	-0.0510	0.0340	-0.0230
	Q16	0.0100	-0.0540	0.0420	-0.0190
	Q24	0.0230	-0.0530	0.0540	-0.0120
	Q32	0.0410	-0.0490	0.0720	-0.0090
	Q36	0.0520	-0.0460	0.0840	-0.0090
	Q40	0.0630	-0.0430	0.0970	-0.0100
	Q48	0.0820	-0.0380	0.1220	-0.0080
Sailboat	Q12	0.0044	-0.0461	0.0150	-0.0282
	Q16	0.0080	-0.0510	0.0210	-0.0310
	Q24	0.0240	-0.0550	0.0300	-0.0440
	Q32	0.0540	-0.0540	0.0400	-0.0760
	Q36	0.0730	-0.0520	0.0460	-0.0770
	Q40	0.0920	-0.0500	0.0540	-0.0780
	Q48	0.1300	-0.0480	0.0690	-0.0920
Flower_Garden	Q12	0.0030	-0.0260	0.0190	-0.0320
	Q16	0.0050	-0.0290	0.0270	-0.0330
	Q24	0.0130	-0.0300	0.0340	-0.0430
	Q32	0.0290	-0.0300	0.0420	-0.0600
	Q36	0.0430	-0.0290	0.0450	-0.0620
	Q40	0.0520	-0.0280	0.0490	-0.0640
	Q48	0.0750	-0.0260	0.0600	-0.0700
Mobile_Celendar	Q12	0.0093	-0.0636	0.0220	-0.0350
	Q16	0.0164	-0.0657	0.0310	-0.0554
	Q24	0.0540	-0.0841	0.0490	-0.1457
	Q32	0.1074	-0.0866	0.0670	-0.2415
	Q36	0.1331	-0.0853	0.0770	-0.2690
	Q40	0.1598	-0.0808	0.0870	-0.2978
	Q48	0.2073	-0.0753	0.1070	-0.3538
Table_Tennis	Q12	0.0080	-0.0730	0.0280	-0.0330
	Q16	0.0100	-0.0800	0.0360	-0.0300
	Q24	0.0200	-0.0840	0.0420	-0.0310
	Q32	0.0370	-0.0850	0.0500	-0.0360
	Q36	0.0470	-0.0850	0.0550	-0.0360
	Q40	0.0590	-0.0850	0.0590	-0.0420
	Q48	0.0790	-0.0850	0.0740	-0.0470
Betes_Pas_Betes	Q12	0.0070	-0.1140	0.0220	-0.1370
	Q16	0.0120	-0.1210	0.0300	-0.1320
	Q24	0.0340	-0.1210	0.0460	-0.1800
	Q32	0.0730	-0.1240	0.0610	-0.2610
	Q36	0.0950	-0.1230	0.0700	-0.2660
	Q40	0.1160	-0.1220	0.0770	-0.2350
	Q48	0.1530	-0.1150	0.0960	-0.2300
Ferris_Wheel	Q12	0.0100	-0.0330	0.0270	-0.1360
	Q16	0.0160	-0.0350	0.0380	-0.1560
	Q24	0.0410	-0.0350	0.0560	-0.1870
	Q32	0.0880	-0.0360	0.0740	-0.2450
	Q36	0.1270	-0.0340	0.0840	-0.2220
	Q40	0.1390	-0.0340	0.0930	-0.1860
	Q48	0.1760	-0.0300	0.1130	-0.1560
Susie	Q12	0.0130	-0.0730	0.0230	-0.1230
	Q16	0.0170	-0.0810	0.0300	-0.1200
	Q24	0.0390	-0.0860	0.0350	-0.1400
	Q32	0.0830	-0.0840	0.0450	-0.1840
	Q36	0.1190	-0.0800	0.0520	-0.1680
	Q40	0.1400	-0.0780	0.0600	-0.1270
	Q48	0.1890	-0.0700	0.0780	-0.0830

Table B.4 (1) Quality Primitive data calculated from training video clips with the 2CABM

Training Video Clips	Q-factor	$Qp\_nfb\_f_1\_gain$	$Qp\_nfb\_f_1\_loss$	$Qp\_nfb\_f_2\_gain$	$Qp\_nfb\_f_2\_loss$
Autumn_Leaves	Q12	0.2600	-0.2810	0.4980	-0.6250
	Q16	0.3920	-0.3390	0.6100	-0.6390
	Q24	0.5740	-0.4320	1.0380	-0.6510
	Q32	0.6100	-0.4890	1.2950	-0.6600
	Q36	0.6550	-0.5100	1.3530	-0.6600
	Q40	0.7370	-0.5280	1.4420	-0.6620
	Q48	0.8450	-0.5500	1.5270	-0.6630
Sailboat	Q12	0.3766	-0.2210	0.5520	-0.6500
	Q16	0.4670	-0.2720	0.6670	-0.6800
	Q24	0.5230	-0.3610	0.9120	-0.7190
	Q32	0.5160	-0.4260	1.0790	-0.7480
	Q36	0.5350	-0.4480	1.1350	-0.7530
	Q40	0.5890	-0.4670	1.2030	-0.7630
	Q48	0.6430	-0.4980	1.2830	-0.7740
Flower_Garden	Q12	0.1860	-0.1950	0.5330	-0.6850
	Q16	0.2390	-0.2420	0.5960	-0.7140
	Q24	0.3060	-0.3330	0.7510	-0.7440
	Q32	0.3390	-0.4070	0.9170	-0.7630
	Q36	0.3840	-0.4340	0.9940	-0.7700
	Q40	0.4390	-0.4590	1.0980	-0.7760
	Q48	0.4930	-0.4930	1.2500	-0.7870
Mobile_Celendar	Q12	0.4494	-0.2100	0.6530	-0.7250
	Q16	0.4940	-0.2580	0.7390	-0.7560
	Q24	0.4711	-0.3510	0.8760	-0.7870
	Q32	0.4216	-0.4260	0.9890	-0.8040
	Q36	0.4183	-0.4530	1.0400	-0.8100
	Q40	0.4405	-0.4760	1.1070	-0.8150
	Q48	0.4517	-0.5100	1.2150	-0.8230
Table_Tennis	Q12	0.6030	-0.3060	0.6290	-0.6240
	Q16	0.7660	-0.3780	0.9240	-0.6470
	Q24	0.8260	-0.4790	1.3400	-0.6750
	Q32	0.8130	-0.5400	1.4480	-0.6930
	Q36	0.8150	-0.5590	1.4740	-0.6990
	Q40	0.8540	-0.5760	1.5280	-0.7060
	Q48	0.8820	-0.5980	1.5660	-0.7100
Betes_Pas_Betes	Q12	0.3150	-0.2590	0.4790	-0.5620
	Q16	0.4160	-0.3240	0.5920	-0.5910
	Q24	0.5010	-0.4210	0.8390	-0.6270
	Q32	0.4900	-0.4780	1.0430	-0.6330
	Q36	0.5050	-0.5010	1.1380	-0.6330
	Q40	0.5770	-0.5180	1.2320	-0.6230
	Q48	0.6820	-0.5410	1.3530	-0.6150
Ferris_Wheel	Q12	0.3000	-0.1890	0.5310	-0.6380
	Q16	0.3850	-0.2300	0.6110	-0.6640
	Q24	0.5090	-0.3090	0.7790	-0.6990
	Q32	0.5120	-0.3710	0.9390	-0.7200
	Q36	0.5590	-0.3960	1.0140	-0.7270
	Q40	0.6700	-0.4170	1.0970	-0.7340
	Q48	0.8180	-0.4500	1.2290	-0.7410
Susie	Q12	0.5610	-0.2340	0.3840	-0.3980
	Q16	0.6970	-0.2900	0.5240	-0.4150
	Q24	0.8360	-0.3850	0.8980	-0.4270
	Q32	0.8320	-0.4430	1.1040	-0.3990
	Q36	0.8690	-0.4620	1.1620	-0.3890
	Q40	0.9540	-0.4770	1.2420	-0.3770
	Q48	1.0710	-0.4970	1.3370	-0.3750

Table B.4 (2) Quality Primitive data calculated from training video clips with the 2CABM

Training Video Clips	Q-factor	*Qp_fb_f1_gain*	*Qp_fb_f1_loss*	*Qp_fb_f2_gain*	*Qp_fb_f2_loss*	*Qp_tb_f1_gain*	*Qp_tb_f1_loss*
Autumn_Leaves	Q12	0.0070	-0.0510	0.0400	-0.2580	0.2600	-0.2900
	Q16	0.0100	-0.0540	0.0490	-0.2660	0.3920	-0.3440
	Q24	0.0230	-0.0530	0.0680	-0.2620	0.5740	-0.4320
	Q32	0.0410	-0.0490	0.0950	-0.2530	0.6100	-0.4880
	Q36	0.0520	-0.0460	0.1110	-0.2420	0.6550	-0.5080
	Q40	0.0630	-0.0430	0.1270	-0.2380	0.7370	-0.5250
	Q48	0.0820	-0.0380	0.1560	-0.2290	0.8450	-0.5470
Sailboat	Q12	0.0044	-0.0461	0.0136	-0.5322	0.3766	-0.2396
	Q16	0.0080	-0.0510	0.0200	-0.5720	0.4670	-0.2900
	Q24	0.0240	-0.0550	0.0280	-0.6240	0.5230	-0.3700
	Q32	0.0540	-0.0540	0.0380	-0.6660	0.5160	-0.4280
	Q36	0.0730	-0.0520	0.0450	-0.6700	0.5350	-0.4480
	Q40	0.0920	-0.0500	0.0540	-0.6790	0.5890	-0.4650
	Q48	0.1300	-0.0480	0.0710	-0.6930	0.6430	-0.4910
Flower_Garden	Q12	0.0030	-0.0260	0.0190	-0.4980	0.1860	-0.2090
	Q16	0.0050	-0.0290	0.0260	-0.5330	0.2390	-0.2550
	Q24	0.0130	-0.0300	0.0350	-0.5720	0.3060	-0.3390
	Q32	0.0290	-0.0300	0.0430	-0.5970	0.3390	-0.4060
	Q36	0.0430	-0.0290	0.0470	-0.6070	0.3840	-0.4310
	Q40	0.0520	-0.0280	0.0520	-0.6140	0.4390	-0.4530
	Q48	0.0750	-0.0260	0.0640	-0.6200	0.4930	-0.4850
Mobile_Celendar	Q12	0.0093	-0.0636	0.0213	-0.6117	0.4494	-0.2155
	Q16	0.0164	-0.0657	0.0289	-0.6578	0.4940	-0.2656
	Q24	0.0540	-0.0841	0.0493	-0.6979	0.4711	-0.3532
	Q32	0.1074	-0.0866	0.0678	-0.7264	0.4216	-0.4230
	Q36	0.1331	-0.0853	0.0787	-0.7332	0.4183	-0.4484
	Q40	0.1598	-0.0808	0.0899	-0.7316	0.4405	-0.4694
	Q48	0.2073	-0.0753	0.1121	-0.7424	0.4517	-0.4995
Table_Tennis	Q12	0.0080	-0.0730	0.0330	-0.4310	0.6030	-0.3150
	Q16	0.0100	-0.0800	0.0410	-0.4560	0.7660	-0.3810
	Q24	0.0200	-0.0840	0.0480	-0.4810	0.8260	-0.4780
	Q32	0.0370	-0.0850	0.0580	-0.4920	0.8130	-0.5370
	Q36	0.0470	-0.0850	0.0620	-0.4930	0.8150	-0.5550
	Q40	0.0590	-0.0850	0.0680	-0.4870	0.8540	-0.5710
	Q48	0.0790	-0.0850	0.0850	-0.4940	0.8820	-0.5930
Betes_Pas_Betes	Q12	0.0070	-0.1140	0.0200	-0.3040	0.3150	-0.2790
	Q16	0.0120	-0.1210	0.0280	-0.3290	0.4160	-0.3400
	Q24	0.0340	-0.1210	0.0450	-0.3680	0.5010	-0.4260
	Q32	0.0730	-0.1240	0.0620	-0.3850	0.4900	-0.4780
	Q36	0.0950	-0.1230	0.0720	-0.3840	0.5050	-0.4980
	Q40	0.1160	-0.1220	0.0810	-0.3820	0.5770	-0.5140
	Q48	0.1530	-0.1150	0.1020	-0.3810	0.6820	-0.5330
Ferris_Wheel	Q12	0.0100	-0.0330	0.0300	-0.3380	0.3000	-0.1980
	Q16	0.0160	-0.0350	0.0420	-0.3610	0.3850	-0.2380
	Q24	0.0410	-0.0350	0.0640	-0.3900	0.5090	-0.3100
	Q32	0.0880	-0.0360	0.0870	-0.4070	0.5120	-0.3650
	Q36	0.1270	-0.0340	0.0990	-0.4220	0.5590	-0.3870
	Q40	0.1390	-0.0340	0.1100	-0.4260	0.6700	-0.4070
	Q48	0.1760	-0.0300	0.1330	-0.4140	0.8180	-0.4360
Susie	Q12	0.0130	-0.0730	0.0330	-0.0750	0.5610	-0.2580
	Q16	0.0170	-0.0810	0.0410	-0.0870	0.6970	-0.3100
	Q24	0.0390	-0.0860	0.0530	-0.1000	0.8360	-0.3930
	Q32	0.0830	-0.0840	0.0720	-0.1000	0.8320	-0.4470
	Q36	0.1190	-0.0800	0.0860	-0.1210	0.8690	-0.4640
	Q40	0.1400	-0.0780	0.0970	-0.1280	0.9540	-0.4780
	Q48	0.1890	-0.0700	0.1250	-0.1740	1.0710	-0.4960

Table B.5 (1) Quality Primitive data calculated from training video clips with the 3CABM

Training Video Clips	Q-fator	*Qp_tb_f2 _gain*	*Qp_tb_f2 _loss*	*Qp_seb _f1_gain*	*Qp_seb_f1 _loss*	*Qp_seb_f2 _gain*	*Qp_seb_f2 _loss*
Autumn_Leaves	Q12	0.5030	-0.6030	0.1640	-0.0820	0.0030	-0.0230
	Q16	0.6700	-0.6140	0.1910	-0.1040	0.0030	-0.0190
	Q24	1.1180	-0.6200	0.2510	-0.1520	0.0040	-0.0120
	Q32	1.3120	-0.6270	0.3230	-0.1880	0.0050	-0.0090
	Q36	1.3610	-0.6270	0.3680	-0.2040	0.0060	-0.0090
	Q40	1.4470	-0.6280	0.4210	-0.2150	0.0070	-0.0100
	Q48	1.5280	-0.6270	0.5210	-0.2370	0.0100	-0.0080
Sailboat	Q12	0.5143	-0.5606	0.3634	-0.0959	0.0048	-0.0282
	Q16	0.6710	-0.5860	0.4080	-0.1220	0.0070	-0.0310
	Q24	0.9450	-0.6180	0.4920	-0.1720	0.0110	-0.0440
	Q32	1.0760	-0.6330	0.5750	-0.2180	0.0140	-0.0760
	Q36	1.1200	-0.6420	0.6220	-0.2390	0.0170	-0.0770
	Q40	1.1830	-0.6500	0.6750	-0.2580	0.0200	-0.0780
	Q48	1.2550	-0.6520	0.7730	-0.2880	0.0250	-0.0920
Flower_Garden	Q12	0.4980	-0.6400	0.3020	-0.0800	0.0050	-0.0320
	Q16	0.5740	-0.6660	0.3330	-0.0990	0.0070	-0.0330
	Q24	0.7510	-0.6930	0.3860	-0.1450	0.0090	-0.0430
	Q32	0.9090	-0.7110	0.4440	-0.1920	0.0110	-0.0600
	Q36	0.9790	-0.7170	0.4770	-0.2150	0.0110	-0.0620
	Q40	1.0810	-0.7240	0.5200	-0.2340	0.0130	-0.0640
	Q48	1.2250	-0.7350	0.6140	-0.2670	0.0150	-0.0700
Mobile_Celendar	Q12	0.6314	-0.6571	0.4581	-0.1154	0.0072	-0.0350
	Q16	0.7043	-0.6906	0.5201	-0.1433	0.0114	-0.0554
	Q24	0.8413	-0.7204	0.6163	-0.2064	0.0188	-0.1457
	Q32	0.9192	-0.7396	0.6983	-0.2612	0.0285	-0.2415
	Q36	0.9617	-0.7436	0.7300	-0.2904	0.0328	-0.2690
	Q40	1.0386	-0.7427	0.7749	-0.3068	0.0380	-0.2978
	Q48	1.1380	-0.7542	0.8570	-0.3451	0.0474	-0.3538
Table_Tennis	Q12	0.6500	-0.5550	0.3570	-0.0770	0.0040	-0.0330
	Q16	1.0500	-0.5700	0.4300	-0.0990	0.0060	-0.0300
	Q24	1.3700	-0.5840	0.5460	-0.1540	0.0090	-0.0310
	Q32	1.4540	-0.5880	0.6590	-0.1990	0.0120	-0.0360
	Q36	1.4730	-0.5900	0.6990	-0.2240	0.0130	-0.0360
	Q40	1.5250	-0.5950	0.7620	-0.2430	0.0150	-0.0420
	Q48	1.5580	-0.5690	0.8430	-0.2740	0.0200	-0.0470
Betes_Pas_Betes	Q12	0.4400	-0.4780	0.2120	-0.1090	0.0070	-0.1370
	Q16	0.5790	-0.5030	0.2520	-0.1400	0.0100	-0.1320
	Q24	0.8450	-0.5240	0.3410	-0.2020	0.0180	-0.1800
	Q32	1.0200	-0.5160	0.4470	-0.2540	0.0270	-0.2610
	Q36	1.0920	-0.5080	0.5220	-0.2750	0.0320	-0.2660
	Q40	1.1810	-0.4890	0.5930	-0.2950	0.0360	-0.2350
	Q48	1.2930	-0.4610	0.7080	-0.3270	0.0450	-0.2300
Ferris_Wheel	Q12	0.4940	-0.5970	0.2410	-0.0870	0.0040	-0.1360
	Q16	0.5840	-0.6220	0.2720	-0.1080	0.0060	-0.1560
	Q24	0.7700	-0.6520	0.3410	-0.1580	0.0080	-0.1870
	Q32	0.9260	-0.6690	0.4130	-0.2020	0.0110	-0.2450
	Q36	0.9940	-0.6730	0.4560	-0.2230	0.0120	-0.2220
	Q40	1.0780	-0.6780	0.4980	-0.2390	0.0140	-0.1860
	Q48	1.2040	-0.6860	0.5930	-0.2690	0.0170	-0.1560
Susie	Q12	0.4320	-0.3860	0.0700	-0.0380	0.0020	-0.1230
	Q16	0.6420	-0.3940	0.0800	-0.0450	0.0030	-0.1200
	Q24	1.0280	-0.3870	0.1080	-0.0750	0.0030	-0.1400
	Q32	1.1730	-0.3510	0.1340	-0.0860	0.0060	-0.1840
	Q36	1.2180	-0.3200	0.1440	-0.0950	0.0060	-0.1680
	Q40	1.2930	-0.2960	0.1610	-0.0900	0.0090	-0.1270
	Q48	1.3790	-0.2400	0.1890	-0.1320	0.0110	-0.0830

Table B.5 (2) Quality Primitive data calculated from training video clips with the 3CABM

Non-training Video Clips	Q-factor	$Qp\_f_1\_gain$	$Qp\_f_1\_loss$	$Qp\_f_2\_gain$	$Qp\_f_2\_loss$
Birches	Q12	0.0216	-0.1972	0.5867	-0.7228
	Q16	0.0301	-0.2413	0.6392	-0.7432
	Q24	0.0419	-0.3283	0.7577	-0.7651
	Q32	0.0554	-0.3950	0.9143	-0.7788
	Q36	0.0628	-0.4227	1.0038	-0.7836
	Q40	0.0716	-0.4455	1.1190	-0.7894
	Q48	0.0884	-0.4795	1.2948	-0.7993
Football	Q12	0.0710	-0.2637	0.9320	-0.6089
	Q16	0.0944	-0.3135	1.0722	-0.6334
	Q24	0.1291	-0.4001	1.2776	-0.6493
	Q32	0.1618	-0.4625	1.3471	-0.6554
	Q36	0.1788	-0.4831	1.3599	-0.6523
	Q40	0.1942	-0.5017	1.4067	-0.6560
	Q48	0.2264	-0.5252	1.4478	-0.6465
Horse_Ridings	Q12	0.0365	-0.2594	0.7188	-0.6202
	Q16	0.0467	-0.3145	0.9716	-0.6355
	Q24	0.0630	-0.4120	1.2723	-0.6457
	Q32	0.0867	-0.4859	1.3988	-0.6514
	Q36	0.0990	-0.5109	1.4379	-0.6492
	Q40	0.1145	-0.5336	1.5002	-0.6487
	Q48	0.1375	-0.5600	1.5642	-0.6556
Tempete	Q12	0.0354	-0.2324	0.6844	-0.6932
	Q16	0.0477	-0.2841	0.7914	-0.7170
	Q24	0.0761	-0.3686	0.9700	-0.7438
	Q32	0.1070	-0.4310	1.1228	-0.7603
	Q36	0.1257	-0.4551	1.1888	-0.7654
	Q40	0.1408	-0.4768	1.2642	-0.7725
	Q48	0.1736	-0.5068	1.3662	-0.7840

Table B.6 Quality Primitive data calculated from non-training video clips with the Wolf-Pinson Metric

Non-training Video Clips	Q-factor	$Qp\_fbf\_P_{hv}$	$Qp\_fbf\_P_{nhv}$	$Qp\_fbi\_P_{hv}$	$Qp\_fbi\_P_{nhv}$	$Qp\_tbf\_P_{hv}$	$Qp\_tbf\_P_{nhv}$	$Qp\_tbi\_P_{hv}$
Birches	Q12	0.7770	4.7525	0.6045	1.9271	0.0113	-0.0170	0.0556
	Q16	0.9611	4.7422	0.7673	1.3017	0.0062	-0.0175	0.1474
	Q24	1.3999	4.6947	0.6602	0.6897	0.0446	-0.0279	0.5587
	Q32	2.0540	4.5332	0.7494	0.2450	0.1794	-0.0381	1.4345
	Q36	2.2832	4.4414	0.6335	0.1190	0.3020	-0.0456	2.1768
	Q40	2.6059	4.3743	0.6009	0.1172	0.4996	-0.0499	3.2651
	Q48	3.0010	4.0560	0.3651	0.1557	1.0561	-0.0585	5.4817
Football	Q12	2.5280	14.8738	5.1879	9.7591	0.5374	-0.0319	1.8424
	Q16	3.4451	13.9651	6.2604	7.8468	1.1256	-0.0407	3.4576
	Q24	4.8670	12.3015	5.6533	5.6424	2.4651	-0.0713	6.7055
	Q32	6.5294	10.4642	5.9068	3.4759	3.9988	-0.1058	8.9715
	Q36	7.3182	9.4515	5.3033	2.7100	5.0242	-0.1316	10.4095
	Q40	7.9364	8.9752	5.1889	2.4946	6.0926	-0.1497	11.7295
	Q48	9.3429	7.8085	3.9086	2.4343	8.1269	-0.1909	12.1846
Horse_Ridings	Q12	2.5491	4.6458	0.9855	2.9395	0.1516	-0.0439	0.6843
	Q16	2.9355	4.2123	1.3419	1.9943	0.3809	-0.0557	1.8611
	Q24	3.8355	3.2757	1.1781	1.1056	1.4208	-0.0996	5.4225
	Q32	4.8233	2.3424	1.3071	0.2784	3.1936	-0.1507	9.3250
	Q36	5.1825	1.8775	0.9908	0.1371	4.5582	-0.1863	12.1723
	Q40	5.5107	1.6660	0.8040	0.1936	5.8979	-0.2147	14.5726
	Q48	5.9685	1.2160	0.4250	0.1834	8.5615	-0.2708	16.5533
Tempete	Q12	2.3660	4.4303	1.2511	1.6592	0.0965	-0.0128	0.6161
	Q16	2.7843	4.1541	1.2630	1.0889	0.1957	-0.0118	1.3051
	Q24	3.4530	3.6894	0.8198	0.6431	0.6293	-0.0203	3.3462
	Q32	4.1915	3.2792	0.8455	0.2102	1.4390	-0.0326	6.0861
	Q36	4.5436	3.0157	0.6424	0.1270	2.1798	-0.0428	8.3459
	Q40	4.7912	2.9379	0.5845	0.1273	2.8587	-0.0485	10.4720
	Q48	5.4625	2.6556	0.3550	0.1269	4.6261	-0.0654	13.7804

Table B.7 (1) Quality Primitive data calculated from non-training video clips with the Bishtawi Metric

Non-training Video Clips	Q-factor	$Qp\_tbi\_P_{n}$ $_{hv}$	$Qp\_sef\_P_{hv}$	$Qp\_sef\_P_{n}$ $_{hv}$	$Qp\_sei\_P_{hv}$	$Qp\_sei\_P_{n}$ $_{hv}$	$Qp\_seb\_P_{h}$ $_{v}$	$Qp\_seb\_P_{n}$ $_{hv}$
Birches	Q12	-0.0261	0.0267	-0.007	0.5693	-0.0206	0.2044	0.0004
	Q16	-0.0301	0.0472	-0.006	1.2111	-0.0281	0.3432	0.0103
	Q24	-0.0546	0.1091	-0.011	2.8050	-0.0517	1.0307	0.0106
	Q32	-0.0825	0.2747	-0.015	3.5204	-0.0757	2.1043	0.0068
	Q36	-0.1010	0.3554	-0.019	5.4066	-0.0929	3.1485	-0.0034
	Q40	-0.1169	0.4576	-0.019	6.4053	-0.1056	4.2467	-0.0106
	Q48	-0.1497	0.7052	-0.020	8.9668	-0.1305	6.4329	-0.0200
Football	Q12	-0.0998	-0.199	0.0060	3.2598	-0.1707	0.8117	-0.0825
	Q16	-0.1429	0.0080	0.0040	5.4860	-0.2595	1.8842	-0.1190
	Q24	-0.2392	0.5272	0.0010	9.3470	-0.3794	3.6800	-0.2110
	Q32	-0.3125	1.1164	-0.005	9.8804	-0.3923	4.5151	-0.2326
	Q36	-0.3659	1.2482	-0.008	11.6751	-0.4492	5.8695	-0.2960
	Q40	-0.4069	1.7583	-0.012	13.4070	-0.5043	7.2821	-0.3264
	Q48	-0.4593	2.3387	-0.009	14.2125	-0.5353	9.0618	-0.3663
Horse_Ridings	Q12	-0.0679	1.2304	-0.054	6.3389	-0.4208	2.5590	-0.0788
	Q16	-0.1014	1.7891	-0.066	6.4827	-0.4377	3.2644	-0.0879
	Q24	-0.2039	2.1467	-0.086	6.3912	-0.4540	4.8668	-0.1483
	Q32	-0.3107	2.9889	-0.121	5.9201	-0.4554	6.0348	-0.1816
	Q36	-0.3815	2.9126	-0.126	7.0003	-0.5375	8.1895	-0.2455
	Q40	-0.4408	3.7650	-0.156	7.2106	-0.5682	9.9449	-0.2920
	Q48	-0.5212	4.2248	-0.164	7.0873	-0.5838	13.3573	-0.3662
Tempete	Q12	-0.0263	1.5858	-0.027	5.2149	-0.1255	1.7953	-0.0342
	Q16	-0.0348	1.9760	-0.026	6.5722	-0.1579	3.1322	-0.0400
	Q24	-0.0707	2.6736	-0.036	8.2078	-0.2092	5.1674	-0.0784
	Q32	-0.1140	3.7489	-0.049	9.2948	-0.2522	7.2006	-0.1059
	Q36	-0.1459	4.3895	-0.060	10.7019	-0.2904	9.5788	-0.1574
	Q40	-0.1708	5.3094	-0.073	12.4270	-0.3327	11.5283	-0.1846
	Q48	-0.2164	6.6395	-0.084	13.9424	-0.3822	15.0506	-0.2359

Table B.7 (2) Quality Primitive data calculated from non-training video clips with the Bishtawi Metric

Non-training Video Clips	Q-factor	$Qp\_f_1\_gain$	$Qp\_f_1\_loss$	$Qp\_f_2\_gain$	$Qp\_f_2\_loss$
Birches	Q12	0.0216	-0.1972	0.5867	-0.7228
	Q16	0.0301	-0.2413	0.6392	-0.7432
	Q24	0.0419	-0.3283	0.7577	-0.7651
	Q32	0.0554	-0.3950	0.9143	-0.7788
	Q36	0.0628	-0.4227	1.0038	-0.7836
	Q40	0.0716	-0.4455	1.1190	-0.7894
	Q48	0.0884	-0.4795	1.2948	-0.7993
Football	Q12	0.0710	-0.2637	0.9320	-0.6089
	Q16	0.0944	-0.3135	1.0722	-0.6334
	Q24	0.1291	-0.4001	1.2776	-0.6493
	Q32	0.1618	-0.4625	1.3471	-0.6554
	Q36	0.1788	-0.4831	1.3599	-0.6523
	Q40	0.1942	-0.5017	1.4067	-0.6560
	Q48	0.2264	-0.5252	1.4478	-0.6465
Horse_Ridings	Q12	0.0365	-0.2594	0.7188	-0.6202
	Q16	0.0467	-0.3145	0.9716	-0.6355
	Q24	0.0630	-0.4120	1.2723	-0.6457
	Q32	0.0867	-0.4859	1.3988	-0.6514
	Q36	0.0990	-0.5109	1.4379	-0.6492
	Q40	0.1145	-0.5336	1.5002	-0.6487
	Q48	0.1375	-0.5600	1.5642	-0.6556
Tempete	Q12	0.0354	-0.2324	0.6844	-0.6932
	Q16	0.0477	-0.2841	0.7914	-0.7170
	Q24	0.0761	-0.3686	0.9700	-0.7438
	Q32	0.1070	-0.4310	1.1228	-0.7603
	Q36	0.1257	-0.4551	1.1888	-0.7654
	Q40	0.1408	-0.4768	1.2642	-0.7725
	Q48	0.1736	-0.5068	1.3662	-0.7840

Table B.8 Quality Primitive data calculated from non-training video clips with the EWPM

Non-training Video Clips	Q-factor	$Qp\_fb\_f_1\_gain$	$Qp\_fb\_f_1\_loss$	$Qp\_fb\_f_2\_gain$	$Qp\_fb\_f_2\_loss$
Birches	Q12	0.0010	-0.0030	0.0200	-0.0040
	Q16	0.0020	-0.0030	0.0270	-0.0060
	Q24	0.0040	-0.0020	0.0360	-0.0090
	Q32	0.0080	-0.0020	0.0440	-0.0140
	Q36	0.0100	-0.0020	0.0490	-0.0150
	Q40	0.0130	-0.0020	0.0560	-0.0160
	Q48	0.0170	-0.0020	0.0690	-0.0190
Football	Q12	0.0440	-0.1170	0.0390	-0.2140
	Q16	0.0610	-0.1300	0.0520	-0.2130
	Q24	0.0840	-0.1510	0.0720	-0.1850
	Q32	0.1090	-0.1620	0.0910	-0.1770
	Q36	0.1250	-0.1630	0.0990	-0.1660
	Q40	0.1390	-0.1640	0.1090	-0.1460
	Q48	0.1740	-0.1650	0.1270	-0.1120
Horse_Ridings	Q12	0.0061	-0.0468	0.0290	-0.0550
	Q16	0.0086	-0.0520	0.0370	-0.0545
	Q24	0.0199	-0.0550	0.0430	-0.0393
	Q32	0.0395	-0.0554	0.0500	-0.0386
	Q36	0.0503	-0.0540	0.0540	-0.0324
	Q40	0.0635	-0.0528	0.0600	-0.0291
	Q48	0.0816	-0.0515	0.0740	-0.0257
Tempete	Q12	0.0053	-0.0182	0.0280	-0.0211
	Q16	0.0094	-0.0193	0.0360	-0.0282
	Q24	0.0242	-0.0186	0.0490	-0.0484
	Q32	0.0439	-0.0182	0.0630	-0.0732
	Q36	0.0546	-0.0174	0.0750	-0.0816
	Q40	0.0698	-0.0165	0.0810	-0.0912
	Q48	0.1015	-0.0149	0.1020	-0.1032

Table B.9 (1) Quality Primitive data calculated from non-training video clips with the 2CABM

Non-training Video Clips	Q-factor	$Qp\_nfb\_f_1\_gain$	$Qp\_nfb\_f_1\_loss$	$Qp\_nfb\_f_2\_gain$	$Qp\_nfb\_f_2\_loss$
Birches	Q12	0.0220	-0.1940	0.5840	-0.7230
	Q16	0.0270	-0.2380	0.6330	-0.7430
	Q24	0.0330	-0.3270	0.7470	-0.7650
	Q32	0.0330	-0.3950	0.9060	-0.7790
	Q36	0.0350	-0.4220	0.9970	-0.7840
	Q40	0.0400	-0.4450	1.1130	-0.7890
	Q48	0.0460	-0.4790	1.2920	-0.7990
Football	Q12	0.7500	-0.2330	0.6760	-0.5830
	Q16	0.8440	-0.2860	0.8250	-0.6080
	Q24	0.9190	-0.3820	1.1250	-0.6260
	Q32	0.8850	-0.4510	1.2510	-0.6310
	Q36	0.8670	-0.4730	1.2820	-0.6290
	Q40	0.9380	-0.4930	1.3340	-0.6340
	Q48	1.0000	-0.5180	1.3870	-0.6270
Horse_Ridings	Q12	0.4286	-0.2470	0.5890	-0.6180
	Q16	0.5632	-0.3060	0.7860	-0.6340
	Q24	0.6705	-0.4100	1.2120	-0.6450
	Q32	0.6687	-0.4860	1.3840	-0.6500
	Q36	0.6848	-0.5110	1.4300	-0.6480
	Q40	0.7451	-0.5340	1.4940	-0.6470
	Q48	0.7873	-0.5600	1.5600	-0.6540
Tempete	Q12	0.2362	-0.2210	0.6390	-0.6930
	Q16	0.2733	-0.2750	0.7360	-0.7170
	Q24	0.2923	-0.3650	0.9330	-0.7440
	Q32	0.2865	-0.4300	1.1070	-0.7600
	Q36	0.3015	-0.4540	1.1800	-0.7650
	Q40	0.3322	-0.4760	1.2560	-0.7720
	Q48	0.3689	-0.5070	1.3610	-0.7840

Table B.9 (2) Quality Primitive data calculated from non-training video clips with the 2CABM

Non-training Video Clips	Q-fator	*Qp_fb_f1_gain*	*Qp_fb_f1_loss*	*Qp_fb_f2_gain*	*Qp_fb_f2_loss*	*Qp_tb_f1_gain*	*Qp_tb_f1_loss*
Birches	Q12	0.0010	-0.0030	0.0200	-0.4930	0.0220	-0.1940
	Q16	0.0020	-0.0030	0.0270	-0.5220	0.0270	-0.2380
	Q24	0.0040	-0.0020	0.0360	-0.5500	0.0330	-0.3230
	Q32	0.0080	-0.0020	0.0460	-0.5660	0.0330	-0.3890
	Q36	0.0100	-0.0020	0.0510	-0.5700	0.0350	-0.4160
	Q40	0.0130	-0.0020	0.0580	-0.5770	0.0400	-0.4380
	Q48	0.0170	-0.0020	0.0720	-0.5840	0.0460	-0.4710
Football	Q12	0.0440	-0.1170	0.0500	-0.2390	0.7500	-0.2500
	Q16	0.0610	-0.1300	0.0660	-0.2490	0.8440	-0.3010
	Q24	0.0840	-0.1510	0.0930	-0.2580	0.9190	-0.3910
	Q32	0.1090	-0.1620	0.1170	-0.2560	0.8850	-0.4560
	Q36	0.1250	-0.1630	0.1280	-0.2520	0.8670	-0.4760
	Q40	0.1390	-0.1640	0.1400	-0.2600	0.9380	-0.4950
	Q48	0.1740	-0.1650	0.1620	-0.2470	1.0000	-0.5180
Horse_Ridings	Q12	0.0061	-0.0468	0.0217	-0.2188	0.4286	-0.2572
	Q16	0.0086	-0.0520	0.0302	-0.2271	0.5632	-0.3129
	Q24	0.0199	-0.0550	0.0379	-0.2286	0.6705	-0.4106
	Q32	0.0395	-0.0554	0.0388	-0.2334	0.6687	-0.4838
	Q36	0.0503	-0.0540	0.0847	-0.2329	0.6848	-0.5083
	Q40	0.0635	-0.0528	0.0931	-0.2257	0.7451	-0.5305
	Q48	0.0816	-0.0515	0.1105	-0.2593	0.7873	-0.5558
Tempete	Q12	0.0053	-0.0182	0.0237	-0.4688	0.2362	-0.2272
	Q16	0.0094	-0.0193	0.0330	-0.4940	0.2733	-0.2785
	Q24	0.0242	-0.0186	0.0498	-0.5205	0.2923	-0.3618
	Q32	0.0439	-0.0182	0.0668	-0.5300	0.2865	-0.4227
	Q36	0.0546	-0.0174	0.0762	-0.5301	0.3015	-0.4456
	Q40	0.0698	-0.0165	0.0857	-0.5389	0.3322	-0.4664
	Q48	0.1015	-0.0149	0.1045	-0.5448	0.3689	-0.4946

Table B.10 (1) Quality Primitive data calculated from non-training video clips with the 3CABM

Non-training Video Clips	Q-factor	*Qp_tb_f2 _gain*	*Qp_tb_f2 _loss*	*Qp_seb _f1_gain*	*Qp_seb_f1 _loss*	*Qp_seb_f2 _gain*	*Qp_seb_f2 _loss*
Birches	Q12	0.5550	-0.7020	0.3220	-0.0920	0.0040	-0.0040
	Q16	0.6030	-0.7210	0.3560	-0.1150	0.0050	-0.0060
	Q24	0.7160	-0.7420	0.4130	-0.1690	0.0060	-0.0090
	Q32	0.8670	-0.7550	0.4740	-0.2190	0.0070	-0.0140
	Q36	0.9550	-0.7600	0.5100	-0.2430	0.0070	-0.0150
	Q40	1.0740	-0.7660	0.5500	-0.2600	0.0090	-0.0160
	Q48	1.2620	-0.7760	0.6350	-0.2940	0.0100	-0.0190
Football	Q12	0.7380	-0.5570	0.1840	-0.0560	0.0030	-0.2140
	Q16	0.9100	-0.5790	0.2120	-0.0730	0.0040	-0.2130
	Q24	1.1930	-0.5880	0.2710	-0.1090	0.0050	-0.1850
	Q32	1.2890	-0.5880	0.3310	-0.1350	0.0060	-0.1770
	Q36	1.3100	-0.5830	0.3770	-0.1490	0.0070	-0.1660
	Q40	1.3590	-0.5820	0.4150	-0.1610	0.0090	-0.1460
	Q48	1.4050	-0.5740	0.4800	-0.1840	0.0100	-0.1120
Horse_Ridings	Q12	0.6127	-0.5929	0.1545	-0.0808	0.0027	-0.0550
	Q16	0.8578	-0.6065	0.1759	-0.1006	0.0038	-0.0545
	Q24	1.2475	-0.6142	0.2373	-0.1493	0.0041	-0.0393
	Q32	1.3893	-0.6132	0.3106	-0.1976	0.0047	-0.0386
	Q36	1.4284	-0.6068	0.3607	-0.2254	0.0054	-0.0324
	Q40	1.4904	-0.6043	0.4159	-0.2484	0.0054	-0.0291
	Q48	1.5517	-0.5915	0.5477	-0.2823	0.0072	-0.0257
Tempete	Q12	0.5816	-0.6592	0.4071	-0.1104	0.0112	-0.0211
	Q16	0.6851	-0.6818	0.4563	-0.1395	0.0128	-0.0282
	Q24	0.8871	-0.7082	0.5538	-0.2015	0.0154	-0.0484
	Q32	1.0603	-0.7247	0.6624	-0.2542	0.0193	-0.0732
	Q36	1.1290	-0.7293	0.7313	-0.2778	0.0256	-0.0816
	Q40	1.2078	-0.7363	0.7963	-0.2968	0.0263	-0.0912
	Q48	1.3093	-0.7461	0.9268	-0.3279	0.0362	-0.1032

Table B.10 (2) Quality Primitive data calculated from non-training video clips with the 3CABM

Training Video Clips	Q-factor	$Qp\_fb\_f_1\_gain$	$Qp\_fb\_f_2\_gain$	$Qp\_tb\_f_1\_loss$	$Qp\_tb\_f_2\_loss$	$Qp\_seb\_f_1\_gain$	$Qp\_seb\_f_2\_loss$	Qp_bf
Autumn_Leaves	Q12	0.0070	0.0400	-0.2900	-0.6030	0.1640	-0.0230	0.0017
	Q16	0.0100	0.0490	-0.3440	-0.6140	0.1910	-0.0190	0.0428
	Q24	0.0230	0.0680	-0.4320	-0.6200	0.2510	-0.0120	0.3006
	Q32	0.0410	0.0950	-0.4880	-0.6270	0.3230	-0.0090	0.6437
	Q36	0.0520	0.1110	-0.5080	-0.6270	0.3680	-0.0090	0.9022
	Q40	0.0630	0.1270	-0.5250	-0.6280	0.4210	-0.0100	0.9517
	Q48	0.0820	0.1560	-0.5470	-0.6270	0.5210	-0.0080	0.9622
Sailboat	Q12	0.0044	0.0136	-0.2396	-0.5606	0.3634	-0.0282	0.0383
	Q16	0.0080	0.0200	-0.2900	-0.5860	0.4080	-0.0310	0.0619
	Q24	0.0240	0.0280	-0.3700	-0.6180	0.4920	-0.0440	0.0887
	Q32	0.0540	0.0380	-0.4280	-0.6330	0.5750	-0.0760	0.0948
	Q36	0.0730	0.0450	-0.4480	-0.6420	0.6220	-0.0770	0.1165
	Q40	0.0920	0.0540	-0.4650	-0.6500	0.6750	-0.0780	0.1254
	Q48	0.1300	0.0710	-0.4910	-0.6520	0.7730	-0.0920	0.1376
Flower_Garden	Q12	0.0030	0.0190	-0.2090	-0.6400	0.3020	-0.0320	0.4087
	Q16	0.0050	0.0260	-0.2550	-0.6660	0.3330	-0.0330	0.4415
	Q24	0.0130	0.0350	-0.3390	-0.6930	0.3860	-0.0430	0.4900
	Q32	0.0290	0.0430	-0.4060	-0.7110	0.4440	-0.0600	0.4950
	Q36	0.0430	0.0470	-0.4310	-0.7170	0.4770	-0.0620	0.5311
	Q40	0.0520	0.0520	-0.4530	-0.7240	0.5200	-0.0640	0.5493
	Q48	0.0750	0.0640	-0.4850	-0.7350	0.6140	-0.0700	0.5893
Mobile_Celendar	Q12	0.0093	0.0213	-0.2155	-0.6571	0.4581	-0.0350	0.3065
	Q16	0.0164	0.0289	-0.2656	-0.6906	0.5201	-0.0554	0.4609
	Q24	0.0540	0.0493	-0.3532	-0.7204	0.6163	-0.1457	0.5707
	Q32	0.1074	0.0678	-0.4230	-0.7396	0.6983	-0.2415	0.6165
	Q36	0.1331	0.0787	-0.4484	-0.7436	0.7300	-0.2690	0.7050
	Q40	0.1598	0.0899	-0.4694	-0.7427	0.7749	-0.2978	0.7519
	Q48	0.2073	0.1121	-0.4995	-0.7542	0.8570	-0.3538	0.8126
Table_Tennis	Q12	0.0080	0.0330	-0.3150	-0.5550	0.3570	-0.0330	0.2691
	Q16	0.0100	0.0410	-0.3810	-0.5700	0.4300	-0.0300	0.5426
	Q24	0.0200	0.0480	-0.4780	-0.5840	0.5460	-0.0310	0.8456
	Q32	0.0370	0.0580	-0.5370	-0.5880	0.6590	-0.0360	0.9020
	Q36	0.0470	0.0620	-0.5550	-0.5900	0.6990	-0.0360	0.9600
	Q40	0.0590	0.0680	-0.5710	-0.5950	0.7620	-0.0420	0.9722
	Q48	0.0790	0.0850	-0.5930	-0.5690	0.8430	-0.0470	0.9756
Betes_Pas_Betes	Q12	0.0070	0.0200	-0.2790	-0.4780	0.2120	-0.1370	0.9004
	Q16	0.0120	0.0280	-0.3400	-0.5030	0.2520	-0.1320	0.9594
	Q24	0.0340	0.0450	-0.4260	-0.5240	0.3410	-0.1800	0.9694
	Q32	0.0730	0.0620	-0.4780	-0.5160	0.4470	-0.2610	0.9800
	Q36	0.0950	0.0720	-0.4980	-0.5080	0.5220	-0.2660	0.9861
	Q40	0.1160	0.0810	-0.5140	-0.4890	0.5930	-0.2350	0.9915
	Q48	0.1530	0.1020	-0.5330	-0.4610	0.7080	-0.2300	0.9922
Ferris_Wheel	Q12	0.0100	0.0300	-0.1980	-0.5970	0.2410	-0.1360	0.5180
	Q16	0.0160	0.0420	-0.2380	-0.6220	0.2720	-0.1560	0.6496
	Q24	0.0410	0.0640	-0.3100	-0.6520	0.3410	-0.1870	0.7365
	Q32	0.0880	0.0870	-0.3650	-0.6690	0.4130	-0.2450	0.7852
	Q36	0.1270	0.0990	-0.3870	-0.6730	0.4560	-0.2220	0.8702
	Q40	0.1390	0.1100	-0.4070	-0.6780	0.4980	-0.1860	0.9026
	Q48	0.1760	0.1330	-0.4360	-0.6860	0.5930	-0.1560	0.9302
Susie	Q12	0.0130	0.0330	-0.2580	-0.3860	0.0700	-0.1230	0.4743
	Q16	0.0170	0.0410	-0.3100	-0.3940	0.0800	-0.1200	0.7337
	Q24	0.0390	0.0530	-0.3930	-0.3870	0.1080	-0.1400	0.8624
	Q32	0.0830	0.0720	-0.4470	-0.3510	0.1340	-0.1840	0.8652
	Q36	0.1190	0.0860	-0.4640	-0.3200	0.1440	-0.1680	0.8709
	Q40	0.1400	0.0970	-0.4780	-0.2960	0.1610	-0.1270	0.8737
	Q48	0.1890	0.1250	-0.4960	-0.2400	0.1890	-0.0830	0.8737

Table B.11 Quality Primitive data calculated from training video clips with the AMAM

Non-training Video Clips	Q-factor	*Qp_fb_f1_gain*	*Qp_fb_f2_gain*	*Qp_tb_f1_loss*	*Qp_tb_f2_loss*	*Qp_seb_f1_gain*	*Qp_seb_f2_loss*	*Qp_bf*
Birches	Q12	0.0010	0.0200	-0.1940	-0.7020	0.3220	-0.0040	0.0002
	Q16	0.0020	0.0270	-0.2380	-0.7210	0.3560	-0.0060	0.0009
	Q24	0.0040	0.0360	-0.3230	-0.7420	0.4130	-0.0090	0.0067
	Q32	0.0080	0.0460	-0.3890	-0.7550	0.4740	-0.0140	0.0389
	Q36	0.0100	0.0510	-0.4160	-0.7600	0.5100	-0.0150	0.0904
	Q40	0.0130	0.0580	-0.4380	-0.7660	0.5500	-0.0160	0.1552
	Q48	0.0170	0.0720	-0.4710	-0.7760	0.6350	-0.0190	0.2876
Football	Q12	0.0440	0.0500	-0.2500	-0.5570	0.1840	-0.2140	0.2470
	Q16	0.0610	0.0660	-0.3010	-0.5790	0.2120	-0.2130	0.8420
	Q24	0.0840	0.0930	-0.3910	-0.5880	0.2710	-0.1850	0.9661
	Q32	0.1090	0.1170	-0.4560	-0.5880	0.3310	-0.1770	0.9506
	Q36	0.1250	0.1280	-0.4760	-0.5830	0.3770	-0.1660	0.9606
	Q40	0.1390	0.1400	-0.4950	-0.5820	0.4150	-0.1460	0.9667
	Q48	0.1740	0.1620	-0.5180	-0.5740	0.4800	-0.1120	0.9667
Horse_Ridings	Q12	0.0061	0.0217	-0.2572	-0.5929	0.1545	-0.0550	0.3713
	Q16	0.0086	0.0302	-0.3129	-0.6065	0.1759	-0.0545	0.4524
	Q24	0.0199	0.0379	-0.4106	-0.6142	0.2373	-0.0393	0.5800
	Q32	0.0395	0.0388	-0.4838	-0.6132	0.3106	-0.0386	0.6769
	Q36	0.0503	0.0847	-0.5083	-0.6068	0.3607	-0.0324	0.8556
	Q40	0.0635	0.0931	-0.5305	-0.6043	0.4159	-0.0291	0.9019
	Q48	0.0816	0.1105	-0.5558	-0.5915	0.5477	-0.0257	0.9196
Tempete	Q12	0.0053	0.0237	-0.2272	-0.6592	0.4071	-0.0211	0.2937
	Q16	0.0094	0.0330	-0.2785	-0.6818	0.4563	-0.0282	0.3335
	Q24	0.0242	0.0498	-0.3618	-0.7082	0.5538	-0.0484	0.3678
	Q32	0.0439	0.0668	-0.4227	-0.7247	0.6624	-0.0732	0.3907
	Q36	0.0546	0.0762	-0.4456	-0.7293	0.7313	-0.0816	0.4867
	Q40	0.0698	0.0857	-0.4664	-0.7363	0.7963	-0.0912	0.5539
	Q48	0.1015	0.1045	-0.4946	-0.7461	0.9268	-0.1032	0.6130

Table B.12 Quality Primitive data calculated from non-training video clips with the AMAM

Video Clips	Q-Factor	Wolf-Pinson Metric	Video Clips	Q-Factor	Wolf-Pinson Metric
Autumn_Leaves	Q12	-0.2238	Table_Tennis	Q12	-0.2927
	Q16	-0.2856		Q16	-0.385
	Q24	-0.3959		Q24	-0.4748
	Q32	-0.4443		Q32	-0.5066
	Q36	-0.456		Q36	-0.5128
	Q40	-0.4756		Q40	-0.5256
	Q48	-0.4911		Q48	-0.5317
Sailboat	Q12	-0.2381	Betes_Pas_Betes	Q12	-0.2392
	Q16	-0.2897		Q16	-0.2917
	Q24	-0.3553		Q24	-0.3582
	Q32	-0.3867		Q32	-0.3938
	Q36	-0.3954		Q36	-0.4099
	Q40	-0.4086		Q40	-0.4278
	Q48	-0.4232		Q48	-0.4457
Flower_Garden	Q12	-0.2061	Ferris_Wheel	Q12	-0.2043
	Q16	-0.2404		Q16	-0.2389
	Q24	-0.3036		Q24	-0.2943
	Q32	-0.352		Q32	-0.3256
	Q36	-0.3715		Q36	-0.3364
	Q40	-0.399		Q40	-0.3562
	Q48	-0.4326		Q48	-0.3819
Mobile_Celendar	Q12	-0.2513	Susie	Q12	-0.2353
	Q16	-0.2807		Q16	-0.2992
	Q24	-0.3158		Q24	-0.3784
	Q32	-0.337		Q32	-0.4045
	Q36	-0.3463		Q36	-0.4088
	Q40	-0.3599		Q40	-0.4227
	Q48	-0.3781		Q48	-0.4303

Table B.13 (1) Automatic quality scores calculated from test video clips with the Wolf-Pinson Metric

Video Clips	Q-Factor	Wolf-Pinson Metric	Video Clips	Q-Factor	Wolf-Pinson Metric
Birches	Q12	-0.2015	Horse_Ridings	Q12	-0.2497
	Q16	-0.227		Q16	-0.3248
	Q24	-0.2827		Q24	-0.4246
	Q32	-0.3388		Q32	-0.4726
	Q36	-0.367		Q36	-0.4862
	Q40	-0.3987		Q40	-0.5032
	Q48	-0.4455		Q48	-0.5189
Football	Q12	-0.2869	Tempete	Q12	-0.2319
	Q16	-0.3289		Q16	-0.2714
	Q24	-0.3955		Q24	-0.3335
	Q32	-0.4225		Q32	-0.3803
	Q36	-0.4266		Q36	-0.3973
	Q40	-0.4384		Q40	-0.417
	Q48	-0.4443		Q48	-0.4391

Table B.13 (2) Automatic quality scores calculated from test video clips with the Wolf-Pinson Metric

B.2 Quality Primitive combinations of the simulated metrics

B.2.1 A Quality Primitive combination of the Bishtawi Metric

The statistical data of linear regression analysis of the Bishtawi Metric is listed in Table B.14, Table B.15, Table B.16 and Table B.17. The data suggests that the Quality Primitives $Qp\_sef\_P_{hv}$, $Qp\_sef\_P_{nhv}$, $Qp\_sei\_P_{hv}$ and $Qp\_seb\_P_{hv}$ in the regression equation B.1 appear to have little effect on the automatic quality scores under 90% confidence interval. In the regression equation B.2 obtained by using the linear regression between the non-automatic quality data and the selected Quality Primitive data (i.e., $Qp\_fbf\_P_{hv}$, $Qp\_fbf\_P_{nhv}$, $Qp\_fbi\_P_{hv}$, $Qp\_fbi\_P_{nhv}$, $Qp\_tbf\_P_{hv}$, $Qp\_tbf\_P_{nhv}$, $Qp\_tbi\_P_{hv}$, $Qp\_tbi\_P_{nhv}$, $Qp\_sei\_P_{nhv}$ and $Qp\_seb\_P_{nhv}$), all ten Quality Primitives have significant effect on the automatic quality scores. Thus, the regression equation B.2 is taken as the Quality Primitive combination of the Bishtawi Metric.

Regression equation (B.1)		
Automatic quality score (i.e. Predicted value of non-automatic quality score)	=	$3.2856 \times Qp\_fbf\_P_{hv} + 3.4992 \times Qp\_fbf\_P_{nhv} - 5.8026 \times Qp\_fbi\_P_{hv} - 4.3801 \times Qp\_fbi\_P_{nhv} + 3.2568 \times Qp\_tbf\_P_{hv} + 247.8803 \times Qp\_tbf\_P_{nhv} - 4.6604 \times Qp\_tbi\_P_{hv} - 211.486 \times Qp\_tbi\_P_{nhv} - 0.9694 \times Qp\_sef\_P_{hv} - 6.9254 \times Qp\_sef\_P_{nhv} - 0.1571 \times Qp\_sei\_P_{hv} + 48.0015 \times Qp\_sei\_P_{nhv} + 2.0165 \times Qp\_seb\_P_{hv} - 81.2959 \times Qp\_seb\_P_{nhv}$

Table B.14 Regression equation achieved with the linear regression between the non-automatic quality data and the fourteen Quality Primitive data in Table B.2

Quality Primitives (i.e., independent variables)	t_i value	$t_{\alpha/2}(n-m)$ $\alpha=0.1$ n=56 m=14	Test if each Quality Primitive in the regression equation B.1 has a significant effect on automatic quality scores (i.e., predicted values of non-automatic quality scores) under 90% confidence interval.
$Qp\_fbf\_P_{hv}$	2.7362	1.302	Because 2.7362> 1.302, $Qp\_fbf\_P_{hv}$ has a significant effect.
$Qp\_fbf\_P_{nhv}$	2.8336	1.302	Because 2.8336> 1.302, $Qp\_fbf\_P_{hv}$ has a significant effect.
$Qp\_fbi\_P_{hv}$	2.5924	1.302	Because 2.5924>1.302, $Qp\_fbi\_P_{hv}$ has a significant effect .
$Qp\_fbi\_P_{nhv}$	2.8502	1.302	Because 2.8502>1.302, $Qp\_fbi\_P_{nhv}$ has a significant effect.
$Qp\_tbf\_P_{hv}$	1.551	1.302	Because 1.551> 1.302, $Qp\_tbf\_P_{hv}$ has a significant effect.
$Qp\_tbf\_P_{nhv}$	4.753	1.302	Because 4.753> 1.302, $Qp\_tbf\_P_{nhv}$ has a significant effect.
$Qp\_tbi\_P_{hv}$	3.2069	1.302	Because 3.2069> 1.302, $Qp\_tbi\_P_{hv}$ has a significant effect.
$Qp\_tbi\_P_{nhv}$	5.0936	1.302	Because 5.0936> 1.302, $Qp\_tbi\_P_{nhv}$ has a significant effect.
$Qp\_sef\_P_{hv}$	0.2126	1.302	Because 0.2126< 1.302, $Qp\_sef\_P_{hv}$ has little effect.
$Qp\_sef\_P_{nhv}$	0.0574	1.302	Because 0.0574<1.302, $Qp\_sef\_P_{nhv}$ has little effect.
$Qp\_sei\_P_{hv}$	0.2086	1.302	Because 0.2086< 1.302, $Qp\_sei\_P_{hv}$ has little effect.
$Qp\_sei\_P_{nhv}$	2.4276	1.302	Because 2.4276> 1.302, $Qp\_sei\_P_{nhv}$ has a significant effect.
$Qp\_seb\_P_{hv}$	1.300	1.302	Because 1.300< 1.302, $Qp\_seb\_P_{hv}$ has little effect.
$Qp\_seb\_P_{nhv}$	1.5613	1.302	Because 1.5613> 1.302, $Qp\_seb\_P_{nhv}$ has a significant effect.

Table B.15 Test of each Quality Primitive in the regression equation B.1

In Table B.15:

- The calculation method of t_i value is described in Appendx A.2.
- n is the number of training video clips.
- m is the number of independent variables in the regression equation B.1.
- Here, α is significant level, which equals to 0.1, while $100(1-\alpha)\%$ (i.e. 90%) is the confidence interval.
- $t_{\alpha/2}(n-m)$ is the $100(1-\alpha/2)$ percentage point of a t distribution with n-m of freedom.

Regression equation (B.2)		
Automatic quality score (i.e. Predicted value of non-automatic quality score)	=	$3.8357\times Qp\_fbf\_P_{hv} + 3.1525\times Qp\_fbf\_P_{nhv} - 5.7492\times Qp\_fbi\_P_{hv} - 3.6886\times Qp\_fbi\_P_{nhv} + 3.2032\times Qp\_tbf\_P_{hv} + 241.7725\times Qp\_tbf\_P_{nhv} - 3.0847\times Qp\_tbi\_P_{hv} - 190.3876 \times Qp\_tbi\_P_{nhv} + 44.3769\times Qp\_sei\_P_{nhv} - 76.6538\times Qp\_seb\_P_{nhv}$

Table B.16 Regression equation achieved with the linear regression between the non-automatic quality data and the selected Quality Primitive data (i.e., $Qp\_fbf\_P_{hv}$, $Qp\_fbf\_P_{nhv}$, $Qp\_fbi\_P_{hv}$, $Qp\_fbi\_P_{nhv}$, $Qp\_tbf\_P_{hv}$, $Qp\_tbf\_P_{nhv}$, $Qp\_tbi\_P_{hv}$, $Qp\_tbi\_P_{nhv}$, $Qp\_sei\_P_{nhv}$ and $Qp\_seb\_P_{nhv}$) in Table B.2

Quality Primitives (i.e., independent variables)	t_i value	$t_{\alpha/2}(n-m)$ $\alpha=0.1$ n=56 m=10	Test if each Quality Primitive in the regression equation B.2 has a significant effect on automatic quality scores (i.e., predicted values of non-automatic quality scores) under 90% confidence interval.
$Qp\_fbf\_P_{hv}$	6.4133	1.302	Because 6.4133> 1.302, $Qp\_fbf\_P_{hv}$ has a significant effect.
$Qp\_fbf\_P_{nhv}$	6.1768	1.302	Because 6.1768> 1.302, $Qp\_fbf\_P_{hv}$ has a significant effect.
$Qp\_fbi\_P_{hv}$	4.8095	1.302	Because 4.8095>1.302, $Qp\_fbi\_P_{hv}$ has a significant effect .
$Qp\_fbi\_P_{nhv}$	4.4204	1.302	Because 4.4204>1.302, $Qp\_fbi\_P_{nhv}$ has a significant effect.
$Qp\_tbf\_P_{hv}$	4.3957	1.302	Because 4.3957> 1.302, $Qp\_tbf\_P_{hv}$ has a significant effect.
$Qp\_tbf\_P_{nhv}$	4.2077	1.302	Because 4.2077> 1.302, $Qp\_tbf\_P_{nhv}$ has a significant effect.
$Qp\_tbi\_P_{hv}$	3.2503	1.302	Because 3.2503> 1.302, $Qp\_tbi\_P_{hv}$ has a significant effect.
$Qp\_tbi\_P_{nhv}$	2.8848	1.302	Because 2.8848> 1.302, $Qp\_tbi\_P_{nhv}$ has a significant effect.
$Qp\_sei\_P_{nhv}$	2.8844	1.302	Because 2.8844> 1.302, $Qp\_sei\_P_{nhv}$ has a significant effect.
$Qp\_seb\_P_{nhv}$	2.4802	1.302	Be cause 2.4802> 1.302, $Qp\_seb\_P_{nhv}$ has a significant effect.

Table B.17 Test of each Quality Primitive in the regression equation B.2

In Table B.17:

- The calculation method of t_i value is described in Appendx A.2.
- n is the number of training video clips.
- m is the number of independent variables in the regression equation

B.2.

- Here, α is significant level, which equals to 0.1, while $100(1-\alpha)\%$ (i.e. 90%) is the confidence interval.
- $t_{\alpha/2}(n-m)$ is the $100(1-\alpha/2)$ percentage point of a t distribution with n-m of freedom.

B.2.2 Quality Primitive combination of the EWPM

The statistical data of linear regression analysis of the EWPM is listed in Table B.18, Table B.19, Table B.20 and Table B.21. The data suggests that the Quality Primitive *Qp_f1_loss* in the regression equation B.3 appears to have little effect on the automatic quality score under 90% confidence interval. In the regression equation B.4 obtained with the linear regression between the non-automatic quality data and the selected Quality Primitive data (i.e., *Qp_ f1_gain, Qp_ f2_gain* and *Qp_ f2_loss*), all three Quality Primitives have significant effect on the automatic quality score. Thus, the regression equation B.4 is taken as the Quality Primitive combination of the EWPM.

Regression equation (B.3)		
Automatic quality score (i.e. Predicted value of non-automatic quality score)	=	156.5047× *Qp_f1_gain* - 16.6081× *Qp_f1_loss* + 24.8069× *Qp_f2_gain* + 39.8434× *Qp_f2_loss*

Table B.18 Regression equation achieved with the linear regression between the non-automatic quality data and the four Quality Primitive data in Table B.3

Quality Primitives (i.e., independent variables)	t_i value	$t_{\alpha/2}(n-m)$ $\alpha=0.1$ n=56 m=4	Test if each Quality Primitive in the regression equation B.3 has a significant effect on automatic quality scores (i.e., predicted values of non-automatic quality scores) under 90% confidence interval.
Qp_f1_gain	5.4862	1.300	Because 5.4862> 1.300, *Qp_f1 _gain* has a significant effect.
Qp_f1_loss	0.4141	1.300	Because 0.4141< 1.300, *Qp_f1_loss* has little effect.
Qp_f2_gain	1.8387	1.300	Because 1.8387> 1.300, *Qp_f2_gain* has a significant effect .
Qp_f2_loss	5.6696	1.300	Because 5.6696> 1.300, *Qp_f2_loss* has a significant effect.

Table B.19 Test of each Quality Primitive in the regression equation B.3

In Table B.19:

- The calculation method of t_i value is described in Appendx A.2.
- n is the number of training video clips.
- m is the number of independent variables in the regression equation B.3.
- Here, α is significant level, which equals to 0.1, while $100(1-\alpha)\%$ (i.e. 90%) is the confidence interval.
- $t_{\alpha/2}$ (n-m) is the $100(1-\alpha/2)$ percentage point of a t distribution with n-m of freedom.

Regression equation (B.4)		
Automatic quality score	=	160.4383× *Qp_f1_gain* + 29.9804× *Qp_f2_gain*
(i.e. Predicted value of	+	38.5314× *Qp_f2_loss*
non-automatic quality score)		

Table B.20 Regression equation achieved with the linear regression between the non-automatic quality data and the selected Quality Primitive data (i.e., *Qp_f1_gain, Qp_f2_gain* and *Qp_f2_loss*) in Table B.3

Quality Primitives (i.e., independent variables)	t_i value	$t_{\alpha/2}(n-m)$ $\alpha=0.1$ n=56 m=3	Test if each Quality Primitive in the regression equation B.4 has a significant effect on automatic quality scores (i.e., predicted values of non-automatic quality scores) under 90% confidence interval.
Qp_f1_gain	6.0116	1.300	Because 6.0116> 1.300, *Qp_f1_gain* has a significant effect.
Qp_f2_gain	5.9297	1.300	Because 5.9297> 1.300, *Qp_f2_gain* has a significant effect .
Qp_f2_loss	6.1910	1.300	Because 6.1910> 1.300, *Qp_f2_loss* has a significant effect.

Table B.21 Test of each Quality Primitive in the regression equation B.4

In Table B.21:

- the calculation method of t_i value is described in Appendx A.2.
- n is the number of training video clips.
- m is the number of independent variables in the regression equation B.4.
- Here, α is significant level, which equals to 0.1, while $100(1-\alpha)\%$ (i.e. 90%) is the confidence interval.
- $t_{\alpha/2}(n-m)$ is the $100(1-\alpha/2)$ percentage point of a t distribution with n-m of freedom.

B.2.3 Quality Primitive combination of the 2CABM

The statistical data of linear regression analysis of the 2CABM is listed in Table B.22, Table B.23, Table B.24 and Table B.25. The data suggests that the Quality Primitives *Qp_fb_f1_loss, Qp_fb_f2_loss, Qp_nfb_f1_gain and Qp_nfb_f2_gain* in the regression equation B.5 appear to have little effect on the automatic quality score under 90% confidence interval. In the regression equation B.6 obtained with the linear regression between the non-automatic quality data and the selected Quality Primitive data (i.e., *Qp_fb_f1_gain, Qp_fb_f2_gain, Qp_nfb_f1_loss* and *Qp_nfb_f2_loss*), all the four Quality Primitives have significant effect on the automatic quality score. Thus, the

regression equation B.6 is taken as the Quality Primitive combination used to calculate the automatic quality score in the 2CABM.

Regression equation (B.5)		
Automatic quality score (i.e. Predicted value of non-automatic quality score)	=	91.3599 × *Qp_fb_f1_gain* + 43.9055 × *Qp_fb_f1_loss* + 239.345× *Qp_fb_f2_gain* + 25.2107 × *Qp_fb_f2_loss* + 5.6326 × *Qp_nfb_f1_gain* - 70.7073 × *Qp_nfb_f1_loss* + 0.7407 × *Qp_nfb_f2_gain* + 33.6584 × *Qp_nfb_f2_loss*

Table B.22 Regression equation achieved with the linear regression between the non-automatic quality data and eight Quality Primitive data in Table B.4

Quality Primitives (i.e., independent variables)	t_i value	$t_{\alpha/2}(n-m)$ $\alpha=0.1$ n=56 m=8	Test if each Quality Primitive in regression equation B.5 has a significant effect on automatic quality scores (i.e., predicted values of non-automatic quality scores) under 90% confidence interval.
Qp_fb_f1_gain	1.8977	1.300	Because 1.8977 > 1.300, *Qp_fb_f1_gain* has a significant effect.
Qp_fb_f1_loss	0.7218	1.300	Because 0.7218 < 1.300, *Qp_fb_f1_loss* has little effect.
Qp_fb_f2_gain	2.5932	1.300	Because 2.5932 > 1.300, *Qp_fb_f2_gain* has a significant effect .
Qp_fb_f2_loss	1.0012	1.300	Because 1.0012<1.300, *Qp_fb_f2_loss* has little effect.
Qp_nfb_f1_gain	0.6844	1.300	Because 0.6844 < 1.300, *Qp_nfb_f1_gain* has little effect.
Qp_nfb_f1_loss	1.7726	1.300	Because 1.7726 > 1.300, *Qp_nfb_f1_loss* has a significant effect.
Qp_nfb_f2_gain	0.0493	1.300	Because 0.0493 < 1.300, *Qp_nfb_f2_gain* has little effect.
Qp_nfb_f2_loss	4.9048	1.300	Because 4.9048 > 1.300, *Qp_nfb_f2_loss* has a significant effect.

Table B.23 Test of each Quality Primitive in the regression equation B.5

In Table B.23:

- The calculation method of t_i value is described in Appendx A.2.
- n is the number of training video clips.
- m is the number of independent variables in the regression equation

B.5.

- Here, α is significant level, which equals to 0.1, while 100(1- α)% (i.e. 90%) is the confidence interval.
- $t_{\alpha/2}$ (n-m) is the 100(1- α /2) percentage point of a t distribution with n-m of freedom.

Regression equation (B.6)		
Automatic quality score	=	47.2359 × *Qp_fb_f1_gain* + 284.3618 ×
(i.e. Predicted value of		*Qp_fb_f2_gain* − 77.1673 × *Qp_nfb_f1_loss* +
non-automatic quality score)		39.5749× *Qp_nfb_f2_loss*

Table B.24 Regression equation achieved with linear regression between non-automatic quality data and selected Quality Primitive data (i.e., *Qp_fb_f1_gain, Qp_fb_f2_gain, Qp_nfb_f1_loss* and *Qp_nfb_f2_loss*) in Table B.4

Primitive Name (i.e., independent variables)	t_i value	$t_{\alpha/2}$(n-m) α =0.1 n=56 m=5	Test if each Quality Primitive in the regression equation B.6 has a significant effect on automatic quality scores (i.e., predicted values of non-automatic quality scores) under 90% confidence interval.
Qp_fb_f1_gain	1.3082	1.30	Because 2.5583 > 1.30, *Qp_fb_f1_gain* has a significant effect.
Qp_fb_f2_gain	3.2456	1.30	Because 3.0896 > 1.30, *Qp_fb_f2_gain* has a significant effect.
Qp_nfb_f1_loss	5.9897	1.30	Because 5.6511 > 1.30, *Qp_nfb_f1_loss* has a significant effect.
Qp_nfb_f2_loss	6.5745	1.30	Because 5.1993>1.30, *Qp_nfb_f2_loss* has a significant effect.

Table B.25 Test of each Quality Primitive in regression equation B.6

In Table B.25:

- The calculation method of t_i value is described in Appendx A.2.
- n is the number of training video clips.
- m is the number of independent variables in the regression equation B.6.
- Here, α is significant level, which equals to 0.1, while 100(1- α)% (i.e. 90%) is the confidence interval.
- $t_{\alpha/2}$ (n-m) is the 100(1- α /2) percentage point of a t distribution with n-m of freedom.

B.2.4 Quality Primitive combination of the 3CABM

The statistical data of linear regression analysis of the 3CABM is listed in Table B.26, Table B.27, Table B.28 and Table B.29. The data suggests that the Quality Primitives *Qp_fb_f1_loss*, *Qp_fb_f2_loss*, *Qp_tb_f1_gain*, *Qp_tb_f2_gain*, *Qp_seb_f2_gain* and *Qp_seb_f1_loss* in the regression equation B.7 appear to have little effect on the automatic quality score under 90% confidence interval. In the regression equation B.8 obtained with the linear regression between the non-automatic quality data and the selected Quality Primitive data (i.e., *Qp_fb_f1_gain*, *Qp_fb_f2_gain*, *Qp_tb_f1_loss*, *Qp_tb_f2_loss*, *Qp_seb_f1_gain* and *Qp_seb_f2_loss*), all the six Quality Primitives have significant effect on the automatic quality score. Thus, the regression equation B.8 is taken as the Quality Primitive combination of the 3CABM.

Regression equation (B.7)		
Automatic quality score (i.e. Predicted value of non-automatic quality score)	=	88.2665 × *Qp_fb_f1_gain* + 2.4862 × *Qp_fb_f1_loss* + 299.089 × *Qp_fb_f2_gain* - 20.8795 × *Qp_fb_f2_loss* –16.6582× *Qp_tb_f1_gain* - 59.957 × *Qp_tb_f1_loss* +12.9192× *Qp_tb_f2_gain* + 32.792× *Qp_tb_f2_loss* - 37.0685 × *Qp_seb_f1_gain* - 5.5101 × *Qp_seb_f1_loss* + 88.7143× *Qp_seb_f2_gain* + 34.1985 × *Qp_seb_f2_loss*

Table B.26 Regression equation achieved with the linear regression between the non-automatic quality data and the twelve Quality Primitive data in Table B.5

Quality Primitive (i.e., independent variables)	t_i value	$t_{\alpha/2}(n-m)$ $\alpha=0.1$ n=56 m=12	Test if each Quality Primitive in the regression equation B.7 has a significant effect on automatic quality scores (i.e., predicted values of non-automatic quality scores) under 90% confidence interval.
Qp_fb_f1_gain	1.3718	1.306	Because 1.3718 > 1.306, *Qp_fb_f1_gain* has a significant effect.
Qp_fb_f1_loss	0.0308	1.306	Because 0.0308 < 1.306, *Qp_fb_f1_loss* has little effect.
Qp_fb_f2_gain	2.3093	1.306	Because2.3093 > 1.306, *Qp_fb_f2_gain* has a significant effect.
Qp_fb_f2_loss	0.8773	1.306	Because 0.8773 < 1.306, *Qp_fb_f2_loss* has little effect.
Qp_tb_f1_gain	1.0317	1.306	Because 1.0317 < 1.306, *Qp_tb_f1_gain* no has effect.
Qp_tb_f1_loss	1.3643	1.306	Because 1.4643 > 1.306, *Qp_tb_f1_loss* has a significant effect.
Qp_tb_f2_gain	0.7255	1.306	Because 0.7255 < 1.306, *Qp_tb_f2_gain* has little effect.
Qp_tb_f2_loss	1.4844	1.306	Because 1.4844 > 1.306, *Qp_tb_f2_loss* has a significant effect.
Qp_seb_f1_gain	1.8642	1.306	Because 1.8642 >1.306, *Qp_seb_f2_gain* has a significant effect.
Qp_seb_f1_loss	0.0981	1.306	Because 0.0981< 1.306, *Qp_seb_f1_loss* has little effect.
Qp_seb_f2_gain	0.2269	1.306	Because 0.2269 < 1.306, *Qp_seb_f1_gain* has little effect.
Qp_seb_f2_loss	1.3245	1.306	Because1.3245 > 1.306, *Qp_seb_f2_loss* has a significant effect.

Table B.27 Test of each Quality Primitive in regression equation B.7

In Table B.27:

- The calculation method of t_i value is described in Appendx A.2.
- n is the number of training video clips.
- m is the number of independent variables in the regression equation B.7.
- Here, α is significant level, which equals to 0.1, while $100(1-\alpha)\%$ (i.e. 90%) is the confidence interval.
- $t_{\alpha/2}(n-m)$ is the $100(1-\alpha/2)$ percentage point of a t distribution with n-m of freedom.

Regression equation (B.8)	
Automatic quality score	= 134.4915 × *Qp_fb_f1_gain* + 224.3640 ×
(i.e. Predicted value of	*Qp_fb_f2_gain* - 59.4794 × *Qp_tb_f1_loss* +
non-automatic quality score)	19.1785 × *Qp_tb_f2_loss* − 15.8416×
	Qp_seb_f1_gain + 41.3030× *Qp_seb_f2_loss*

Table B.28 Regression equation achieved with the linear regression between the non-automatic quality data and the selected Quality Primitive data (i.e., *Qp_fb_f1_gain, Qp_fb_f2_gain, Qp_tb_f1_loss, Qp_tb_f2_loss, Qp_seb_f1_gain* and *Qp_seb_f2_loss*) in Table B.5

Quality Primitive (i.e., independent variables)	t_i value	$t_{\alpha/2}(n-m)$ $\alpha=0.1$ n=56 m=6	Test if each Quality Primitive in regression equation B.8 has a significant effect on automatic quality scores (i.e., predicted values of non-automatic quality scores) under 90% confidence interval.
Qp_fb_f1_gain	3.0596	1.30	Because 3.0596 > 1.30, *Qp_fb_f1_gain* has a significant effect.
Qp_fb_f2_gain	3.7052	1.30	Because 3.7052 >1.30, *Qp_tb_f1_gain* has a significant effect.
Qp_tb_f1_loss	5.3740	1.30	Because 5.3740 >1.30, *Qp_tb_f1_loss* has a significant effect.
Qp_tb_f2_loss	2.7918	1.30	Because 2.7918 >1.30, *Qp_tb_f2_loss* has a significant effect.
Qp_seb_f1_gain	2.1112	1.30	Because 2.1112 >1.30, *Qp_seb_f2_gain* has a significant effect.
Qp_seb_f2_loss	2.7936	1.30	Because 2.7936 >1.30, *Qp_seb_f2_loss* has a significant effect.

Table B.29 Test of each Quality Primitive in the regression equation B.8

In Table B.29:

- The calculation method of t_i value is described in Appendx A.2.
- n is the number of training video clips.
- m is the number of independent variables in the regression equation B.8.
- Here, α is significant level, which equals to 0.1, while $100(1-\alpha)\%$ (i.e. 90%) is the confidence interval.
- $t_{\alpha/2}(n-m)$ is the $100(1-\alpha/2)$ percentage point of a t distribution with n-m of freedom.

B.2.5 Quality Primitive combination of the AMAM

The data listed in Table B.30 and Table B.31 suggests that each Quality Primitive in

the regression equation B.9 has a significant effect on the automatic quality scores under 90% confidence interval. Thus, the regression equation B.9 is taken as a Quality Primitive combination of the AMAM.

Regression equation B.9		
Automatic quality score (i.e. Predicted value of non-automatic quality score)	=	196.1270 × *Qp_fb_f1_gain* + 138.1984× *Qp_fb_f2_gain* - 46.4135 × *Qp_tb_f1_loss* + 11.9942 × *Qp_tb_f2_loss* - 17.8626× *Qp_seb_f1_gain* + 69.8896 × *Qp_seb_f2_loss* + 10.6713×*Qp_bf*

Table B.30 Regression Equation B.9 achieved with linear regression between non-automatic data and thirteen Quality Primitive data in Table B.9

Quality Primitive (i.e., regression variables)	t_i value	$t_{\alpha/2}$ (n-m) α =0.1 n=56 m=7	Test if each Quality Primitive has a significant effect on automatic quality scores (i.e., predicted values of non-automatic quality scores) under 90% confidence interval.
Qp_fb_f1_gain	3.8476	1.306	Because 3.8476 > 1.306, *Qp_fb_f1_gain* has a significant effect.
Qp_fb_f2_gain	1.9590	1.306	Because 1.9590 > 1.306, *Qp_fb_f2_gain* has a significant effect.
Qp_tb_f1_loss	3.7902	1.306	Because 3.7902> 1.306, *Qp_tb_f1_loss* has a significant effect.
Qp_tb_f2_loss	1.6205	1.306	Because 1.6205> 1.306, *Qp_tb_f2_loss* has a significant effect.
Qp_seb_f1_gain	2.4481	1.306	Because 2.4481 > 1.306, *Qp_seb_f2_gain* has a significant effect.
Qp_seb_f2_loss	3.6061	1.306	Because 3.6061> 1.306, *Qp_seb_f2_loss* has a significant effect.
Qp_bf	2.1781	1.306	Because 2.1781> 1.306, *Qp_bf* has a significant effect.

Table B.31 Test of each Quality Primitive in the regression equation B.9

In Table B.31:

- Calculation method of the t_i value is described in Appendx A.2.
- n is the number of training video clips.
- m is the number of regression variables in the regression equation B.9.
- Here, α is significant level, which equals to 0.1, while 100(1- α)% (i.e. 90%) is the confidence interval.
- $t_{\alpha/2}$ (n-m) is the 100(1- α /2) percentage point of a t distribution with n-m of freedom.