

# NOTE TO USERS

This reproduction is the best copy available.

**UMI<sup>®</sup>**



# **Automatic Measurement of Video Quality**

## **using Fuzzy Logic**

Zhen Cai

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  
Concordia University  
Montreal, Quebec, Canada

October 2004

© Concordia University, 2004



Library and  
Archives Canada

Bibliothèque et  
Archives Canada

Published Heritage  
Branch

Direction du  
Patrimoine de l'édition

395 Wellington Street  
Ottawa ON K1A 0N4  
Canada

395, rue Wellington  
Ottawa ON K1A 0N4  
Canada

*Your file    Votre référence*

*ISBN: 0-494-04362-8*

*Our file    Notre référence*

*ISBN: 0-494-04362-8*

#### NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

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

#### AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

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

---

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

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

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

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

  
**Canada**

## **ABSTRACT**

### **Automatic Measurement of Video Quality using Fuzzy Logic**

Zhen Cai

Compression of digital video systems introduces artifacts (i.e., typical types of degradations), such as Blocking, Blurring, and Ringing. In this thesis, a new temporal artifact-‘Flashing Blocks’ is introduced. This new temporal artifact with other spatial artifacts is integrated into an equation to simulate the goodness of subjective ratings. In the past, linear mapping was used for this mapping. However, in this thesis, a non-linear mapping using Fuzzy Logic is proposed to create the mapping.

In this thesis, the performance of these automatic metrics is evaluated by measuring the Mean Square Error between the measures predicted by humans and that predicted automatically. The simulation results indicate that the performance of the automatic metric using Fuzzy Logic is significantly better than those automatic metrics using linear mapping for MPEG-2 test video sequences.

Keywords: digital video system, automatic metric, Human Visual System (HVS), Block Flashing, Fuzzy Logic.

## **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

<b>LIST OF FIGURES .....</b>	<b>Viii</b>
<b>LIST OF TABLES .....</b>	<b>xii</b>
<b>LIST OF ABBREVIATIONS AND SYMBOLS .....</b>	<b>xiv</b>
<b>Chapter 1 Introduction .....</b>	<b>1</b>
1.1 Video Compression and Quality Measurement .....	1
1.2 Weakness of MSE.....	3
1.3 Problem Statement .....	4
1.4 Thesis Summary.....	4
<b>Chapter 2 The Background of the Quality Measurement Metric .....</b>	<b>6</b>
2.1 Introduction .....	6
2.2 Subjective Measurement of Video Quality.....	7
2.3 HVS Based Method of Automatically Measuring Video Quality-Watson's Method .....	10
2.4 Quality Primitives .....	12
2.4.1 Wolf's method .....	13
2.4.2 the Bistawhi's method .....	18
2.5 Introductions to Fuzzy Logic .....	23
2.6 Conclusions .....	32
<b>Chapter 3 Fuzzy Metrics with Spatial Quality Primitives .....</b>	<b>34</b>
3.1 Introduction .....	34
3.2 Fuzzy Model using 4 QP's .....	35
3.2.1 Proposed Fuzzy Metric.....	36

3.2.1.1 The Premise Part of the Fuzzy Metric.....	38
3.2.1.2 The Consequent Part of the Fuzzy Metric.....	40
3.2.2 The Tuning Method .....	40
3.2.3 Example of Implementing the Fuzzy Metric .....	44
3.3 Fuzzy Model using Fuzzy Clustering Method (8 QP's) .....	47
3.3.1 Proposed Fuzzy Metric .....	47
3.3.1.1 The Premise Part of the Fuzzy Metric .....	49
3.3.1.2 The Consequent Part of the Fuzzy Metric .....	52
3.3.2 The Tuning Method .....	52
3.3.3 Example of Implementing the Fuzzy Metric .....	54
3.4 Fuzzy Model using 12 QP's .....	58
3.4.1 Proposed Fuzzy Metric .....	58
3.4.1.1 The Premise Part of the Fuzzy Algorithm .....	60
3.4.1.2 The Consequent Part of the Fuzzy Algorithm .....	61
3.4.2 The Tuning Method of the Hybrid algorithm .....	62
3.5 Simulation Result .....	64
3.6 Conclusions .....	77
<b>Chapter 4 Flashing Block Artifact .....</b>	<b>78</b>
4.1 Introduction .....	78
4.2 The Definition of the 'Flashing Blocks' artifact .....	79
4.3 The Method to find the 'Flashing Blocks' .....	86
4.3.1 The Temporal Method .....	87
4.3.2 The Spatial Frequency Method .....	88

4.3.3 The Over Dark and the Over Bright Method .....	89
4.3.4 Selection of the Temporal Threshold (TI).....	89
4.4 Method to compute the‘Flahsing Blocks’QP’s .....	93
4.5 Fuzzy Model using Temporal Artifact as the Input .....	94
4.5.1 Fuzzy Model with 4+1 QP’s .....	95
4.5.2 Fuzzy Model with 8+1 QP’s.....	96
4.5.3 Fuzzy Model with 12+1 QP’s .....	98
4.6 Simulation Result of Fuzzy Algorithms .....	100
4.7 Conclusions .....	105
<b>5. Conclusions and Further Research .....</b>	<b>106</b>
5.1 Contributions .....	106
5.2 Conclusions .....	106
5.3 Future Research .....	107
<b>BIBLIOGRAPHY .....</b>	<b>108</b>

## LIST OF FIGURES

Figure 2.1 Overview of the Objective Quality Metric .....	10
Figure 2.2 Wolf's method to compute the Quality Primitives .....	13
Figure 2.3 Sobel edge enhance filters.....	14
Figure 2.4 Classification of HV $\overline{HV}$ .....	15
Figure 2.5 An original video clip.....	19
Figure 2.6 Flat Blocks .....	19
Figure 2.7 Texture Blocks .....	20
Figure 2.8 Sharp Edge Blocks .....	20
Figure 2.9 the Approximation Ability of the Fuzzy Logic .....	32
Figure 3.1 Figure of the Fuzzy Algorithm .....	37
Figure 3.2 Premise Part of the Fuzzy System .....	39
Figure 3.3 Membership Function of $x_1$ .....	46
Figure 3.4 Membership Function of $x_2$ .....	46
Figure 3.5 Membership Function of $x_3$ .....	46
Figure 3.6 Membership Function of $x_4$ .....	46
Figure 3.7 Figure of the Fuzzy Metric(3.7) ..	48
Figure 3.8 Premise Part of the Fuzzy System .....	51
Figure 3.9 Membership Function of $x_1$ .....	56
Figure 3.10 Membership Function of $x_2$ .....	56
Figure 3.11 Membership Function of $x_3$ .....	56
Figure 3.12 Membership Function of $x_4$ .....	56

Figure 3.13 Membership Function of x5 .....	57
Figure 3.14 Membership Function of x6.....	57
Figure 3.15 Membership Function of x7.....	57
Figure 3.16 Membership Function of x6.....	57
Figure 3.17 Figure of the Hybrid Algorithm .....	58
Figure 3.18 the Premise Part of the Fuzzy Algorithm(3.18) .....	61
Figure 3.19 Scatter Plot of LMA versus Subjective Ratings (4QP's).....	68
Figure 3.20 Scatter Plot of the 4QPFA versus Subjective Ratings (4 QP's) .....	68
Figure 3.21 Scatter Plot of LMA versus Subjective Ratings(8 QP's) .....	72
Figure 3.22 Scatter Plot of the 8QPFA( 8 QP's) versus Subjective Ratings .....	72
Figure 3.23 Scatter Plot of the LMA versus Subjective Ratings (12 QP's) .....	76
Figure 3.24 Scatter Plot of the 12QPFA versus Subjective Ratings (12 QP's).....	76
Figure 4.1 Spatial Locations of 'Flashing Blocks' and 'non Flashing Blocks' .....	79
Figure 4.2 the Temporal Plot of the 'Flashing Blocks' (Spatial Domain).....	80
Figure 4.3 the Temporal Plot of 'Flashing Blocks' (DCT Domain) .....	81
Figure 4.4 the Temporal Plot the 'non Flashing Blocks' (Spatial flat Domain) .....	82
Figure 4.5 the Temporal Plot of the 'non Flashing Blocks'(DCT Domain) .....	83
Figure 4.6 the Temporal Plot of the 'non flashing block' (Spatial non-flat Domain).....	84
Figure.4.7 the Temporal Plot of the 'non Flashing Blocks' (DCT Domain) .....	85
Figure 4.8 'Flashing Blocks' Detection Method .....	86
Figure 4.9 Spatial Locations of the 'Flashing Blocks' found by eyes .....	91
Figure 4.10 the relationship curve between the False Alarm Rate and the Miss rate .....	93

Figure 4.11 the Procedure for generating the QP (T) for the ‘Flashing Blocks’ artifact	94
Figure 4.12 Figure of the Fuzzy Algorithm (4.9)	96
Figure 4.13 Figure of the Fuzzy Metric	97
Figure 4.14 Figure of the Hybrid Algorithm	99
Figure 4.15 Scatter Plot of the Output of the 4QPFAT versus Subjective Rating Values	101
Figure 4.16 Scatter Plot of the Output of the 4QPFA versus Subjective Rating Values	101
Figure 4.17 Scatter Plot of the Output of the 8QPFAT versus Subjective Rating Values	103
Figure 4.18 Scatter Plot of the Output of the 8QPFA versus Subjective Rating Values	103
Figure 4.19 Scatter Plot of the Output of the 12QPFAT versus Subjective Rating Values	104
Figure 4.20 Scatter Plot of the Output of the 12QPFA versus Subjective Rating Values	105

## LIST OF TABLES

Table 2.1 Rating Scale for Subjective Evaluations .....	8
Table 2.2 MPEG-2 Encoded at 7.5 Mbps and pure I-frame with constant Q-factor .....	9
Table 2.3 Blocks Classes and Their Classification Conditions .....	21
Table 2.4 8 Quality Primitives .....	22
Table 2.5 12 Quality Primitives .....	22
Table 2.6 Membership Functions of the Fuzzy System .....	25
Table 2.7 the initial values of the Nonlinear Parameters of the Fuzzy Scheme .....	30
Table 2.8 the initial values of the Linear Parameters of the Fuzzy Scheme.....	30
Table 2.9 Nonlinear Parameters of the Fuzzy Scheme .....	31
Table 2.10 Linear Parameters of the Fuzzy Scheme.....	31
Table 3.1 the Definition of the Membership Functions .....	38
Table 3.2 Premise Part of the Fuzzy Algorithm .....	39
Table 3.3 Nonlinear Parameters of the Fuzzy Algorithm (3.3) .....	44
Table 3.4 the value of the Membership Functions .....	45
Table 3.5 the value of the Premise Part of the Fuzzy Algorithm .....	45
Table 3.6 the value of the Consequent Part of the Fuzzy Algorithm .....	45
Table 3.7 the Definitions of the Membership Functions .....	50
Table 3.8 the Premise Part of the Fuzzy Metric .....	52
Table 3.9 Nonlinear Parameters of the Fuzzy algorithm .....	53
Table 3.10 the value of the Membership Functions .....	54
Table 3.11 the values of the Premise Part of the Fuzzy Algorithm .....	55
Table 3.12 the values of the Consequent Part of the Fuzzy Algorithm .....	55

Table 3.13 the Membership Functions of the Fuzzy Algorithm .....	60
Table 3.14 the Linear Mapping Parameters of the Hybrid Algorithm .....	62
Table 3.15 Nonlinear Parameters of the Fuzzy Algorithm .....	63
Table 3.16 the Simulation Result calculated from the training set of the 4QPFA (4QP) .....	66
Table 3.17 the Simulation Result calculated from the non training set of the 4QPFA (4QP).....	67
Table 3.18 Simulation Result of 4QPFA (4 QP's) .....	67
Table 3.19 the Simulation Result calculated from the training set with the 8QPFA (8QP) .....	70
Table 3.20 the Simulation Result calculated from the non training set with the 8QPFA (8QP) .....	71
Table 3.21 Simulation Result of 8QPFA (8 QP's) .....	71
Table 3.22 the Simulation Result calculated from the training set with the 12QPFA (12 QP) .....	74
Table 3.23 the Simulation Result calculated from the non-training set with the 12QPFA (12QP) .....	75
Table 3.24 Simulation Results of 12QPFA (12 QP's) .....	75
Table 4.1 2D DCT coefficients of the block in the first frame in Figure 4.3 .....	81
Table 4.2 2D DCT coefficients of the block in the first frame in Figure 4.5 .....	83
Table 4.3 2D DCT coefficients of the block in the first frame in Figure 4.7 .....	84
Table 4.4 'Flashing Blocks' found by eye and FBDM .....	92
Table 4.5 Mean Square Error (MSE) between the given Quality Score (4QPFAT and	

4QPFA) and the subjective rating values .....	100
Table.4.6 Mean Square Error (MSE) between the given Quality Score (8QPFAT and 8QPFA) and the subjective rating values .....	102
Table 4.7 Mean Square Error (MSE) between the given Quality Score (12QPFAT and 12QPFA) and the subjective rating values .....	104

## LIST OF ABBREVIATIONS AND SYMBOLS

PSNR	Peak Signal-to-Noise Ratio
MSE	Mean Squared Error
Mb	Mega byte
MPEG	Moving Pictures Expert Group
MPEG-2	Video Compression Standard
VQEG	Video Quality Expert Group
HVS	Human Visual System
Q-factor	Quantization factor
S-T region	Spatial-Temporal region
MOS	Mean Opinion Score
QF	Quality Features
QM	Quality Measures
QP	Quality Primitives
QS	Quality Scores
4QPFA	4 Quality Primitives Fuzzy Algorithms
8QPFA	8 Quality Primitives Fuzzy Algorithms
12QPFA	12 Quality Primitives Fuzzy Algorithms
LMA	Linear Mapping Algorithm
TM	The Temporal Method
SFM	Spatial Frequency Method
FBDM	‘Flashing Blocks’ Detection Method
ODOBM	The Over Dark and Over Bright Method

4QPFAT	Fuzzy Algorithm that includes 4 Spatial Quality Primitives and one temporal Quality Primitives
8QPFAT	Fuzzy Algorithm that includes 8 Spatial Quality Primitives and one temporal Qulaity Primitves
12QPFAT	Fuzzy algorithm that includes 12 Spatial Quality Primitives and one temporal Quality Primitves
2D DCT	Two Dimension Discrete Cosine Transform
DC	Direct Current
AC	Alternating Current
Eqn.	Equation

# Chapter 1

## Introduction

### 1.1 Video Compression and Quality Measurement

Recent years have seen the introduction and widespread acceptance of several varieties of digital video systems. These include digital television broadcasts from satellites (DBS-TV), the US Advanced Television System (ATV), digital movies on a compact disk (DVD), and digital video cassette recorders(DV) and teleconference. In the near future it can be expected to see widespread terrestrial broadcast and cable distribution of digital televisions.

All digital video systems produce an enormous amount of data. In fact, the amount of data generated may be so great that it results in impractical storage, processing, and communication requirements. For example, assuming one byte per pixel, a  $360 \times 288$  pixel monochrome still image occupies 262,144 bytes of storage. Color images may require as much as three times the storage amount required by monochrome images. If the frame rate is 30 pictures/second[1], then the raw data is 60 Mbit/s. This massive storage and bandwidth requirements are a serious concern for many applications involving digital video. Video compression alleviates this problem by reducing the number of bits. Standard lossy digital video compression techniques such as block coding techniques (JPEG[2], MPEG[3], H.261[4]...etc), result in different types of impairments in the reconstructed videos. Among the common types of artifacts that such compression techniques produce are:

- Blocking: A distortion of the image characterized by the appearance of an underlying block encoding structure [5].
- Blurring: A global distortion over the entire image, characterized by reduced sharpness of edges and spatial details [5].
- Spatial Edge Noise (Ringing): A form of busyness characterized by spatially varying distortion in close proximity to the edge of objects [5].
- Block Flashing: A temporal artifact where changes are made across a block from frame to frame, making the block very visible. See Chapter 4.

Detection and measurement techniques of these types of artifacts will be discussed in Chapter 2. Because the human eye is the final arbiter of video quality, subjective assessment methods which utilize human observers to view and rate the video quality are the best methods for assessing coding quality as perceived by the human observers [6]. But they have the disadvantage of being costly and time consuming. Automatic measures which are repeatable and do not depend on the viewing conditions or the mood of the viewer, need to be developed. Such measures are necessarily needed for comparing videos produced by different compression algorithms. The currently used automatic measures such as the Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR) do not compute the perceptual error as detected by the human eye. Section 1.2 will discuss the weaknesses of such measures. The subject of this thesis is to attempt to develop a reliable automatic measure that has high correlation with the subjective ratings.

## 1.2 Weakness of MSE

Today, Mean Square Error (MSE) is widely used as an automatic quality measure, but it is equally widely criticized [6]. MSE is given by:

$$MSE = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N [i(m,n) - \hat{i}(m,n)]^2 \quad (1.1)$$

while Peak Signal to Noise Ratio (PSNR) is given by

$$PSNR = 10 \times \log \left( \frac{255^2}{MSE} \right) \quad (1.2)$$

where M and N are the number of rows and columns of an image respectively and 255 represents the number of bits for quantized images (8 bits),  $i(m,n)$  and  $\hat{i}(m,n)$  are original image pixels and reconstructed image pixels located at (m,n), respectively. In [7] [8], experiments were made to test the correlation between the MSE and the perceived measure of the video quality. It showed that MSE is not a good measure. The fact that the properties of the Human Visual System are not considered in calculating Mean Square Error [9] may result in MSE being badly correlated with subjective quality measures. MSE suffers from several significant weaknesses such as [10]:

1. An arbitrary increase in the MSE does not always lead to a decrease in video quality.
2. Equal values of MSE for the two degraded videos do not imply similar visual quality.
3. MSE does not locate the types of artifacts in degraded videos.

These weaknesses of MSE are not surprising; it is a fact that the human observers do not sum the error over the entire video but deal with it in much more complicated way [3, 6, 7, 8].

### **1.3 Problem Statement**

As mentioned earlier, the only reliable measures of perceived video quality are subjective assessment methods. These methods are expensive and time consuming. Current automatic video quality assessment methods such as MSE are easy to compute but do not measure well the perceptual distortions of videos. The development of an automatic technique that combines the best of subjective and automatic measures, by being easy to compute and sensitive to the properties of the HVS, is needed. This thesis will restrict itself to considering only degradations that result from MPEG-2 video compression.

### **1.4 Figures of Merit**

The automatic quality metrics developed in this thesis should predict the video quality like a human being. And how close the outputs of the automatic metric is to the evaluation of human beings need to be determined.

The performance of an automatic quality metric can be evaluated by the correlation with the subjective ratings. To test the correlation between the output of the automatic metrics and subjective metrics, the Mean Square Error (MSE) is used in this thesis.

### **1.5 Thesis summary**

There are different types of artifacts that are associated with different compression techniques. In here, only three common types of artifacts, (blocking and blurring, spatial edge noise) are concerned. One approach to automatically evaluate the quality of

compressed videos is to first measure these individual artifacts and then to combine them to form an overall quality measurement. The measurements of individual artifacts are called Quality Primitives (QP's). The combination of all of the QP's is called Quality Score (QS). The QS gives the overall estimated quality of the video sequence. This is the approach that will be taken in this thesis.

There has been previous work in developing QP's which is outlined in Chapter 2. One innovation that is presented in this thesis is to consider the context of an artifact, i.e. for examples is it in a high frequency area of the image where the artifact may be masked. The context is taken into account by creating separate QP's for the same artifact, for example, one for "flat" areas, one for "sharp edge" areas. See Chapter 2.

Previous authors [28] [33] have used linear regression to combine the QP's to form the QS. In this thesis a non-linear mapping of the QP's to get a QS using Fuzzy logic [35] is proposed. See Chapter 3.

Finally in Chapter 4, a new QP to measure the 'Flashing Blocks' artifact is introduced.

## Chapter 2

### The Background of the Quality Measurement Metric

#### 2.1 Introduction

This chapter reviews methods for measuring video quality. The methods for measuring video quality can be divided into two kinds, namely, subjective metrics and automatic metrics. In the section 2.2, the subjective metric is introduced. Because the human beings are end user of the video sequences, the subjective metric is the most satisfying metric. But, it is time consuming and money consuming. In this section, automatic methods are mainly introduced.

Automatic methods for measuring video quality (i.e. those that do not require human intervention) can be broken into two classes: those that attempt to model the Human Visual System (HVS Based Methods) and those that attempt to measure expected artifacts and then map them to what human observers score on certain test sets (Artifact Based Methods). Generally speaking the advantage of the HVS Based Methods is that they can deal with a variety of degradations, perhaps, even those not contemplated by the original designer of the method. On the other hand they tend to be complex and suffer from not being able to perfectly model the HVS. The weakness of Artifact Based Methods is that they depend on the designer knowing *a priori* what artifacts to expect. If however, there is an important application that results in well understood degradations then the relative simplicity of Artifact Based Methods may make them appropriate. Video compression is such an application, and in this thesis only degradations from MPEG –2 video compressions are considered.

In Section 2.3, Winkler’s metric [30] an HVS Based metric is presented.

Artifact Based Methods require two distinct operations: first the calculation of the QP's and second the combination of the QP's to form a QS. In Section 2.4, the popular methods for forming QP's (i.e measuring artifacts are presented). Section 2.5 reviews fuzzy logic which will be used in Chapter 3 to combine QP's to form a QS.

## **2.2 Subjective Measurement of Video Quality**

As the end users of videos are human beings, the most reliable video quality measure is subjective rating by human observers[7]. Generally in the subjective rating experiments, videos and their perceptual errors are checked and rated by humans and then the observers' Mean Opinion Score (MOS) is statistically calculated. See[7] for details.

In such experiments, both expert and non-expert observers may be used; non-experts represent the average viewer while experts give better assessments of video quality since they are familiar with videos and their distortions. There are a certain viewing conditions that should be set before evaluation takes place [23]. Among them, viewing distance, viewing angle, monitor size, peak luminance of screen, room lighting, and number of assessors should be considered. These are different evaluation techniques:

1. The absolute evaluation: The observers view a video and assess its quality by assigning it to the category in a given rating scale. See Table 2.1 for examples of categories.
2. The comparative evaluation: A set of videos is ranked from best to worst by observers.
3. Bubble sort evaluation: The observer compares two videos A and B from a group of videos and determines their order. Assuming that the order is AB, the observer takes

a third video and compares it with B to establish the order ABC or ACB. If the order is ACB, then another comparison is made to determine the new order. The procedure continues until all the videos have been used, allowing the best videos to bubble to the top if no ties are accepted [3, 13].

The most commonly used technique is the first one. It uses the rating scale that has been accepted by [26] and appeared in the relevant literature [3, 4, 5, 7]. Table 2.1 lists the rating scale. The mean rating (Mean Opinion Score) of a group of observers who take part in the evaluation is usually computed by [7]:

$$R = \frac{\left( \sum_{k=1}^n s_k n_k \right)}{\left( \sum_{k=1}^n n_k \right)} \quad (2.1)$$

where  $s_k$ =the score corresponding to the  $k^{\text{th}}$  rating,  $n_k$ =the number of observers with that rating, and  $n$ =the number of grades in the scale. It is important to note that the results of subjective rating are affected by a number of factors including

1. type of videos
2. level of expertise of the observers, and
3. experimental conditions

Note	Impairments	Quality
5	Imperceptible	Excellent
4	Perceptible, but not annoying	Good
3	Slightly annoying	Fair
2	Annoying	Poor
1	Very Annoying	Bad

Table 2.1 Rating Scale for Subjective Evaluation [26]

Table 2.1 is an example of Subjective Ratings. But it is very simple, and it is difficult to use this as the standard to evaluate the goodness of the automatic metrics. In this thesis, a different subjective ratings table is used. There are 12 original video sequences compressed by the MPEG-2 [6] using seven different quantization constants. These reconstructed video sequences have been rated in subjective tests. They are shown in Table 2.2. In this table, 0 means the quality of the video sequences is very good and 100 means the quality of the video sequence is very bad. This data is used in this thesis because it is available.

Q-factor	12	16	24	32	36	40	48
Autumn Leaves	2.2	12	17.8	39.7	51.3	54.3	57.2
Bette Pas Bette	1.3	8.7	18.6	22.2	41.3	49.2	46.6
Birches	4.5	6.5	7.5	8.3	8.4	21.3	22.6
Ferris Wheel	4.6	7	9.3	21.5	35.5	37.3	41.6
Flower Garden	8.5	8	10.7	22.7	24.1	30.7	35
Football	1.5	10.2	9.4	37.3	36.8	56.3	46.6
Horseback Riding	3.2	10.6	19.4	28.8	50	52.3	55.5
Mobile&Calendar	0	0.5	11.7	9.2	29.8	29.5	33.6
Sailboat	5.3	9.3	6.5	10.8	14.6	26.1	25.8
Susie	7.2	15.1	11.1	31.8	53	68	70.4
Table Tennis	1.5	7.2	10.2	19.6	31.1	38.7	36.9
Tempette	0.6	2.5	12.1	11.1	23.3	23.7	29.8

Table 2.2 MPEG-2 Encoded at 7.5 Mbps and pure I-frame with constant Q-factor

### 2.3 HVS Based Method of Automatically Measuring Video Quality – Watson’s Method

In this section, an automatic metric that incorporates a simplified model of the human visual system is introduced [30]. The block diagram of this metric is shown in the Figure 2.1

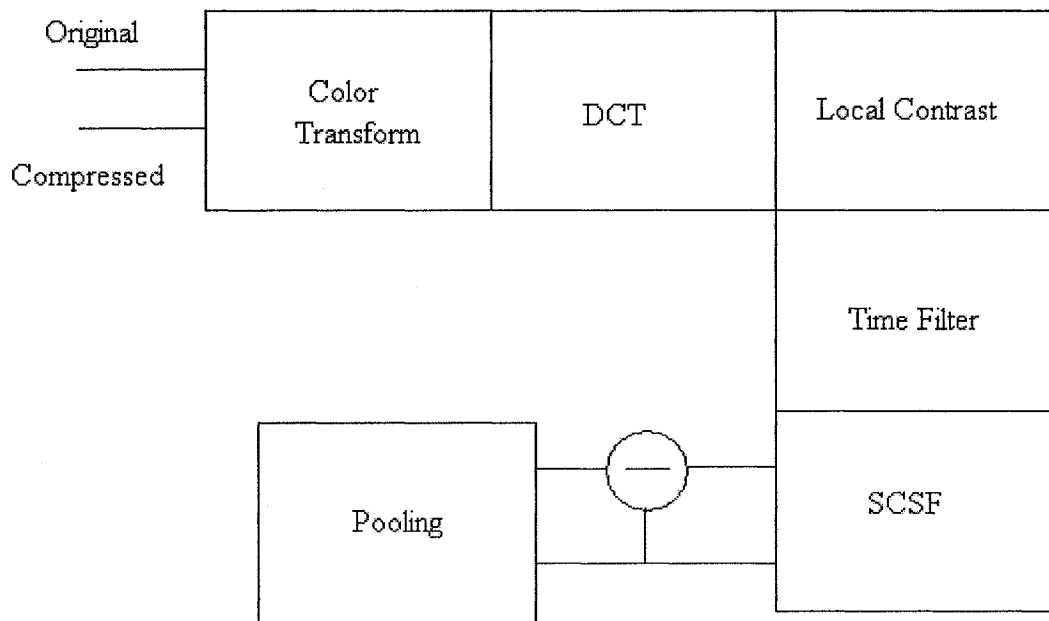


Figure 2.1 Overview of the Objective Quality Metric [30]

- The input to the metric is a pair of color video sequences. The first of the two sequences is the reference and the second is the test sequence.
- The color transform in the process is the conversion of both video sequences to the YUV color space.
- An 8×8 block DCT is applied to each Y frame.
- To establish a better HVS model, DCT coefficients are converted to local contrast coefficients whose ranges are from 1 to -1. The following formula is used :

$LC(i,j)=DCT(i,j) \times H/DC$ . Here DC means the DC coefficient of the DCT coefficients.  $H = \left(\frac{DC}{256}\right)^{0.65}$ . 256 is the mean DCT values for 8 bit images. 0.65 is a psychological parameter.

- Both sequences are then filtered temporally.

The DCT coefficients, now expressed in local contrast form, are converted to Just-Noticeable-Differences (JNDS) by rounding the coefficients of the spatial contrast sensitivity function matrix. At first, the coefficients of each block divide the corresponding coefficients of the following block. Second, rounding the coefficients in the nearest integer. The spatial contrast sensitivity function matrix has the following forms:

$$\begin{bmatrix} 8 & 16 & 19 & 22 & 26 & 27 & 29 & 34 \\ 16 & 16 & 22 & 24 & 27 & 29 & 34 & 37 \\ 19 & 22 & 26 & 27 & 29 & 34 & 34 & 38 \\ 22 & 22 & 26 & 27 & 29 & 34 & 37 & 40 \\ 22 & 26 & 27 & 29 & 32 & 35 & 40 & 48 \\ 26 & 27 & 29 & 32 & 35 & 40 & 48 & 58 \\ 26 & 27 & 29 & 34 & 38 & 46 & 56 & 69 \\ 27 & 29 & 35 & 38 & 46 & 56 & 69 & 83 \end{bmatrix}$$

- The pooling is done by using a weighted pooling. First the two sequences are subtracted to form a difference image (diff). Then the following calculations are made:

$$M1 = 1000 \times \text{mean}_t(\text{mean}_{i,j}(\text{abs}(\text{diff}(i,j,t))))$$

This calculation is done in the two steps. The first step is done in the spatial direction and the average value of each video clip is computed when t is a constant.

The second step is done in the temporal direction and the average value of the first step is computed.

$$M2=1000 \times \text{maximum}_t(\text{maximum}_{i,j}(\text{abs}(\text{diff}(i,j,t))))$$

The calculation is the same as the calculation of the M1. The first step is done in the spatial direction and the maximum value of each video clip is computed. The second step is done in the temporal direction and the maximum value of the output of the first step is computed

$$VQM=(M1+0.005 \times M2)$$

This work was done by Waton [30]. Readers can refer to [30] for the details. Because the human visual system is very complex, it is hard to describe it in the mathematical form. To increase the correlation between the output of this metric and subjective ratings, more complex human visual system model may be needed. It is a very difficult to describe HVS in detail. As an alternative to HVS Based Methods, Artifact Based Methods may be used. The work in this thesis is in Artifact Based Methods which are reviewed in the next two sections.

## **2.4 Quality Primitives (QP's)**

In this section, the first part of the artifact-based automatic video quality measurement is introduced. Measurement of artifacts are first made on small Spatial-Temporal (ST) regions, that are  $8 \times 8 \times 1$  (the 1 being in the temporal direction), the numbers that measure the artifacts on each ST region will be referred to in this thesis as Quality Features (QFs). In Watson (33), one QF represents the magnitude of the gain or loss spatial activity due to the spatial artifacts of the blocking or ringing and the other represents the angle of the

gain or the loss spatial activity due to the blurring[33]. Here  $f_1$  is used to represent the QF of the magnitude of the spatial activities and  $f_2$  is used to represent the QF of the angle of the spatial activities. These QFs are in turn refined into Quality Measures (QMs) which distinguish between the gain or loss of the spatial activity in question (see Equation 2.5). The QMs, like the QFs are calculated for each ST region. To provide information about the entire sequence, all of the QMs in the sequence are pooled to form Quality Primitives (QPs).

#### 2.4.1 Wolf's method

In this section, Wolf's method [28] to compute the QP's is introduced. It is outlined in the following figure:

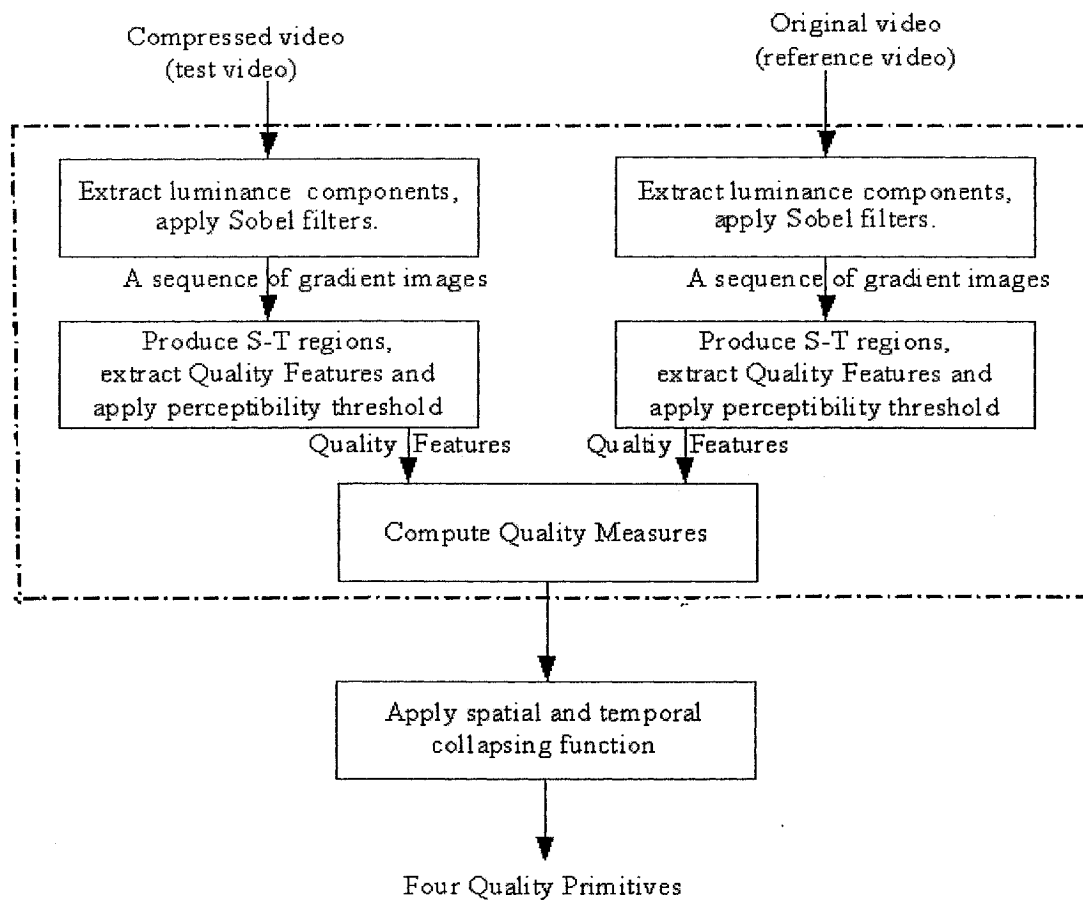


Figure 2.2 Wolf's method to compute the Quality Primitives[33]

From Figure 2.2, the input and output video sequences are first processed with horizontal and vertical edge enhancement filters that increase edges while reducing noise. It has been shown that [11-18] the Sobel filters shown in Figure 2.3 work very well for this step. These two Sobel filters are applied to the original and degraded video sequences separately

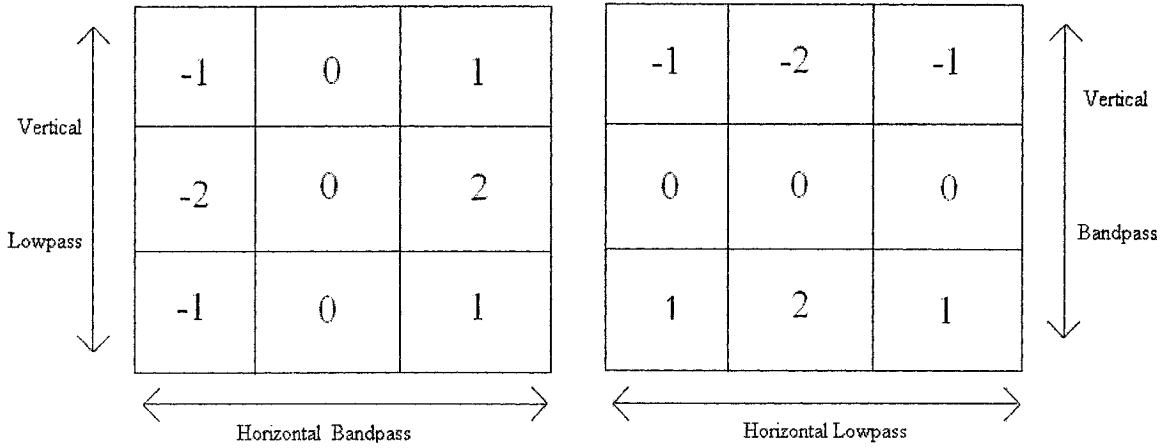


Figure 2.3 Sobel edge enhancement filters

The horizontal and vertical edge enhanced input and output video streams are each divided into localized Spatial-Temporal (S-T) regions. For MPEG-2 video systems, S-T regions size of 8 horizontal pixels  $\times$  8 vertical lines  $\times$  1 are used. [10]. The filters shown in Figure 2.3 increase spatial gradients in the horizontal (H) direction while the transpose of these filters increase spatial gradients in the vertical (V) direction. For a given spatial pixel located at row  $i$ , column  $j$ , and time  $t$ , the H and V filter responses will be noted as  $H(i,j,t)$  and  $V(i,j,t)$ , respectively. These responses can be changed into the form of polar coordinates  $(R, \theta)$  using the following algorithms

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

$$\theta(i, j, t) = \tan^{-1} \left[ \frac{V(i, j, t)}{H(i, j, t)} \right] \quad (2.2)$$

The first QF,  $f_1$ , is computed simply as the standard deviation (stdev) over the S-T region of the  $R(i,j,t)$  samples, and the minimum value of  $f_1$  is  $P$ , namely

$$f_1(i,j,t) = \text{maximum} \{ \text{stdev}[R(i,j,t)], P \}, \quad i,j,t \in \{S-T \text{ region}\}$$

and  $P$  is the perceptibility threshold and is empirically set. This feature measures the magnitude of the spatial activity within a given S-T region.

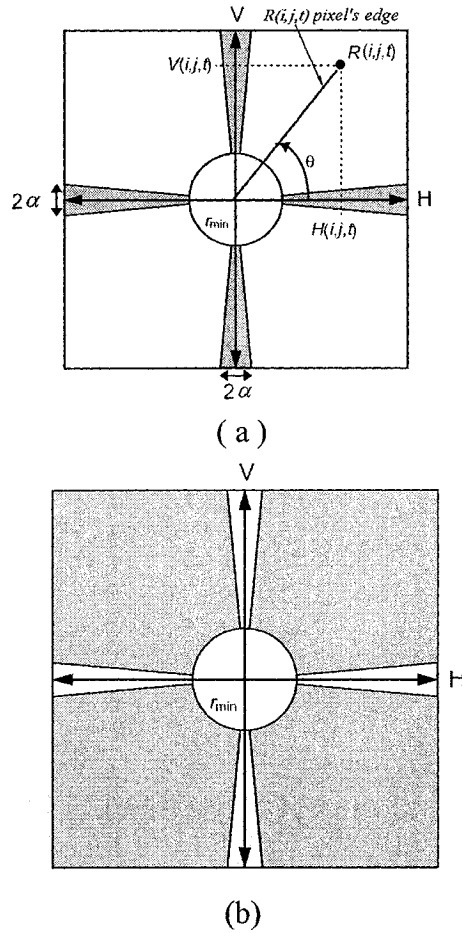


Figure 2.4 Classification of  $\overline{HV}$ . For each pixel the horizontal gradient component (H) and vertical gradient component (V) are plotted. Gradients that fall into the shaded region of (a) form part of the HV image. Those that fall into the shaded region of (b) form part of  $\overline{HV}$  [33].

The second QF,  $f_2$ , is sensitive to changes in the angular distribution, or orientation, of spatial activity. The HV image is made up of the  $R(i,j,t)$  pixels that are in horizontal or

vertical edges. The image  $\overline{HV}$  is made up of the  $R(i,j,t)$  that are in the diagonal edges. Gradient magnitudes  $R(i,j,t)$  less than  $r_{\min}$  are zeroed in both images. Pixels in  $HV$  and  $\overline{HV}$  can be represented mathematically as

$$HV(i,j,t) = \begin{cases} R(i,j,t) & \text{if } R(i,j,t) \geq r_{\min} \quad \text{and} \quad m\frac{\pi}{2} - \Delta\theta < \theta(i,j,t) < m\frac{\pi}{2} + \Delta\theta (m=0,1,2,3) \\ 0 & \text{otherwise} \end{cases}$$

and

$$\overline{HV}(i,j,t) = \begin{cases} R(i,j,t) & \text{if } R(i,j,t) \geq r_{\min} \quad \text{and} \quad m\frac{\pi}{2} + \Delta\theta < \theta(i,j,t) < (m+1)\frac{\pi}{2} + \Delta\theta (m=0,1,2,3) \\ 0 & \text{otherwise} \end{cases}$$

$$(i,j,t) \in (\text{S-T region}) \quad (2.3)$$

QF  $f_2$  for one S-T region is then given by the ratio of the mean of  $HV$  to the mean of  $\overline{HV}$ , the minimum value of  $f_2$  is  $P$ , namely

$$f_2 = \frac{\text{maximum}\{\text{mean}[HV(i,j,t)], P\}}{\text{maximum}\{\text{mean}[\overline{HV}(i,j,t)], P\}} \quad (2.4)$$

$f_2$  is perceptive to changes in the angular distribution of spatial activity with a given S-T region.

The following provides a general description of how Quality Measures are calculated from the original and degraded video sequences. For this discussion, the Quality Features of the original sequences are noted as  $f_{1_o}(i,j,t)$  and  $f_{2_o}(i,j,t)$ . The corresponding Quality Features of the degraded sequences are noted as the  $f_{1_d}(i,j,t)$  and  $f_{2_d}(i,j,t)$ . And  $i, j$  and  $t$  are indices that note the spatial and temporal positions, respectively, of S-T region within the original and degraded video streams. Four Quality Measures needed to be computed due to the gain and loss of the spatial activity (e.g. loss of spatial activity due to blurring and gain of spatial activity due to noise or blocking). They are noted as the  $fl\_gain(i,j,t)$

and  $f1\_loss(i,j,t)$  and  $f2\_gain(i,j,t)$  and  $f2\_loss(i,j,t)$ . For a given S-T region, they are computed using:

$$\begin{aligned}
 f1\_gain(i,j,t) &= pp \left\{ \log_{10} \left[ \frac{f1\_d(i,j,t)}{f1\_o(i,j,t)} \right] \right\} \\
 f1\_loss(i,j,t) &= np \left\{ \log_{10} \left[ \frac{f1\_d(i,j,t) - f1\_o(i,j,t)}{f1\_o(i,j,t)} \right] \right\} \\
 f2\_gain(i,j,t) &= pp \left\{ \log_{10} \left[ \frac{f2\_d(i,j,t)}{f2\_o(i,j,t)} \right] \right\} \\
 f2\_loss(i,j,t) &= np \left\{ \log_{10} \left[ \frac{f2\_d(i,j,t) - f2\_o(i,j,t)}{f2\_o(i,j,t)} \right] \right\} \quad (2.5)
 \end{aligned}$$

where  $pp$  is the positive part operator (i.e., negative values are replaced with zero), and  $np$  is the negative part operator (i.e., positive values are replaced with zero).

Quality Primitives are then computed using pooling. For spatial pooling, for each temporal index  $t$  the average of the largest 5% of the quality measures in eqn.2.5 over the spatial index  $i,j$ . Four numbers are then obtained for each frame, namely:  $f1\_gain(t)$  and  $f1\_loss(t)$  and  $f2\_gain(t)$  and  $f2\_loss(t)$ . The temporal pooling function is computed as the mean of the spatial Quality Primitives over the temporal sequences. The result of the temporal pooling is the QP's of the video sequences. They are  $f1\_gain$  and  $f1\_loss$  and  $f2\_gain$  and  $f2\_loss$ .

Here no account was taken of the context in which an artifact occurs. The next section describes a method to classify S-T regions according to their spatial activity. New Quality Primitives that are obtained by pooling only QM'S that are in the same

classification may result in Quality Scores with increased correlation to human subjective scores.

#### **2.4.2 Bistawhi's method**

Studies of the HVS indicate that not only the strength of an artifact is important, but also its surrounding. The same degradation may be visible in one context but not in another [25]. In this section, work by Bishtawi is explained that attempts to increase the correlation between the QP and the subjective ratings, by putting Quality Primitives (QP's) into context [34].

To this end,  $8 \times 8 \times 1$  blocks of a video clip are classified into three categories: Flat Blocks, Texture Blocks, and Sharp Edge Blocks based on the gradient image. The gradient image is constructed by using the Sobel filter operators shown in Figure 2.3. Following Table 2.3, each block is classified into one of the three categories. See [34] for details.

Figure 2.6 and Figure 2.7 and Figure 2.8 are the results of these classifications.



Figure 2.5 an original video clip



Figure 2.6 Flat Blocks (unblacked blocks)



Figure 2.7 Texture Blocks (unblackened blocks)

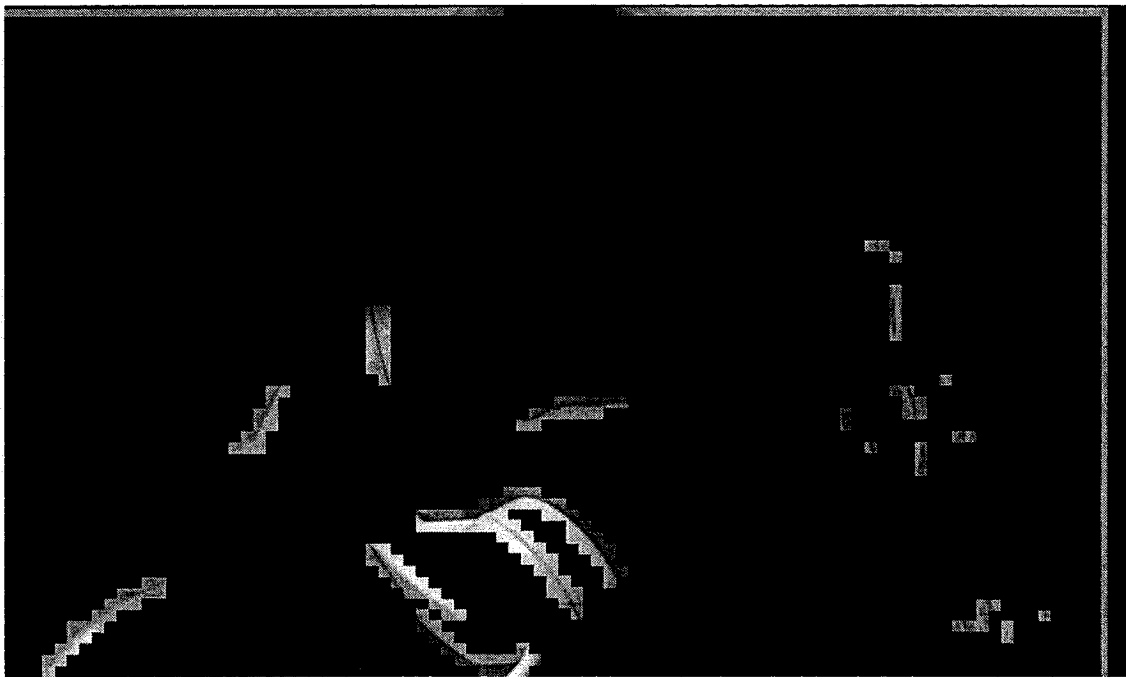


Figure 2.8 Sharp edge blocks (unblackened blocks)

Block Class	Conditions of Acceptance
Flat Blocks	The sobel magnitude of each pixel in a block is less than the threshold
Sharp Edge Blocks	a) A block fall with one of the two situations 1) It has one edge area and two non-edge areas 2) It has one edge area and one non-edge area and b) The sharp edge inside the block has a contour length larger than the threshold.
Texture Area	All the blocks that do not belong to the Flat blocks or the sharp edge blocks.

Table 2.3 Blocks Classes and Their Classification Conditions [34]

The Quality Primitives that describe the artifacts of videos can be obtained by combining Wolf's method and Bishtawi's method. In this combination there are the formulas for the 4 QP's given by Wolf are used. In one scheme Sharp Edge and Texture blocks are not distinguished and so there are only two classifications. This method would yield  $(4 \times 2)$  8 QP's. It is shown in the Table 2.4

Quality Primitive Class	Quality Primitive
Quality Primitives of Flat Blocks	fbf <sub>1</sub> _gain
	fbf <sub>1</sub> _loss
	fbf <sub>2</sub> _gain
	fbf <sub>2</sub> _loss
Quality Primitives of Non-flat Blocks	nfbf <sub>1</sub> _gain
	nfbf <sub>1</sub> _loss
	nfbf <sub>2</sub> _gain
	nfbf <sub>2</sub> _loss

Table 2.4 Eight Quality Primitives

When all three classifications are used there are (4×3) 12 QP's

Quality Primitive Class	Quality Primitive
Quality Primitives of Flat Blocks	fbf <sub>1</sub> _gain
	fbf <sub>1</sub> _loss
	fbf <sub>2</sub> _gain
	fbf <sub>2</sub> _loss
Quality Primitives of Texture Blocks	tbf <sub>1</sub> _gain
	tbf <sub>1</sub> _loss
	tbf <sub>2</sub> _gain
	tbf <sub>2</sub> _loss
Quality Primitives of Sharp Edge Blocks	sebf <sub>1</sub> _gain
	sebf <sub>1</sub> _loss
	sebf <sub>2</sub> _gain
	sebf <sub>2</sub> _loss

Table 2.5 Twelve Quality Primitives

Bishtawi did not have access to subjectively rated video. Thus he was not able to go further.

In Chapter 3, 84 video sequences described in section 2.2 are used to design the artifact-based objective metric that simulates the subjective ratings. The QP's obtained

from these video sequences are used as the input to a nonlinear mapping that simulates subjective ratings. The strategy for generating a non-linear mapping is to use fuzzy logic. Some basics of fuzzy logic are given in the following section.

## 2.5 Introduction to Fuzzy Logic

The next step in the objective artifact-based metric is how to put the QP's of video sequences into an equation that simulates the subjective ratings. In previous publications[32,33], a linear mapping is used. Because of the complexity of the HVS a non-linear mapping may be better suited to creating a QS from the QP's.

Fuzzy Logic will be used to generate such a non-linear mapping because of its function approximation ability.

To illustrate this point, consider a nonlinear function used to describe the subjective rating:

$$g : \alpha \rightarrow \beta$$

where  $\alpha \subset \mathbb{R}^n$  and  $\beta \subset \mathbb{R}$ . This function is represented by a finite number of input-output associations  $(\mathbf{x}, y)$ . With the goal of constructing a fuzzy system

$$\hat{g} : X \rightarrow Y$$

is defined where  $\hat{g}$  is an approximation of  $g$ ,  $X \subset \alpha$  and  $Y \subset \beta$  are some domain and range of interest. By choosing a parameter vector  $\mathbf{P}$  (which includes membership function centers, widths, etc). So that

$$g(\mathbf{x}) = \hat{g}(\mathbf{x}|\mathbf{P}) + e(\mathbf{x})$$

for all  $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^T$  where the approximation error  $e(\mathbf{x})$  is as small as possible.

A dynamic Takagi-Sgeno (T-S or T-S-K) fuzzy scheme [36] is described by a set of fuzzy “IF-THEN” rules, with fuzzy sets in the antecedents (the “if” conditions) and local linear functions in the consequents (the “then” action). Every  $i$ -th rule of a T-S fuzzy scheme has the following form:

$i^{\text{th}}$  Rule

IF  $x_1$  is  $R_{i1}$ , ...,  $x_j$  is  $R_{ij}$ , ...,  $x_n$  is  $R_{in}$

Then  $g_i(\mathbf{x}) = a_{i1}x_1 + \dots + a_{ij}x_j + \dots + a_{in}x_n + b_i$

where  $i=1, \dots, r$  with  $r$  the number of rules; the  $x_j (j=1, \dots, n)$  are the input variables. The  $R_{ij}$  are the linguistic values to describe the input variables. Examples of linguistic variables are “big” or “small”. In fuzzy logic, these linguistic values describe fuzzy sets. They are mathematically represented by membership functions ( $mf_{ij}$ ). Here  $mf_{ij}(x_j)$  returns a number which describes the degree to which  $x_j$  belongs to  $R_{ij}$ .

The  $a_{ij}$  and  $b_i$  are linear coefficients of the output of rule  $i$ .  $g_i$  is the output of rule  $i$ .

The final output of the fuzzy scheme is inferred as follows:

$$\hat{g} = \frac{\sum_{i=1}^R B_i(\mathbf{x}) g_i(\mathbf{x})}{\sum_{i=1}^R B_i(\mathbf{x})} \quad (2.6)$$

where  $B_i(\mathbf{x})$  can be defined (and in this thesis is defined) as the product of the membership functions in rule  $i$ , i.e.  $B_i(\mathbf{x}) = mf_{i1}(x_1)mf_{i2}(x_2)\dots mf_{in}(x_n)$ .

According to [36], the parameters can be tuned by using a hybrid learning procedure- Adaptive-Network-Based Fuzzy Inference System. The parameters of the Takagi-Sugeno can be divided into two kinds: antecedent parameters (nonlinear parameters) and consequent parameters (linear parameters). In [36], the tuning method for the antecedent

parameters is gradient decent. The tuning method for the consequent parameters is least squares.

For example, consider a nonlinear function

$$g(x) = \frac{\sin(3x)}{\exp\left(\frac{x}{2}\right)} \quad (2.7)$$

Here, four linguistic values are used to describe the input variable (x). They are “very big” and “big” and “small” and “very small”. The fuzzy scheme has the following forms:

If x is very small, then  $g_1(x) = a_1x + b_1$

If x is small, then  $g_2(x) = a_2x + b_2$

If x is big, then  $g_3(x) = a_3x + b_3$

If x is very big, then  $g_4(x) = a_4x + b_4$

The membership functions are described as the following tables:

‘Linguist Values’	Membership Function
‘very small’	$mf_1(x) = \exp\left(\frac{(x - m_1)^2}{-2\sigma_1^2}\right)$
‘small’	$mf_2(x) = \exp\left(\frac{(x - m_2)^2}{-2\sigma_2^2}\right)$
‘big’	$mf_3(x) = \exp\left(\frac{(x - m_3)^2}{-2\sigma_3^2}\right)$
‘very big’	$mf_4(x) = \exp\left(\frac{(x - m_4)^2}{-2\sigma_4^2}\right)$

Table 2.6 the Membership Functions of the Fuzzy System

$B_i(x)$  is the product of the membership functions in the rule  $i$  and can be expressed as the following forms:

$$B_1(x) = mf_1(x)$$

$$B_2(x) = mf_2(x)$$

$$B_3(x) = mf_3(x)$$

$$B_4(x) = mf_4(x)$$

Note that here since there is only a single input variable the argument of  $B_i$  is scalar and there is only one membership function per rule.

According to (2.6), the mathematical form of the fuzzy system is:

$$\hat{g}(x) = \frac{\sum_{i=1}^4 B_i(x) g_i(x)}{\sum_{i=1}^4 B_i(x)} \quad (2.8)$$

The parameters of the fuzzy system can be divided into two kinds: nonlinear parameters antecedent parameters  $(m_i, \sigma_i \dots)$   $i=1:4$  and linear consequent parameters  $(a_i, b_i \dots)$   $i=1:4$ .

To determine the linear parameters least squares is used. Here training data is given in the form of  $(x_k, g(x_k))$  where  $x_k$  is an element of  $\mathcal{R}^n$  and  $k=1..K$ . Here  $K$  is the number of points in the training set. Note that in the illustrative example here  $n=1$  and  $x_k$  is a scalar.

Define

$$S(x) = \sum_{i=1}^4 B_i(x) = \exp\left(\frac{(x-m_1)^2}{-2\sigma_1^2}\right) + \exp\left(\frac{(x-m_2)^2}{-2\sigma_2^2}\right) + \exp\left(\frac{(x-m_3)^2}{-2\sigma_3^2}\right) + \exp\left(\frac{(x-m_4)^2}{-2\sigma_4^2}\right)$$

which is the denominator of Equation 2.8.

Also define

$$\mathbf{A} = \begin{bmatrix} \frac{B_1(x_1)x_1}{S(x_1)} & \frac{B_1(x_1)}{S(x_1)} & \frac{B_2(x_1)x_1}{S(x_1)} & \frac{B_2(x_1)}{S(x_1)} & \frac{B_3(x_1)x_1}{S(x_1)} & \frac{B_3(x_1)}{S(x_1)} & \frac{B_4(x_1)x_1}{S(x_1)} & \frac{B_4(x_1)}{S(x_1)} \\ \frac{B_1(x_2)x_2}{S(x_2)} & \frac{B_1(x_2)}{S(x_2)} & \frac{B_2(x_2)x_2}{S(x_2)} & \frac{B_2(x_2)}{S(x_2)} & \frac{B_3(x_2)x_2}{S(x_2)} & \frac{B_3(x_2)}{S(x_2)} & \frac{B_4(x_2)x_2}{S(x_2)} & \frac{B_4(x_2)}{S(x_2)} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \frac{B_1(x_K)x_K}{S(x_K)} & \frac{B_1(x_K)}{S(x_K)} & \frac{B_2(x_K)x_K}{S(x_K)} & \frac{B_2(x_K)}{S(x_K)} & \frac{B_3(x_K)x_K}{S(x_K)} & \frac{B_3(x_K)}{S(x_K)} & \frac{B_4(x_K)x_K}{S(x_K)} & \frac{B_4(x_K)}{S(x_K)} \end{bmatrix}$$

Note that  $\mathbf{A}$  is  $K \times 8$  in this example.

Furthermore define

$$\mathbf{LP} = \begin{bmatrix} a_1 \\ b_1 \\ a_2 \\ b_2 \\ a_3 \\ b_3 \\ a_4 \\ b_4 \end{bmatrix}$$

which contains the linear parameters of the model. Then

$$\hat{\mathbf{G}} = \mathbf{A} \times \mathbf{LP} \tag{2.9}$$

where  $\hat{\mathbf{G}}$  is a  $K \times 1$  vector. Each element of  $\hat{\mathbf{G}}$  is the  $\hat{g}(x_k)$  (see Equation 2.8) for each of the data points.

Define:

$$\mathbf{G} = \begin{bmatrix} g(x_1) \\ g(x_2) \\ \cdot \\ g(x_K) \end{bmatrix}$$

where  $k=1 \dots K$ . Each element of  $G$  contains the value of the function to be estimated for each data point.

Using the least square method,  $\mathbf{LP}$  can be obtained so that  $\|\hat{G} - G\|^2$  is minimized.

For the tuning of the nonlinear parameters  $(m_i, \sigma_i)$   $i=1:4$ , the gradient method is used.

Consider the mean square error between the output of the fuzzy system and the nonlinear function

$$E_k = (g(x_k) - \hat{g}(x_k))^2 \quad (2.10)$$

where  $k=1..K$  and define a vector containing all the nonlinear parameters:

$$\mathbf{NP} = \begin{bmatrix} m_1 \\ \sigma_1 \\ m_2 \\ \sigma_2 \\ m_3 \\ \sigma_3 \\ m_4 \\ \sigma_4 \end{bmatrix}$$

The formula is as follows:

$$\mathbf{NP}(K) = \mathbf{NP}(K-1) - \ell \nabla_{\mathbf{NP}} E_k \quad (2.11)$$

where  $\mathbf{NP}(k)$  and  $\mathbf{NP}(k-1)$  represent the nonlinear parameters. The gradient descent does  $K$  steps, one for each data point in the training set. Note that  $\mathbf{NP}(0)$  is some initial condition the first time gradient descent is run, and is  $\mathbf{NP}(K)$  of the last gradient descent in gradient descents run after the first one.

In this illustrative example

$$\nabla_{NP} = \left[ \frac{\partial E_k}{\partial (NP)_i} \right] = \begin{bmatrix} \frac{\partial E_k}{\partial m_1} \\ \frac{\partial E_k}{\partial \sigma_1} \\ \frac{\partial E_k}{\partial m_2} \\ \frac{\partial E_k}{\partial \sigma_2} \\ \frac{\partial E_k}{\partial m_3} \\ \frac{\partial E_k}{\partial \sigma_3} \\ \frac{\partial E_k}{\partial m_4} \\ \frac{\partial E_k}{\partial \sigma_4} \end{bmatrix} = \begin{bmatrix} \frac{2(\hat{g}(x_k) - g(x_k))\partial \hat{g}(x_k)}{\partial m_1} \\ \frac{2(\hat{g}(x_k) - g(x_k))\partial \hat{g}(x_k)}{\partial \sigma_1} \\ \frac{2(\hat{g}(x_k) - g(x_k))\partial \hat{g}(x_k)}{\partial m_2} \\ \frac{2(\hat{g}(x_k) - g(x_k))\partial \hat{g}(x_k)}{\partial \sigma_2} \\ \frac{2(\hat{g}(x_k) - g(x_k))\partial \hat{g}(x_k)}{\partial m_3} \\ \frac{2(\hat{g}(x_k) - g(x_k))\partial \hat{g}(x_k)}{\partial \sigma_3} \\ \frac{2(\hat{g}(x_k) - g(x_k))\partial \hat{g}(x_k)}{\partial m_4} \\ \frac{2(\hat{g}(x_k) - g(x_k))\partial \hat{g}(x_k)}{\partial \sigma_4} \end{bmatrix} \quad (2.12)$$

Note that  $\ell$  is a learning rate which can be expressed as

$$\ell = \frac{m}{\sqrt{\sum_i \left( \frac{\partial E_k}{\partial NP_i} \right)^2}} \quad (2.13)$$

m is the step size. The value of m can be changed to vary the speed of convergence

The tuning procedure can be divided into the following steps:

1. Arbitrarily choose the initial value of the premise (non linear) parameters  $(m_1, \sigma_1)$ .
2. Use the least squares (2.9) method to tune the consequent (linear) parameters  $(a_1, b_1)$ ;
3. Use the gradient method (2.11) for K points in the training set to tune the premise parameters;
4. Repeat steps 2 and step 3, until the difference of the mean square error (2.10) in two iterations is small enough. Here  $E_k(n)$  is mean square error of kth point in the

nth iterations.  $E_k(n-1)$  is mean square error of kth point in the (n-1)th iterations. It is obtained that  $|E_k(n) - E_k(n-1)| < \varepsilon$ .  $\varepsilon$  is a small number and is set 1 in this example.

For this illustrative example,  $\alpha \in \mathfrak{R}$  is a scalar and  $\beta \in \mathfrak{R}$ . And the domain of interest,  $X \in \alpha$  is from 0 to 10. The range of interest  $Y \in \beta$  is the corresponding range of the function  $\sin(3x)/\exp(x/2)$ . The number (K) of the training data (x,y) is 100. In this example the  $x_k$  are chosen as being evenly spaced between 0 and 10. The  $g(x_k)$  are generated from the function without noise. After choosing the following initial values and finishing the 10000 iterations (i.e. repeating steps 2 and 3 10000 times):

$m_1=1$	$\sigma_1 = 0.5$
$m_2=2$	$\sigma_2 = 1.0$
$m_3=5$	$\sigma_3 = 1.5$
$m_4=6$	$\sigma_4 = 2.0$

Table 2.7 the initial values of Nonlinear Parameters of the Fuzzy Scheme

$a_1=2$	$b_1 = 3$
$a_2=-1.5$	$b_2 = 2.5$
$a_3=-2$	$b_3 = 3.3$
$a_4=-4$	$b_4 = -3$

Table 2.8 the initial values of the Linear Parameters of the Fuzzy Scheme

The result is as follows:

For the nonlinear parameters:

$m_1=1.005$	$\sigma_1 = 0.3082$
$m_2=3.412$	$\sigma_2 = 0.4594$
$m_3=4.271$	$\sigma_3 = 0.6261$
$m_4=9.002$	$\sigma_4 = 3.689$

Table 2.9 Nonlinear Parameters of the Fuzzy Scheme

For the linear Parameters:

$a_1=-1.724$	$b_1 = 1.864$
$a_2=-1.197$	$b_2 = 3.642$
$a_3=-0.3015$	$b_3 = 1.64$
$a_4=0.01205$	$b_4 = -0.09853$

Table 2.10 Linear Parameters of the Fuzzy Scheme

The result can be expressed as the following figures:

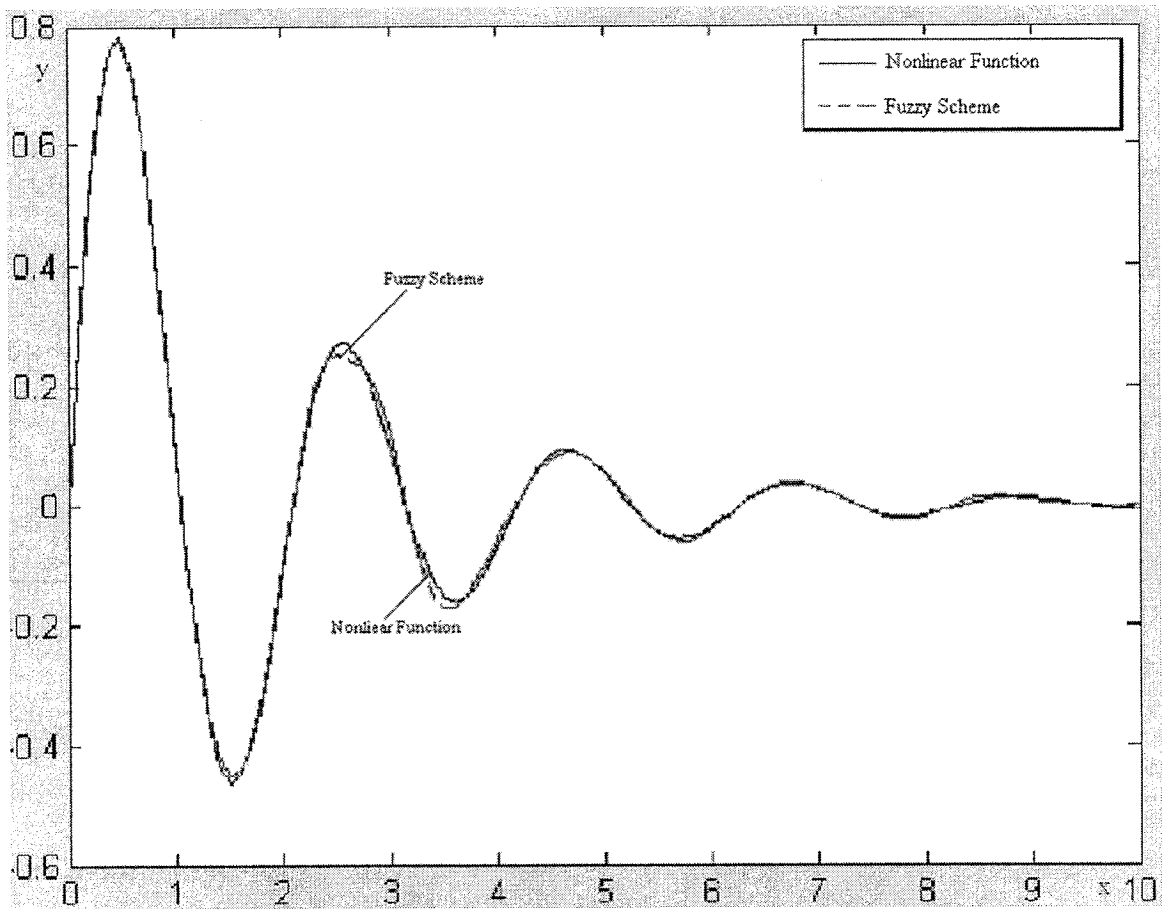


Figure 2.9 the Approximation Ability of the Fuzzy Logic

In Chapter 3, fuzzy logic is used to approximate the nonlinear mapping that simulates the subjective ratings of human beings. It is the main contribution of this thesis.

## 2.6 Conclusions:

In this chapter, some video quality measurement techniques are presented. First, the non-automatic measurement techniques are presented. These techniques are done by humans. Although they are the most reliable quality measures, they are costly, time consuming, and depend on the test conditions. Second, objective measures are presented. These objective metrics can be divided into two kinds. One kind is based the human

visual system (HVS), the other is based on artifacts of the video sequences. Because the HVS is very complex, it is difficult to design a satisfying objective metric that is based on the HVS. In the remainder of this thesis, the artifact-based objective metric is used to simulate the subjective ratings. The first step of this metric is to obtain the Quality Primitives of the video sequences. The Quality Primitives obtained are used as the input to a nonlinear mapping that simulates the subjective ratings. In the remainder of this thesis fuzzy logic is used as the framework for the non-linear mapping.

## Chapter 3

### Fuzzy Metrics with Spatial Quality Primitives

#### 3.1 Introduction

Since the human eye weighs video impairments in a complex way[34], a nonlinear mapping of QP's may result in a QS which more closely matches the MOS's of human observers than a linear mapping. Because of strong function approximation ability of the fuzzy system mentioned in Section 2.5, the Takagi-Sugeno (T-S) fuzzy system [36] is used to approximate as the framework for the nonlinear mapping. As was mentioned in Chapter 2, the QP's combining Wolf's method [28] and Bishtawi's method [33], are used as the input to the fuzzy system.

In the Section 3.2, four Quality Primitives selected from Table 2.4 (fbf<sub>1</sub>\_gain, fbf<sub>2</sub>\_loss, nfbf<sub>1</sub>\_loss, nfbf<sub>2</sub>\_loss) as the input to the fuzzy system. In Section 3.3, 8 QP's (Table 2.4) is used as the input to the fuzzy system. In Section 3.4, 12 QP's (Table 2.5) is used as the input to the fuzzy system. To reduce the complexity of the fuzzy algorithm, fuzzy c-means [36] is used to select the rules of the fuzzy algorithm. Each of Sections 3.2, 3.3 and 3.4 can be divided into three parts. In the first part the fuzzy algorithm used to simulate the subjective evaluations is used in detail so that readers can implement it. In the second part, the method to tune the parameters of the fuzzy model-Least Square and Gradient Descent is given. In the last part, an example to compute the fuzzy algorithm using the tuned parameters is given. In the Section 3.5, the simulation result is given. And in the last section, the conclusion is given.

### 3.2. Fuzzy Model Using 4 QP's:

In this section, 4 QP's selected from Table.2.4 (fbf<sub>1</sub>\_gain, fbf<sub>2</sub>\_loss, nfbf<sub>1</sub>\_loss, nfbf<sub>2</sub>\_loss) are used as the input to the nonlinear mapping that simulate the subjective ratings. It is also possible to choose 4 QP's from the Table 2.5. But it is too complex. Another method is discussed in the Section 3.4 when inputs are composed 12 QP's. They are represented as  $x=[x_1 \ x_2 \ x_3 \ x_4]$ . Two linguistic values (big and small) are used to describe every QP's. If all the possibilities combining linguistic values are enumerated, there will be  $2^4=16$  rules. It is very complex. To reduce the number of the rules, a technique-Fuzzy Clustering is used to select the rules. Fuzzy Clustering is the partitioning of data into subsets of groups based on the similarities between the data and can be implemented by using an algorithm called Fuzzy C-Means [35]. Fuzzy C-Means is an iterative algorithm used to find cluster centers  $c^j$  (vectors of dimension  $4 \times 1$ ) to satisfy the following inequalities.

$$\left| c_n^j - c_{n-1}^j \right| < \varepsilon_c \quad (3.1)$$

Let choose  $\varepsilon_c = 0.001$  and the update formula for Fuzzy C Means is:

$$c_n^j = \frac{\sum_{i=1}^M x_i (u_{ij})}{\sum_{i=1}^M (u_{ij})} \quad (3.2)$$

$$\text{where } u_{ij} = \left[ \sum_{k=1}^R \left( \frac{\left| x_i - c_{n-1}^j \right|^2}{\left| x_i - c_{n-1}^k \right|^2} \right) \right]$$

Here, M is the number of input-output data pairs in the training data set. And  $n=1, \dots, N$  represent the number of the iterations.  $c_0^j$  represents the initial values of the iterations.

From the section 2.2, 8 video sequences compressed at 7 different quantization levels are used as the training set. They are Autumn Leaves, Sailboat, Flower Garden, Mobile&Calendar, Table Tennis, Bette Pas Bette, Susie, Ferris Wheel.  $M$  is equal to  $8 \times 7 = 56$ .  $R$  is the number of the rules needed to calculate, in here  $R=2$ .  $\mathbf{x}_i$  for  $i=1, \dots, M$  is the input vector of the input-output training data pairs,  $\mathbf{c}^j = [c_1^j \ c_2^j \ c_3^j \ c_4^j]^T$  for  $j=1, \dots, R$  are the cluster centers. The method is as follows:

1. Give the initial values (  $\mathbf{c}_0^j$  )
2. Using the (3.2) to compute the  $\mathbf{c}_n^j$
3. If  $|\mathbf{c}_n^j - \mathbf{c}_{n-1}^j| < \varepsilon_c$ , the number of the rules is decided, otherwise, it is needed to repeat step 1 and step 2, until this inequality  $|\mathbf{c}_n^j - \mathbf{c}_{n-1}^j| < \varepsilon_c$  is set up. In here, the initial cluster centers are  $\mathbf{c}_0^1 = [0.1 \ 0.2 \ 0.3 \ 0.5]$   $\mathbf{c}_0^2 = [0.3 \ 0.5 \ 0.8 \ 0.7]$  and the number of the iteration is 1000, and the (3.1) is satisfied. So  $R=2$ .

In [36], the author gives a detail description of the fuzzy c-means and the readers can refer to it.

### 3.2.1 Proposed Fuzzy Metric.

After using fuzzy c-means, the proposed fuzzy algorithm is as follows:

If  $x_1$  is big and  $x_2$  is small and  $x_3$  is small and  $x_4$  is small, then

$\text{join}_1(\mathbf{x})$

If  $x_1$  is small and  $x_2$  is big and  $x_3$  is big and  $x_4$  is big, then

$\text{join}_2(\mathbf{x})$

There are other possibilities combining linguistic values (for example  $x_1$  is small and  $x_2$  is small and  $x_3$  is big and  $x_4$  is small) . But by using fuzzy c-means, it is found that the above rules are good enough to describe the fuzzy metric. This fuzzy metric can be divided into two parts-premise part (if section) and consequent part (then section) and can be describe in mathematically as follow forms:

$$\text{Join}(\mathbf{x}) = \frac{\sum_{i=1}^2 B_i(\mathbf{x}) \text{join}_i(\mathbf{x})}{\sum_{i=1}^2 B_i(\mathbf{x})} = \sum_{i=1}^2 \text{Normal}_i(\mathbf{x}) \text{join}_i(\mathbf{x}) \quad (3.3)$$

The following figure can be used to express (3.3)

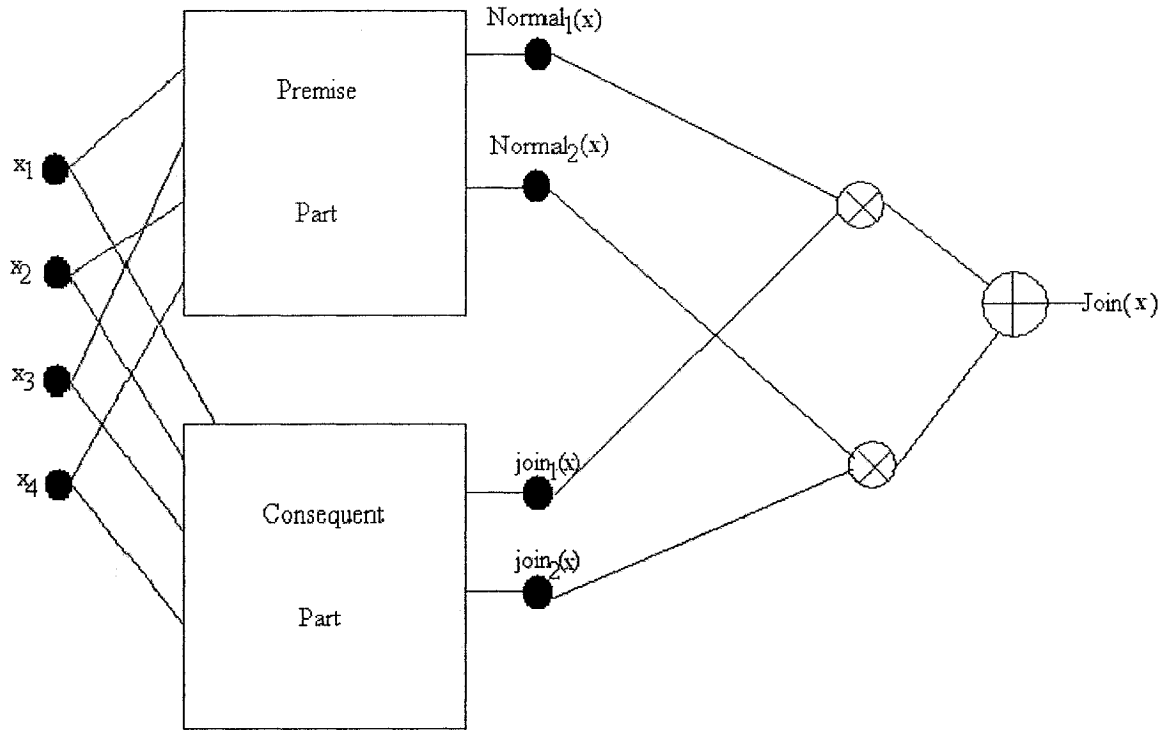


Figure 3.1 Figure of the Fuzzy Algorithm (3.3)

where  $\text{join}_i(\mathbf{x})$  is the output of the consequent part and  $\text{Normal}_i(\mathbf{x})$  is the normalized output of the premise part. The Sum of the product of the output of the premise part and the consequent part ( $\text{join}_i(\mathbf{x}) \times \text{Normal}_i(\mathbf{x})$ ) is the output of the fuzzy system i.e.  $\text{Join}(\mathbf{x})$ .

### 3.2.1.1 The Premise Part of the Fuzzy Metric

For the premise part, the following membership functions are used:

Quality Primitives	Linguistic Values	Membership Function
$x_1$	'big'	$mf_1(x_1) = e^{\frac{-(x_1 - m_{11})^2}{2(\sigma_{11})^2}}$
$x_1$	'small'	$mf_2(x_1) = e^{\frac{-(x_1 - m_{12})^2}{2(\sigma_{12})^2}}$
$x_2$	'big'	$mf_3(x_2) = e^{\frac{-(x_2 - m_{21})^2}{2(\sigma_{21})^2}}$
$x_2$	'small'	$mf_4(x_2) = e^{\frac{-(x_2 - m_{22})^2}{2(\sigma_{22})^2}}$
$x_3$	'big'	$mf_5(x_3) = e^{\frac{-(x_3 - m_{31})^2}{2(\sigma_{31})^2}}$
$x_3$	'small'	$mf_6(x_3) = e^{\frac{-(x_3 - m_{32})^2}{2(\sigma_{32})^2}}$
$x_4$	'big'	$mf_7(x_4) = e^{\frac{-(x_4 - m_{41})^2}{2(\sigma_{41})^2}}$
$x_4$	'small'	$mf_8(x_4) = e^{\frac{-(x_4 - m_{41})^2}{2(\sigma_{41})^2}}$

Table 3.1 the Definition of the Membership Functions

The premise part of the fuzzy system can be expressed as follows:

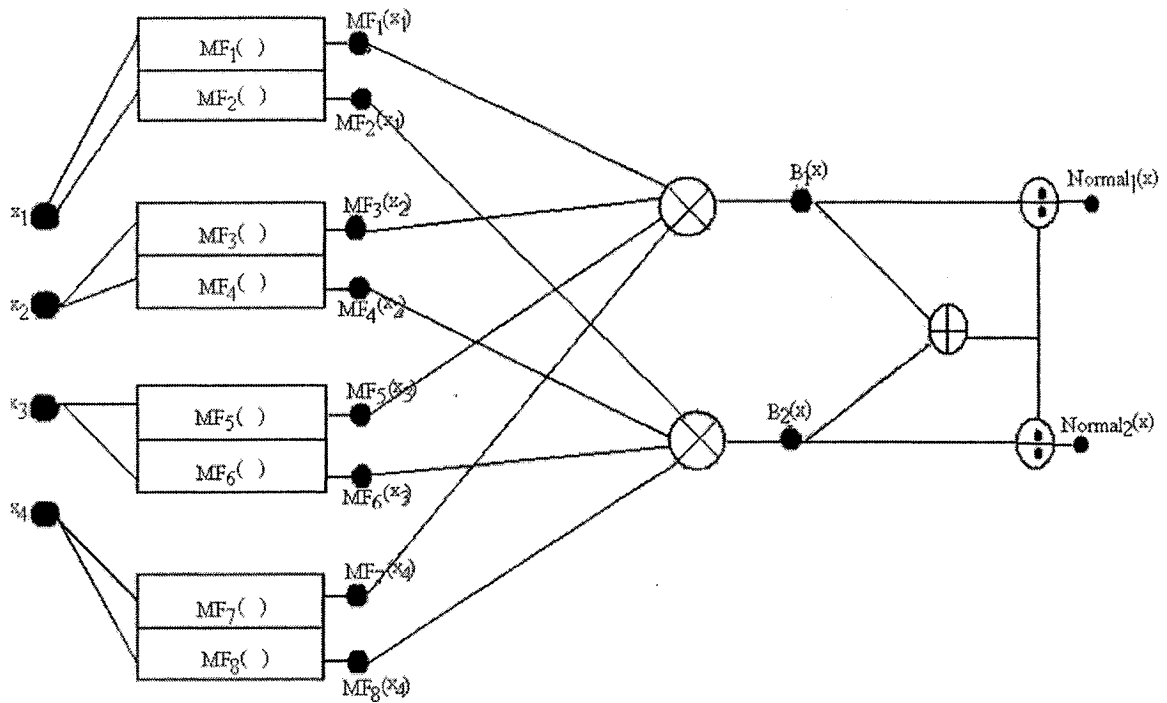


Figure 3.2 Premise Part of the Fuzzy System

The input of this figure is the QP's measuring the spatial artifacts of the video sequences.

The output of this figure is the normalized product of the membership functions, for

example,  $Normal_1(x) = \frac{B_1(x)}{\sum_{i=1}^2 B_i(x)}$ . To illustrate this figure more clearly, let consider the

table where the input  $B_i(x)$  can be thought as the product of the memberships whose label are dots.

	$MF_1(x_1)$	$MF_2(x_1)$	$MF_3(x_2)$	$MF_4(x_2)$	$MF_5(x_3)$	$MF_6(x_3)$	$MF_7(x_4)$	$MF_8(x_8)$
$B_1(x)$	●		●		●		●	
$B_2(x)$		●		●		●		●

Table 3.2 Premise Part of the Fuzzy Algorithm

### 3.2.1.2 The Consequent Part of the Fuzzy Metric

The consequence parts of the fuzzy system can be expressed as the following matrix forms:

$$\begin{bmatrix} join_1(x) \\ join_2(x) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & b_1 \\ a_{21} & a_{22} & a_{23} & a_{24} & b_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ 1 \end{bmatrix} \quad (3.4)$$

The equation (3.4) can be expressed in the Matrix forms:

$$\mathbf{JOIN} = \mathbf{L} \times \mathbf{xI} \quad (3.5)$$

where **JOIN** is the 2×1 vector in equation (3.4), **L** is the 2×5 matrix in equation (3.4), **xI** is 5×1 vector containing the QP's and one in equation (3.4).

### 3.2.2 Tuning Method

As mentioned in Section 2.5, the parameters of the Takagi-Sugeno fuzzy scheme can be divided into two kinds-premise parameters (NP) (nonlinear parameters for example  $\sigma_{11}, m_{11}$ ) and consequent parameters (LP) (linear parameters for example  $a_{11}, b_1$ ). The tuning method for the premise parameters is gradient descent. The tuning method for the consequent parameters is least squares. Here training data is given in the form of (  $\mathbf{x}_k, s_k$ ) where  $\mathbf{x}_k$  is a vector whose elements ( $x_{k1}, x_{k2}, x_{k3}, x_{k4}$ ) are the QP's for the corresponding video sequences and  $s_k$  is the corresponding subjective ratings for this video and  $k=1..K$ .

Here K is the number of points in the training set and  $K=56$

At first, from the equation (3.3), for every  $\mathbf{x}_k$ , there exists a  $Join(\mathbf{x}_k)$ . Putting all the  $\mathbf{x}_k$  and  $Join(\mathbf{x}_k)$   $k=1..K$  together, the matrix equation can be obtained:

$$A \times LP = \hat{S} \quad (3.6)$$

Because there are 10 linear parameters in the fuzzy scheme (3.3) and the number of points in the training set is K (K=56), the size of A is a matrix of K×10. And A has the following forms:

$$A = \begin{bmatrix} \frac{B_1(x_1)x_{11}}{\sum_{i=1}^2 B_i(x_1)} & \frac{B_1(x_1)x_{12}}{\sum_{i=1}^2 B_i(x_1)} & \frac{B_1(x_1)x_{13}}{\sum_{i=1}^2 B_i(x_1)} & \cdot & \cdot & \cdot & \frac{B_2(x_1)}{\sum_{i=1}^2 B_i(x_1)} \\ \frac{B_1(x_2)x_{21}}{\sum_{i=1}^2 B_i(x_2)} & \frac{B_1(x_2)x_{22}}{\sum_{i=1}^2 B_i(x_2)} & \frac{B_1(x_2)x_{23}}{\sum_{i=1}^2 B_i(x_2)} & \cdot & \cdot & \cdot & \frac{B_2(x_2)}{\sum_{i=1}^2 B_i(x_2)} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \frac{B_1(x_K)x_{K1}}{\sum_{i=1}^2 B_i(x_K)} & \frac{B_1(x_K)x_{K2}}{\sum_{i=1}^2 B_i(x_K)} & \frac{B_1(x_K)x_{K2}}{\sum_{i=1}^2 B_i(x_K)} & \cdot & \cdot & \cdot & \frac{B_2(x_K)}{\sum_{i=1}^2 B_i(x_K)} \end{bmatrix}$$

LP is an unknown vector whose elements are linear parameters

$$LP = \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{14} \\ b_1 \\ a_{21} \\ a_{22} \\ a_{23} \\ a_{24} \\ b_2 \end{bmatrix}$$

where  $\hat{S}$  is a 56×1 vector. Each element of  $\hat{S}$  is the **Join**( $x_k$ ) (see Equation 3.3) for each of the data points.

Define:

$$\mathbf{S} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_k \end{bmatrix}$$

Each element of  $\mathbf{S}$  contains the subjective ratings for each video sequence. Using the least square method,  $\mathbf{LP}$  can be obtained so that  $\|\mathbf{S} - \hat{\mathbf{S}}\|^2$  is minimized.

Let assume the mean square error between the fuzzy output and subjective rating is

$$E_k = (s_k - \text{Join}(\mathbf{x}_k))^2 \quad (3.7)$$

where  $k=1..K$  and define a vector containing all the nonlinear parameters:

$$\mathbf{NP} = \begin{bmatrix} m_{11} \\ \sigma_{11} \\ m_{12} \\ \sigma_{12} \\ \vdots \\ \vdots \\ \vdots \\ \sigma_{42} \end{bmatrix}$$

where  $\mathbf{NP}$  is a  $16 \times 1$  vector.  $\mathbf{NP}_i$  represents the nonlinear parameters.  $i=1,..,16$ .

The formula is as follows:

$$\mathbf{NP}(k) = \mathbf{NP}(k-1) - \ell \nabla_{\mathbf{NP}} E_k \quad (3.8)$$

where  $\mathbf{NP}(k)$  and  $\mathbf{NP}(k-1)$  represent the nonlinear parameters. The gradient descent does  $K$  steps, one for each sequence in the training set. Note that  $\mathbf{NP}(0)$  is some initial condition when the first time of gradient descent is run, and is  $\mathbf{NP}(K)$  of the last gradient descent in gradient descents run after the first one.

In this section

$$\nabla_{NP} = \left[ \frac{\partial E_k}{\partial (NP)_i} \right] = \begin{bmatrix} \frac{\partial E_k}{\partial m_{11}} \\ \frac{\partial E_k}{\partial \sigma_{11}} \\ \frac{\partial E_k}{\partial m_{12}} \\ \frac{\partial E_k}{\partial \sigma_{12}} \\ \cdot \\ \cdot \\ \frac{\partial E_k}{\partial \sigma_{42}} \end{bmatrix} = \begin{bmatrix} \frac{2(\text{Join}(x_k) - s_k) \partial \text{Join}(x_k)}{\partial m_{11}} \\ \frac{2(\text{Join}(x_k) - s_k) \partial \text{Join}(x_k)}{\partial \sigma_{11}} \\ \frac{2(\text{Join}(x_k) - s_k) \partial \text{Join}(x_k)}{\partial m_{12}} \\ \frac{2(\text{Join}(x_k) - s_k) \partial \text{Join}(x_k)}{\partial \sigma_{12}} \\ \cdot \\ \cdot \\ \frac{2(\text{Join}(x_k) - s_k) \partial \text{Join}(x_k)}{\partial \sigma_{42}} \end{bmatrix} \quad (3.9)$$

Note that  $\ell$  is a learning rate which can be expressed as

$$\ell = \frac{m}{\sqrt{\sum_i \left( \frac{\partial E_k}{\partial NP_i} \right)^2}} \quad (3.10)$$

$m$  is the step size. The value of  $m$  can be changed to vary the speed of convergence

The tuning procedure can be divided into the following steps:

1. Arbitrary Choice of the initial value of the premise parameters.
2. Use the Least Square method (3.6) to tune the consequent parameters;
3. Use the gradient method (3.8) for  $K$  video sequences in the training set to tune the premise parameters;
4. Repeat steps 2 and step 3, until the difference of the mean square error (3.7) in two iterations is small enough. Here  $E_k(n)$  is mean square error of  $k$ th video sequences in the  $n$ th iterations.  $E_k(n-1)$  is mean square error of  $k$ th video sequences in the  $(n-1)$ th iterations. It is obtained that  $|E_k(n) - E_k(n-1)| < \varepsilon$ .  $\varepsilon$  is a small number and is set 1 in this example.

Below are the parameters of the fuzzy system.

Membership Function	The Mean Value	The Variance Value
$MF_1(x_1)$	$m_{11} = 0.03272$	$\sigma_{11} = 0.03858$
$MF_2(x_1)$	$m_{12} = -0.1834$	$\sigma_{12} = 0.04143$
$MF_3(x_2)$	$m_{21} = -0.1834$	$\sigma_{21} = 0.04143$
$MF_4(x_2)$	$m_{22} = -0.04484$	$\sigma_{22} = 0.2418$
$MF_5(x_3)$	$m_{31} = -0.4799$	$\sigma_{31} = 0.1333$
$MF_6(x_3)$	$m_{32} = -0.4267$	$\sigma_{32} = 0.2202$
$MF_7(x_4)$	$m_{41} = -0.7806$	$\sigma_{41} = 0.2202$
$MF_8(x_4)$	$m_{42} = -0.5129$	$\sigma_{42} = 0.09383$

Table 3.3 Nonlinear Parameters of the Fuzzy Algorithm (3.3)

The consequent matrix in (3.5)

$$\mathbf{L} = \begin{bmatrix} 74.2 & 22.56 & -56.71 & 8.361 & -2.887 \\ 7.418 & 36.33 & -10.38 & 66.43 & 94.63 \end{bmatrix}$$

### 3.3.3 Example of Implementing the Fuzzy Metric

In this section, to illustrate the implication of this algorithm (3.9) clearly, let assume the following input:

$$x_1=0.019, x_2=-0.090, x_3=-0.281, x_4=-0.625$$

The membership function has the following values:

$MF_1(x_1)$	0.9387
$MF_2(x_1)$	$5.3989 \times 10^{-5}$
$MF_3(x_2)$	0.0788
$MF_4(x_2)$	0.987
$MF_5(x_3)$	0.3285
$MF_6(x_3)$	0.7115
$MF_7(x_4)$	0.7501
$MF_8(x_4)$	0.4898

Table 3.4 the value of The Membership Functions

The premise parts of the fuzzy algorithm have the following values:

$B_1(x)$	0.0182
$B_2(x)$	$1.848 \times 10^{-5}$ ;

Table 3.5 the value of the Premise Parts of the Fuzzy Algorithm

From Table 3.5, it is found that the value of  $B_1(x)$  is bigger than the value of  $B_2(x)$ . It shows that it is more appropriate to describe the QP's ( $x_1, x_2, x_3, x_4$ ) using the first rule ( $x_1$  is big and  $x_2$  is small and  $x_3$  is small and  $x_4$  is small).

The consequent parts of the fuzzy algorithm have the following values

$Join_1(x)$	7.2023
$Join_2(x)$	52.8903

Table 3.6 the values of the consequent part of the fuzzy algorithm

Finally, the output of the fuzzy algorithm (3.3) is:

$$\text{Join}(\mathbf{x}) = \text{Join}_1(\mathbf{x}) \times B_1(\mathbf{x}) / (B_1(\mathbf{x}) + B_2(\mathbf{x})) + \text{Join}_2(\mathbf{x}) \times B_2(\mathbf{x}) / (B_1(\mathbf{x}) + B_2(\mathbf{x})) = 7.2485$$

Below are the figures of Membership Functions for  $x_1$ ,  $x_2$ ,  $x_3$ ,  $x_4$ , and the input points represent the value of the QP's used in this example. The value of these membership functions represent the degree to describe the values of the QP's using the linguistic values-'big' or 'small'.

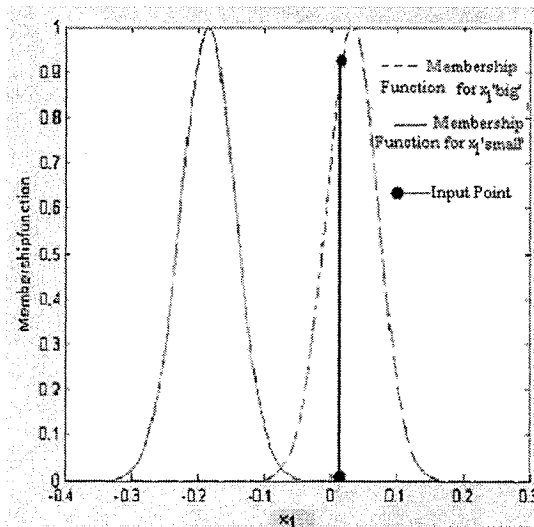


Figure 3.3 Membership Function of  $x_1$

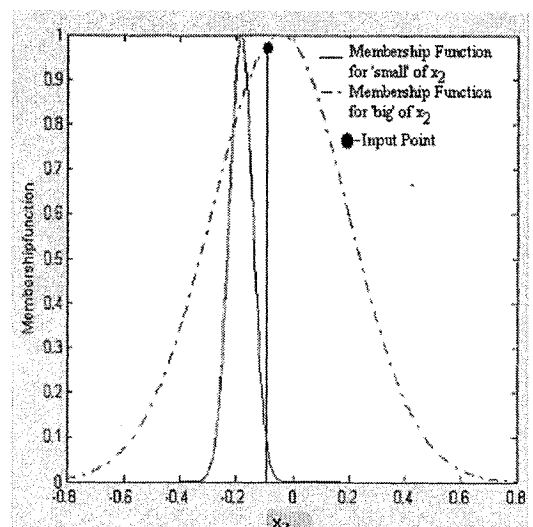


Figure 3.4 Membership Function of  $x_2$

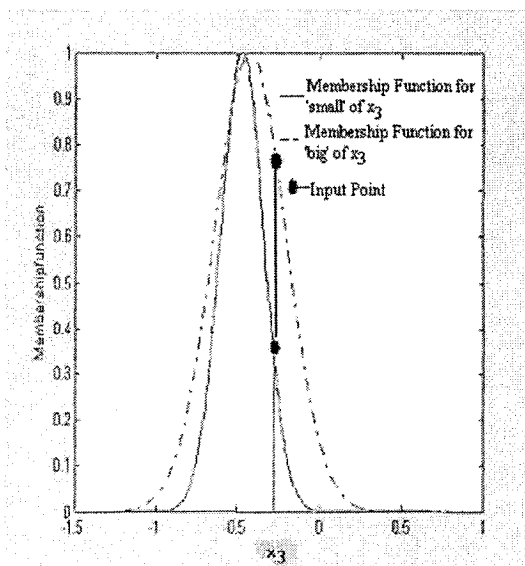


Figure 3.5 Membership Function of  $x_3$

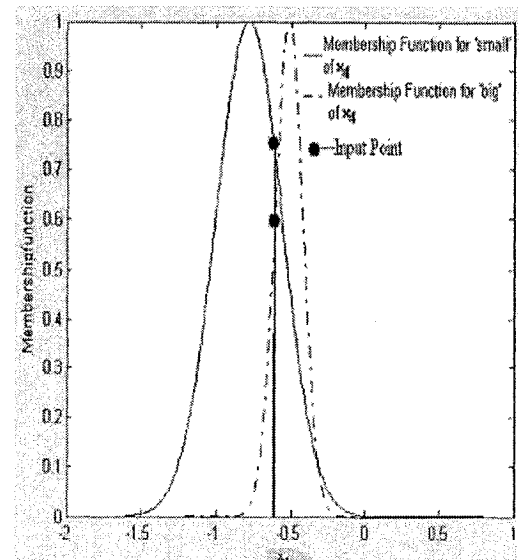


Figure 3.6 Membership Function of  $x_4$

From Figure 3.3 and Figure 3.6, it is found that the maximum value of the membership functions '1' and the minimum value of the membership function is '0'. Here, '1' means it is the most accurate to describe the values of the QP's using linguistic values. '0' means it is the least accurate to describe the values of the QP's using linguistic values.

### 3.3 Fuzzy Model Using Fuzzy Clustering Method (8 QP's)

In this section, 8 QP's (Table 2.4) (flat and non-flat) combining Wolf and Bistahwi's method are used as the input to the fuzzy system. This fuzzy system will be used to simulate the subjective goodness. QP's are represented as  $x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8$ . Define a vector  $\mathbf{x}=[x_1 \ x_2 \ x_3 \ x_4 \ x_5 \ x_6 \ x_7 \ x_8]$  that contains all the QP's. Every QP's is described using two linguistic variable-'small' and 'large'. And as did in the Section 3.2, Fuzzy Clustering [36] is used to select the rules.

#### 3.3.1 Proposed Fuzzy Metric

After using Fuzzy C-Means, the proposed fuzzy algorithm is as follows:

If the  $x_1$  is small and  $x_2$  is small and  $x_3$  is small and  $x_4$  is small and  $x_5$  is small and  $x_6$  is small and  $x_7$  is small and  $x_8$  is small, Then

Join<sub>1</sub>( $\mathbf{x}$ )

If the  $x_1$  is big and  $x_2$  is big and  $x_3$  is big and  $x_4$  is big and  $x_5$  is big and  $x_7$  is big and  $x_8$  is big, Then

Join<sub>2</sub>( $\mathbf{x}$ )

This fuzzy metric can be divided into two parts-premise part (if section) and consequent part (then section) and can be described in mathematically as follow forms:

$$\text{Join}(\mathbf{x}) = \frac{\sum_{i=1}^2 B_i(\mathbf{x}) \text{join}_i(\mathbf{x})}{\sum_{i=1}^2 B_i(\mathbf{x})} = \sum_{i=1}^2 \text{Normal}_i(\mathbf{x}) \text{join}_i(\mathbf{x}) \quad (3.11)$$

Below figure is used to express (3.11)

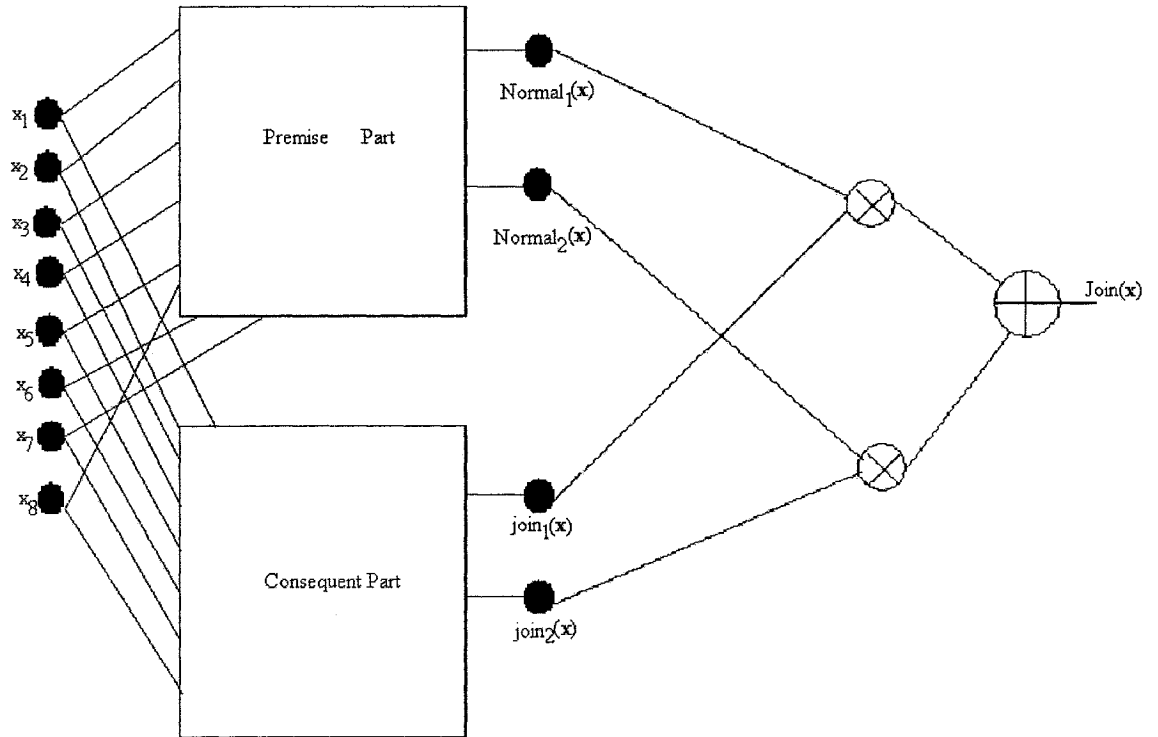


Figure 3.7 Figure of the Fuzzy Metric (3.11)

where  $\text{join}_i(\mathbf{x})$  ( $i=1,2$ ) is the output of the consequent part and  $\text{Normal}_i(\mathbf{x})$  ( $i=1,2$ ) is the normalized output of the premise part. The Sum of the product of the output of the premise part and the consequent part ( $\text{join}_i(\mathbf{x}) \times \text{Normal}_i(\mathbf{x})$ ) is the output of the fuzzy system- $\text{Join}(\mathbf{x})$ .

### 3.3.1.1 The Premise Part of the Fuzzy Metric

For the premise part, the following membership functions are used.

Quality Primitives	Linguistic Values	Membership Function
$x_1$	'big'	$mf_1(x_1) = e^{\frac{-(x_1 - m_{11})^2}{2(\sigma_{11})^2}}$
$x_1$	'small'	$mf_2(x_1) = e^{\frac{-(x_1 - m_{12})^2}{2(\sigma_{12})^2}}$
$x_2$	'big'	$mf_3(x_2) = e^{\frac{-(x_2 - m_{21})^2}{2(\sigma_{21})^2}}$
$x_2$	'small'	$mf_4(x_2) = e^{\frac{-(x_2 - m_{22})^2}{2(\sigma_{22})^2}}$
$x_3$	'small'	$mf_5(x_3) = e^{\frac{-(x_3 - m_{31})^2}{2(\sigma_{31})^2}}$
$x_3$	'big'	$mf_6(x_3) = e^{\frac{-(x_3 - m_{32})^2}{2(\sigma_{32})^2}}$
$x_4$	'small'	$mf_7(x_4) = e^{\frac{-(x_4 - m_{41})^2}{2(\sigma_{41})^2}}$
$x_4$	'big'	$mf_8(x_4) = e^{\frac{-(x_4 - m_{42})^2}{2(\sigma_{42})^2}}$
$x_5$	'big'	$mf_9(x_5) = e^{\frac{-(x_5 - m_{51})^2}{2(\sigma_{51})^2}}$
$x_5$	'small'	$mf_{10}(x_5) = e^{\frac{-(x_5 - m_{52})^2}{2(\sigma_{52})^2}}$

Quality Primitives	Linguistic Values	Membership Function
$x_6$	'small'	$mf_{11}(x_6) = e^{\frac{-(x_6 - m_{61})^2}{2(\sigma_{61})^2}}$
$x_6$	'big'	$mf_{12}(x_6) = e^{\frac{-(x_6 - m_{62})^2}{2(\sigma_{62})^2}}$
$x_7$	'small'	$mf_{13}(x_7) = e^{\frac{-(x_7 - m_{71})^2}{2(\sigma_{71})^2}}$
$x_7$	'big'	$mf_{14}(x_7) = e^{\frac{-(x_7 - m_{72})^2}{2(\sigma_{72})^2}}$
$x_8$	'small'	$mf_{15}(x_8) = e^{\frac{-(x_8 - m_{81})^2}{2(\sigma_{81})^2}}$
$x_8$	'big'	$mf_{16}(x_8) = e^{\frac{-(x_8 - m_{82})^2}{2(\sigma_{82})^2}}$

Table 3.7 the Definition of the Membership Functions

The premise parts (if section) of the fuzzy system can be expressed as the following figure:

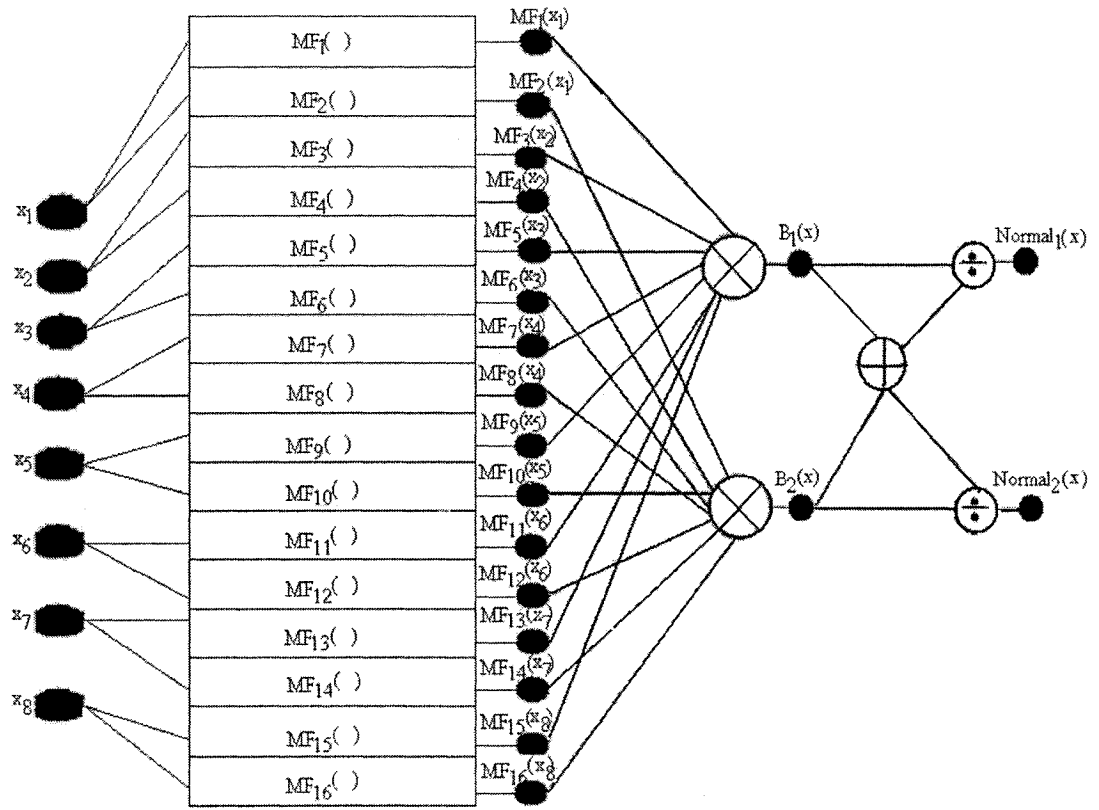


Figure 3.8 Premise Part of the Fuzzy System

The input of this figure is the QP's measuring the quality of the video sequences. The output of this figure is the normalized product of membership functions, for example,

$$\text{Normal}_1(\mathbf{x}) = \frac{B_1(\mathbf{x})}{\sum_{i=1}^2 B_i(\mathbf{x})}, \text{ where } B_i(\mathbf{x}) \text{ can be thought as the product of the memberships}$$

whose labels are dot in the following tables.

	Mf <sub>1</sub> (x <sub>1</sub> )	Mf <sub>2</sub> (x <sub>2</sub> )	Mf <sub>3</sub> (x <sub>3</sub> )	Mf <sub>4</sub> (x <sub>4</sub> )	Mf <sub>5</sub> (x <sub>5</sub> )	Mf <sub>6</sub> (x <sub>6</sub> )	Mf <sub>7</sub> (x <sub>7</sub> )	Mf <sub>8</sub> (x <sub>8</sub> )
B <sub>1</sub> (x)	●		●		●		●	
B <sub>2</sub> (x)		●		●		●		●

	Mf <sub>9</sub> (x <sub>9</sub> )	Mf <sub>10</sub> (x <sub>10</sub> )	Mf <sub>11</sub> (x <sub>11</sub> )	Mf <sub>12</sub> (x <sub>12</sub> )	Mf <sub>13</sub> (x <sub>13</sub> )	Mf <sub>14</sub> (x <sub>14</sub> )	Mf <sub>15</sub> (x <sub>15</sub> )	Mf <sub>16</sub> (x <sub>16</sub> )
B <sub>1</sub> (x)	●		●		●		●	
B <sub>2</sub> (x)		●		●		●		●

Table 3.8 Premise Part of the Fuzzy Metric

### 3.3.1.2 The Consequent Part of the Fuzzy Metric

The consequence parts (then section) of the fuzzy system can be expressed as the following matrix forms:

$$\begin{bmatrix} \text{join}_1(x) \\ \text{join}_2(x) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} & a_{18} & b_1 \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} & a_{28} & b_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ 1 \end{bmatrix} \quad (3.12)$$

The equation (3.13) can be expressed in the Matrix forms:

$$\mathbf{JOIN} = \mathbf{L} \times \mathbf{xI} \quad (3.13)$$

where **JOIN** is the 2×1 vector in equation (3.12), **L** is the 2×9 matrix in equation (3.12), **xI** is 9×1 in equation (3.12).

### 3.3.2 The Tuning Method

The parameters of the fuzzy metric can be divided into two parts (premise parameters and consequent parameters) and for the premise parameters, the tuning method is gradient

decent and for the consequent parameters, the tuning method is linear regression. The tuning method is similar to that described in Section 3.2.2. The result is as follows:

For the nonlinear parameters:

Membership Function	The Mean Value	The Variance Value
$MF_1(x_1)$	$m_{11} = 0.1866$	$\sigma_{11} = 0.2723$
$MF_2(x_1)$	$m_{12} = -0.04961$	$\sigma_{12} = 0.1111$
$MF_3(x_2)$	$m_{21} = -0.1548$	$\sigma_{21} = 0.09201$
$MF_4(x_2)$	$m_{22} = -0.3151$	$\sigma_{22} = 0.2167$
$MF_5(x_3)$	$m_{31} = 0.7041$	$\sigma_{31} = 0.642$
$MF_6(x_3)$	$m_{32} = 1.126$	$\sigma_{32} = 0.612$
$MF_7(x_4)$	$m_{41} = -0.08068$	$\sigma_{41} = 0.1303$
$MF_8(x_4)$	$m_{42} = -0.2101$	$\sigma_{42} = 0.262$
$MF_9(x_5)$	$m_{51} = 0.1088$	$\sigma_{51} = 0.1933$
$MF_{10}(x_5)$	$m_{52} = -0.01677$	$\sigma_{52} = -0.1208$
$MF_{11}(x_6)$	$m_{61} = -0.4082$	$\sigma_{61} = 0.2624$
$MF_{12}(x_6)$	$m_{62} = -0.5152$	$\sigma_{62} = 0.07182$
$MF_{13}(x_7)$	$m_{71} = 1.069$	$\sigma_{71} = 0.6216$
$MF_{14}(x_7)$	$m_{72} = 1.356$	$\sigma_{72} = 0.6193$
$MF_{15}(x_8)$	$m_{81} = -0.7608$	$\sigma_{81} = 0.2639$
$MF_{16}(x_8)$	$m_{82} = -0.4941$	$\sigma_{82} = 0.219$

Table 3.9 Nonlinear Parameters of the Fuzzy Algorithm (3.11)

The consequent matrix in (3.13) is as follows:

$$\mathbf{L} = \begin{bmatrix} 124.1 & 111.7 & 34.72 & 28.51 & 127.5 & -81.79 & -65.61 & -10 & 14.01 \\ 540.9 & 1550 & 226 & 351.8 & 1497 & -823.1 & -258.8 & 29.4 & 108.1 \end{bmatrix}$$

### 3.3.3 Example of Implementing the Fuzzy Metric

In this section, an example is given to compute the fuzzy algorithm (3.11) in detail. Let

assume the QP's is  $x_1=0.019$ ,  $x_2=-0.207$ ,  $x_3=0.365$ ,  $x_4=-0.09$ ,  $x_5=0.034$ ,  $x_6=-0.281$ ,  $x_7=0.498$ ,  $x_8=-0.625$

The membership functions have the following values:

$MF_1(x_1)$	0.8274
$MF_2(x_1)$	0.8264
$MF_3(x_2)$	0.8514
$MF_4(x_2)$	0.8830
$MF_5(x_3)$	0.8659
$MF_6(x_3)$	0.4616
$MF_7(x_4)$	0.9974
$MF_8(x_4)$	0.9003
$MF_9(x_5)$	0.942
$MF_{10}(x_5)$	0.9155
$MF_{11}(x_6)$	0.8891
$MF_{12}(x_6)$	0.0108
$MF_{13}(x_7)$	0.6558
$MF_{14}(x_7)$	0.383
$MF_{15}(x_8)$	0.876
$MF_{16}(x_8)$	0.8363

Table 3.10 the value of the Membership Functions

The premise part of the fuzzy algorithm has the following values:

$B_1(\mathbf{x})$	0.2928
$B_2(\mathbf{x})$	0.001

Table 3.11 the value of the Premise Part of the Fuzzy Algorithm

The consequent parts of the fuzzy algorithm have the following values

$\text{join}_1(\mathbf{x})$	4.2471
$\text{join}_2(\mathbf{x})$	-17.72

Table 3.12 the values of the Consequent Part of the Fuzzy Algorithm

Finally, the output of the fuzzy algorithm (3.15) has the following values:

$$\text{Join}(\mathbf{x}) = \text{join}_1(\mathbf{x}) \times B_1(\mathbf{x}) / (B_1(\mathbf{x}) + B_2(\mathbf{x})) + \text{join}_2(\mathbf{x}) \times B_2(\mathbf{x}) / (B_1(\mathbf{x}) + B_2(\mathbf{x})) = 4.1845$$

Below are the figures of the membership functions in this fuzzy algorithm, and the input points represent the values of the QP's used in this example. The value of these membership functions represent the degree to describe the values of the QP's using the linguistic values- 'big' or 'small'.

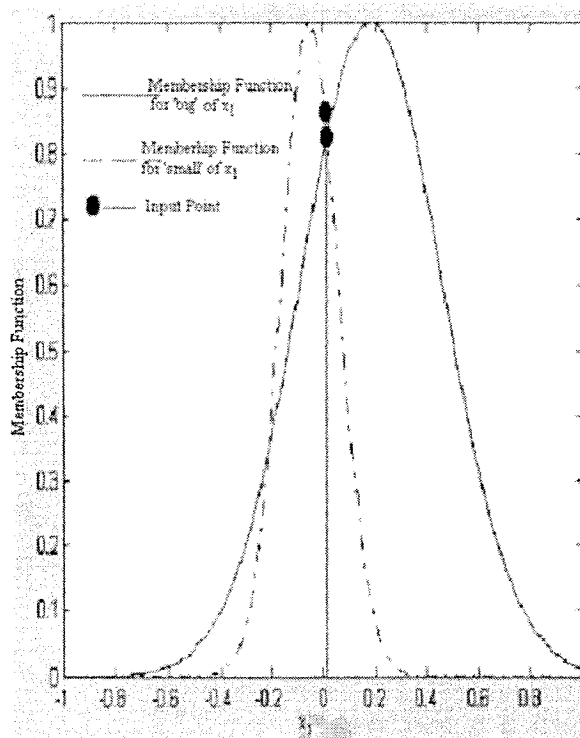


Figure 3.9 Membership Function of  $x_1$

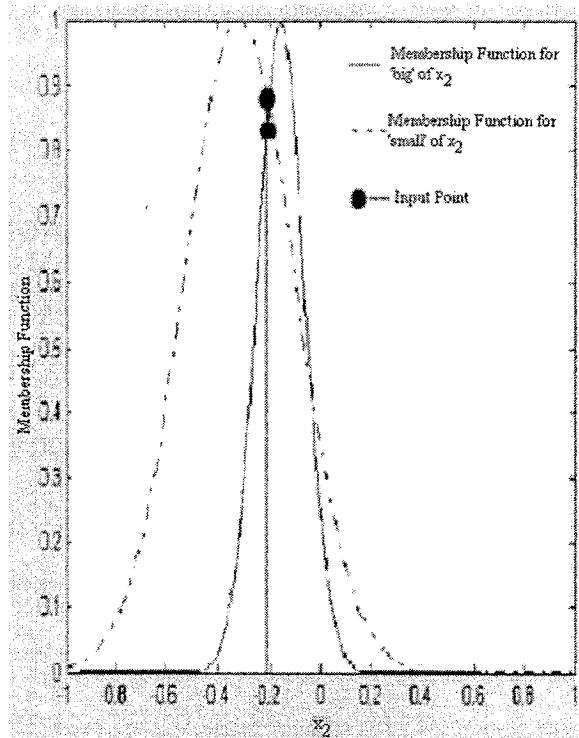


Figure 3.10 Membership Function of  $x_2$

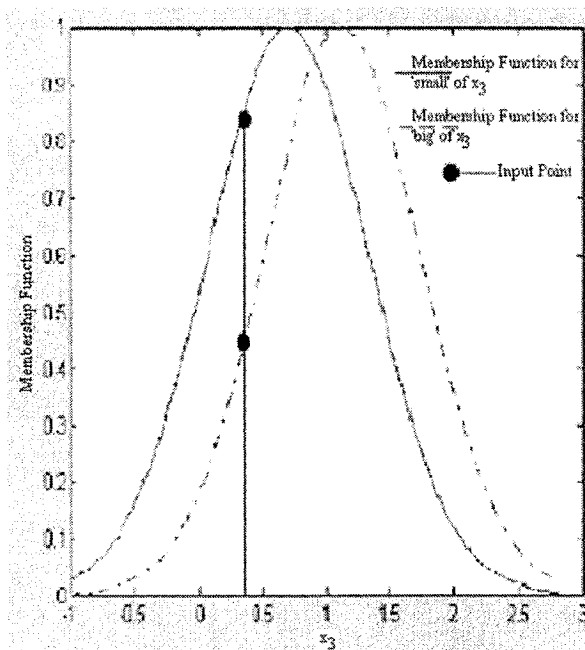


Figure 3.11 Membership Function of  $x_3$

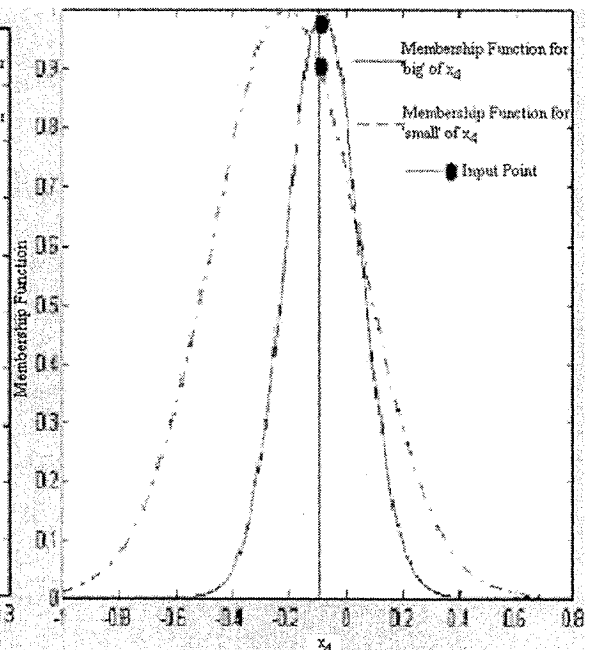


Figure 3.12 Membership Function of  $x_4$

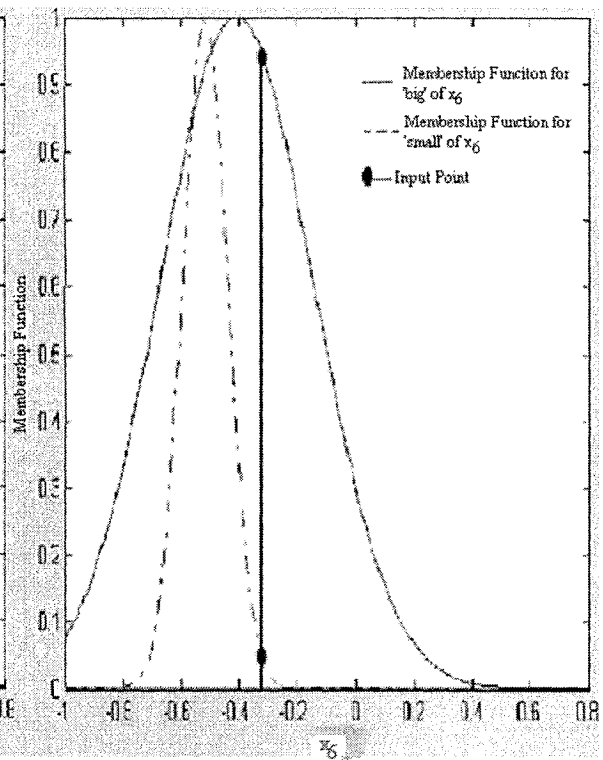
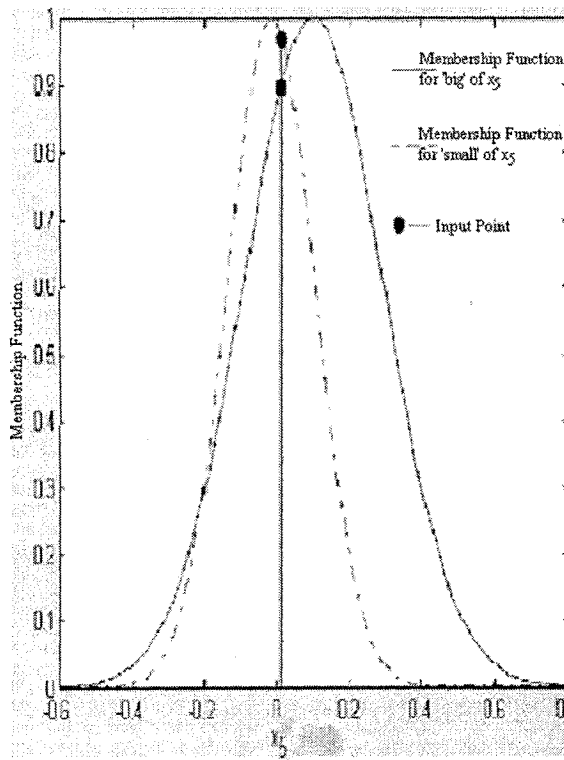


Figure 3.13 Membership Function of  $x_5$       Figure 3.14 Membership Function of  $x_6$

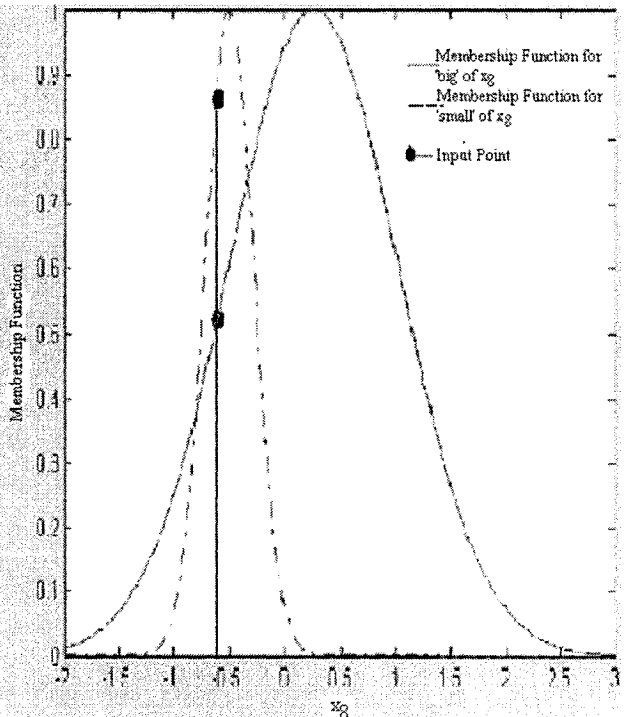
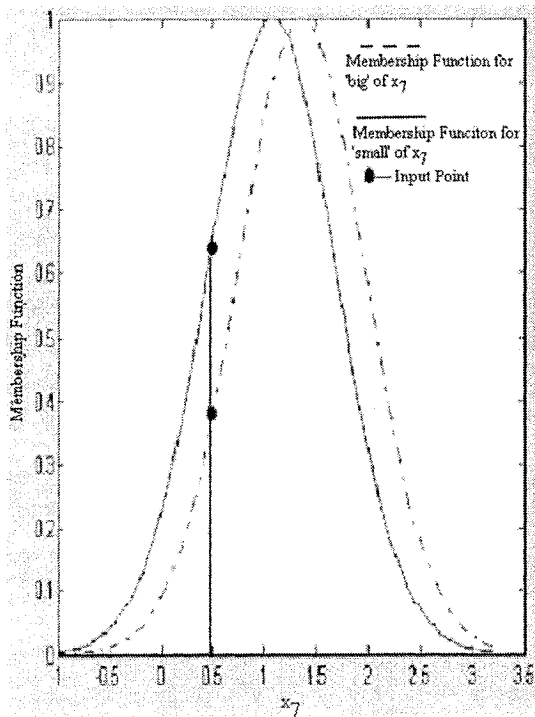


Figure 3.15 Membership Function of  $x_7$       Figure 3.16 Membership Function of  $x_8$

From Figure 3.9 and Figure 3.16, it is also found that the maximum value of the membership functions '1' and the minimum value of the membership function is '0'. Here, '1' means it is the most accurate to describe the values of the QP's using linguistic values. '0' means it is the least accurate to describe the values of the QP's using linguistic values.

### 3.4 Fuzzy Model Using 12 QP's

For the twelve QP's (table 2.5) which are represented as  $x_i$  ( $i=1..12$ ), it is difficult to tune the parameters of the fuzzy metric used to simulate the subjective ratings. To solve this problem, a hybrid algorithm combining the linear mapping and fuzzy logic is used in this section.

#### 3.4.1 Proposed Fuzzy Metric

The hybrid algorithm has the following forms:

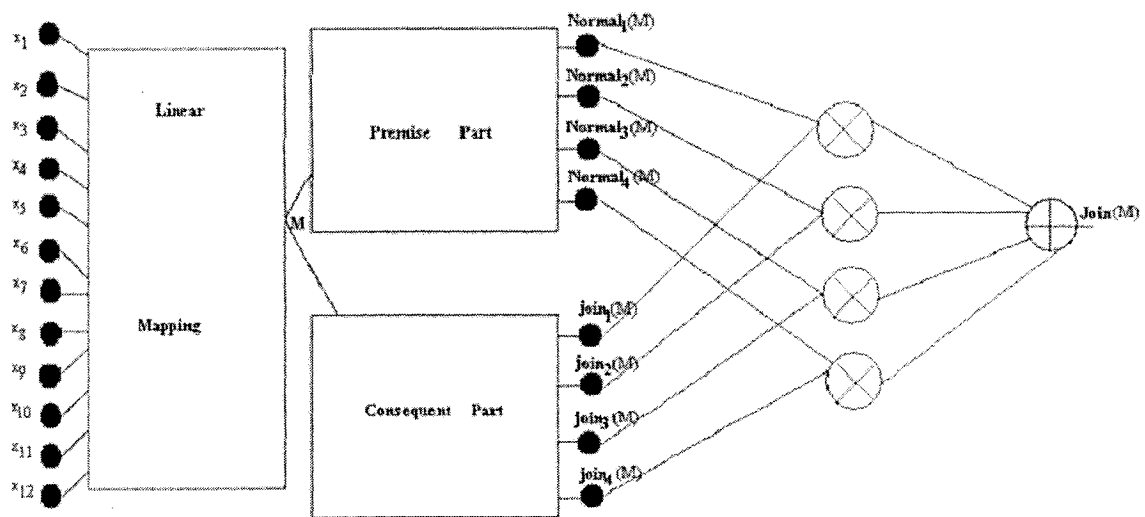


Figure 3.17 Figure of the Hybrid Algorithm

At first, the linear mapping method is used to get the input of the fuzzy system. The parameters of linear mapping are obtained by using the linear regression [28] [34]. Then the fuzzy method is used for the nonlinear mapping to get the final result. This method will give a simpler algorithm comparing with other algorithm (3.3, 3.11).

Step 1:

Let assume the output of the linear mapping is M, then it can be obtained:

$$M = \sum_{i=1}^{12} REG_i x_i$$

$REG_i$   $i=1..12$  is linear parameters produced by linear regression.

Step 2:

If M is very small

Then  $join_1(M)$

If M is small

Then  $join_2(M)$

If M is big

Then  $join_3(M)$

If M is very big

Then  $join_4(M)$

The second part (fuzzy algorithm) can be expressed in mathematically as follows:

$$Join(M) = \frac{\sum_{i=1}^4 MF_i(M) join_i(M)}{\sum_{i=1}^4 MF_i(M)} = \sum_{i=1}^4 Normal_i(M) join_i(M) \quad (3.14)$$

### 3.4.1.1 The Premise Part of the Fuzzy Algorithm

Because there is only one parameter (M) as the input to the fuzzy algorithm (3.14), the premise part of this fuzzy algorithm is simple. The following membership functions are used:

Linguistic Variables	Membership Function
'very small'	$mf_1(M) = e^{\frac{-(M-m_{11})^2}{2(\sigma_{11})^2}}$
'small'	$mf_2(M) = e^{\frac{-(M-m_{12})^2}{2(\sigma_{12})^2}}$
'big'	$mf_3(M) = e^{\frac{-(M-m_{13})^2}{2(\sigma_{13})^2}}$
'very big'	$mf_4(M) = e^{\frac{-(M-m_{14})^2}{2(\sigma_{14})^2}}$

Table 3.13 the Membership Function of the Fuzzy Algorithm

The following figure is used to express the premise part of the fuzzy algorithm

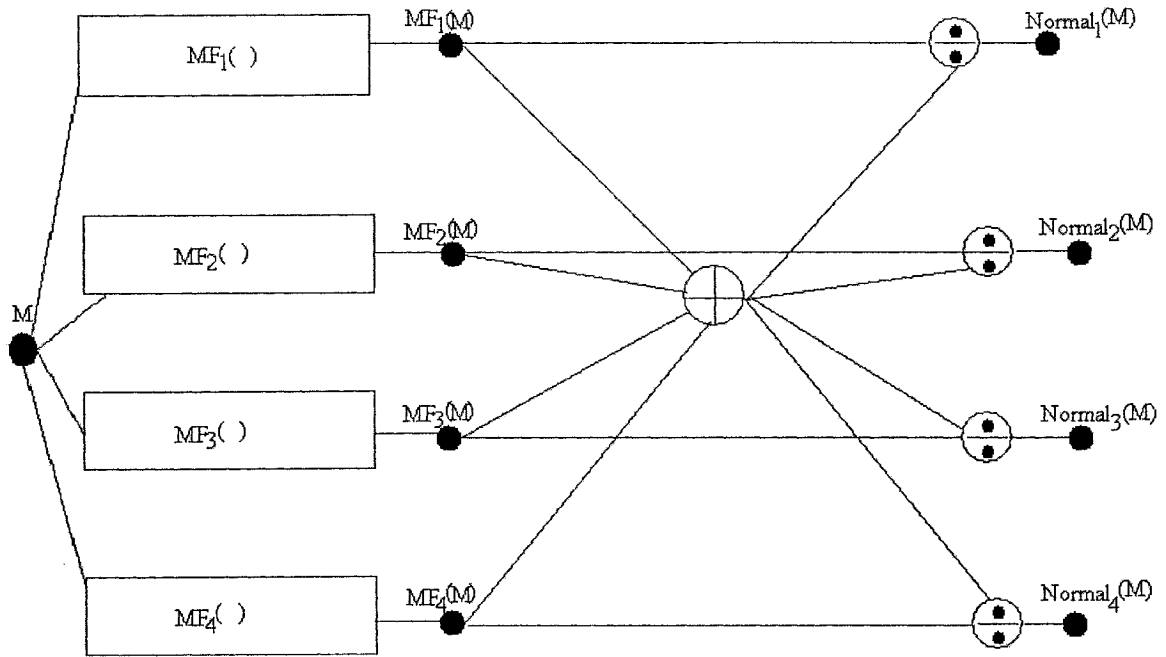


Figure 3.18 the Premise Part of the Fuzzy Algorithm (3.14)

The input of this figure is the Middle parameter (M) (The output of the linear mapping).

The output of this figure is the normalized membership function in each rule, for example

$$Normal_1(M) = \frac{MF_1(M)}{\sum_{i=1}^4 MF_i(M)}.$$

#### 3.4.1.2 The Consequent Part of the Fuzzy Algorithm

The consequence parts (then section) of the fuzzy system can be expressed as the following matrix forms:

$$\begin{bmatrix} Join_1(M) \\ Join_2(M) \\ Join_3(M) \\ Join_4(M) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} \begin{bmatrix} M \\ 1 \end{bmatrix} \quad (3.15)$$

The equation (3.15) can be expressed in the Matrix forms:

$$JOIN = L \times xI \quad (3.16)$$

where **JOIN** is the  $4 \times 1$  vector in equation (3.15), **L** is the  $4 \times 2$  matrix in equation (3.15), **xl** is  $2 \times 1$  in (3.15).

### 3.4.2 The Tuning Method of the Hybrid Algorithm.

The Tuning Method of the Hybrid Algorithm can be divided into two steps. The first step is linear regression in order to get the linear parameters. The second step is the method used to tune the parameters of the fuzzy algorithm (3.14). The tuning method of the fuzzy algorithm is similar to that proposed in the Section 3.2.

The tuning result is as follows:

For the linear mapping parameters:

REG <sub>1</sub>	91.2728
REG <sub>2</sub>	3.9207
REG <sub>3</sub>	-16.3810
REG <sub>4</sub>	35.0267
REG <sub>5</sub>	292.74
REG <sub>6</sub>	-60.12
REG <sub>7</sub>	12.8
REG <sub>8</sub>	32.22
REG <sub>9</sub>	89.36
REG <sub>10</sub>	-7.02
REG <sub>11</sub>	-37.10
REG <sub>12</sub>	-20.23

Table 3.14 the linear mapping parameters of the hybrid algorithm

For the nonlinear parameters of the fuzzy algorithm (3.13)

Membership Function	The Mean Value	The Variance Value
$MF_1(M)$	$m_{11} = 1.969$	$\sigma_{11} = 9.887$
$MF_2(M)$	$m_{12} = 10.62$	$\sigma_{12} = 9.73$
$MF_3(M)$	$m_{13} = 47.04$	$\sigma_{13} = 1.724$
$MF_4(M)$	$m_{14} = 73.79$	$\sigma_{14} = 3.056$

Table 3.15 Nonlinear Parameters of the Fuzzy Algorithm

The consequent matrix in (3.15) is as follows:

$$\mathbf{L} = \begin{bmatrix} 1.842 & -40.58 \\ 2.686 & -121 \\ 3.313 & -30.45 \\ 5.141 & -50.12 \end{bmatrix}$$

Although the input to the hybrid algorithm is complex (12 spatial QP's), the tuning process is easy comparing other algorithms. Because the human visual system weights the video artifact in a nonlinear way, nonlinear mapping (fuzzy algorithm) should be used to simulate the subjective ratings. But the tuning of the parameters in the fuzzy algorithm becomes more complex with the increase of the number of the input variables. In this algorithm, the fuzzy method is used to finish the nonlinear mapping that simulates the subjective ratings and the linear mapping is used to reduce the number of the input variable for the nonlinear mapping. This algorithm can be thought of the main contribution of this thesis.

### 3.5 Simulation Result

In this section, simulation results for the fuzzy algorithms are given. In the Section 3.2, 4 QP's are used as the input to the fuzzy algorithm. It is called four quality primitives fuzzy algorithms(4QPFA). In the Section 3.3, 8 QP's are used as the input to the fuzzy algorithm. It is called eight quality primitives fuzzy algorithms(8QPFA). In the Section 3.4, 12 QP's are used as the input to the fuzzy algorithm. It is called twelve quality primitives fuzzy algorithms(12QPFA). The evaluation standard is the mean square error between the results predicted by the algorithms in question and the subjective rating of human beings. As mentioned in chapter 2, the experimental video sequences contain 12 video sequences. Each video sequences is compressed using 7 different Q factors. The training set includes Autumn Leaves, Sailboat, Flower Garden, Mobile&Calendar, Table Tennis, Bette Pas Bette, Susie, Ferris Wheel and the non training set includes Birches, Football, Horseback Riding, Tempete. They are all scored subjectively by humans under the supervision of the Communications Research Centre in Ottawa, Canada. The detail data of the subjective ratings can refer to Table 2.2. The simulation result of linear mapping algorithm used in many previous papers [28][34] is also compared with the simulation result of the fuzzy algorithms (4QPFA, 8QPFA, 12QPFA) to show the goodness of the fuzzy algorithms.

The raw simulation result of the four quality primitives fuzzy algorithm (4QPFA) is given in the Table 3.16 and Table 3.17 for the training set and the non training set. Linear mapping algorithm (LMA) [28, 34] using the same QP's is also computed to show the

goodness of the fuzzy algorithm. Table 3.18 is the summary of the simulation result of the 4QPTA and LMA. It is found that the MSE between 4QPFA and the subjective ratings is smaller than that between the LMA and the subjective rating. It shows that the 4QPFA is better than the LMA [28] [34].

Training Video Clips	Q factor	Subjective Rating	4QPFA (4QP's)	LMA (4QP)
Autumn Leaves	12	2.2	7.6791	5.3576
	16	12	11.8241	10.9302
	24	17.8	21.4656	23.0592
	32	39.7	35.603	33.9674
	36	51.3	48.3765	39.4926
	40	54.3	50.5032	44.3686
Sailboat	48	57.2	53.4056	51.8037
	12	5.3	4.2175	-0.4311
	16	9.3	6.8656	3.5723
	24	6.5	13.3491	12.7366
	32	10.8	20.0859	22.0971
	36	14.6	24.9671	27.7183
Flower Garden	40	26.1	31.978	32.9485
	48	25.8	36.458	41.3257
	12	8.5	1.6648	-3.6564
	16	8	3.8034	-0.3008
	24	10.7	8.88	7.3469
	32	22.7	14.586	15.4465
Mobile& Calendar	36	24.1	18.3881	20.2397
	40	30.7	21.0629	23.7546
	48	35.1	27.002	30.7212
	12	0	2.7842	-2.0618
	16	0.5	5.141	2.024
	24	11.7	11.8234	13.1789
Table Tennis	32	9.2	21.179	26.2755
	36	29.8	23.2242	31.8001
	40	29.5	23.8911	37.3095
	48	33.6	29.5698	46.4717
	12	1.5	8.7104	6.9682
	16	7.2	13.0985	13.0673
Bette Pas Bette	24	10.2	20.3948	22.8967
	32	19.6	26.8533	31.1595
	36	31.1	31.6599	34.7063
	40	38.7	40.2918	38.4116
	48	36.9	43.1254	43.9466
	12	1.3	2.7858	0.9563
Ferris Wheel	16	8.7	6.9042	6.7837
	24	18.6	13.449	16.9233
	32	22.2	32.3348	26.1562
	36	41.3	39.6267	31.7945
	40	49.2	44.0153	38.3166
	48	46.6	48.5494	47.6569
Susie	12	4.6	-2.1749	-5.8125
	16	7	-0.2314	-2.3864
	24	9.3	6.8867	7.6889
	32	21.5	15.6325	19.2595
	36	35.5	28.9433	28.2558
	40	37.3	33.8895	33.6403
	48	41.6	38.7737	43.8719
	12	7.2	7.5102	5.5006
	16	15.1	11.4027	10.8457
	24	11.1	19.8095	22.3056
	32	31.8	29.9424	34.7172
	36	53	48.99	42.9894
	40	68	63.843	49.7104
	48	70.4	72.7711	61.8353

Table 3.16 the Simulation Result calculated from the training set of the 4QPFA (4QP)

Non-Training Video Clips	Q factor	Subjective Rating	4QPFA (4QP's)	LMA (4QP)
Birches	12	4.5	2.2111	-3.8092
	16	6.5	4.0334	-0.8586
	24	7.5	8.6147	6.0992
	32	8.3	12.246	11.6058
	36	8.4	14.0875	14.2179
	40	21.3	15.5627	16.3429
Football	48	22.6	17.6997	19.4658
	12	1.5	3.3229	2.8107
	16	10.2	8.8695	10.0688
	24	9.4	22.214	23.8703
	32	37.3	44.8374	35.2264
	36	36.8	47.5331	40.6923
Horseback Riding	40	56.3	49.2241	45.4566
	48	46.6	54.8299	55.2207
	12	3.2	4.8684	1.5254
	16	10.6	8.3327	6.4894
	24	19.4	17.0874	17.8354
	32	28.8	24.9144	27.9326
Tempete	36	50	29.668	32.5825
	40	52.3	42.9802	37.0933
	48	55.5	49.9054	43.002
	12	0.6	3.9609	-0.8237
	16	2.5	6.8905	3.6704
	24	12.1	13.1181	12.77
	32	11.1	18.9767	21.0866
	36	23.3	22.1932	25.2584
	40	23.7	26.3583	29.712
	48	29.8	32.994	37.6546

Table 3.17 the Simulation Result calculated from the non training set of the 4QPFA (4QP)

	The Mean Square Error between the 4QPFA (4QP's) and the subjective rating	The Mean Square Error between the LMA (4 QP's) and the subjective rating
Training Set	33.1433	62.35
Non Training Set	44.5922	55.20
Total	36.9596	59.97

Table 3.18 Simulation Result of 4QPFA (4 QP's)

Figure 3.20 is the figure of the scatter plot of the output of 4QPFA versus the subjective rating values and Figure 3.19 is the scatter plot of the output of LMA and subjective rating values.

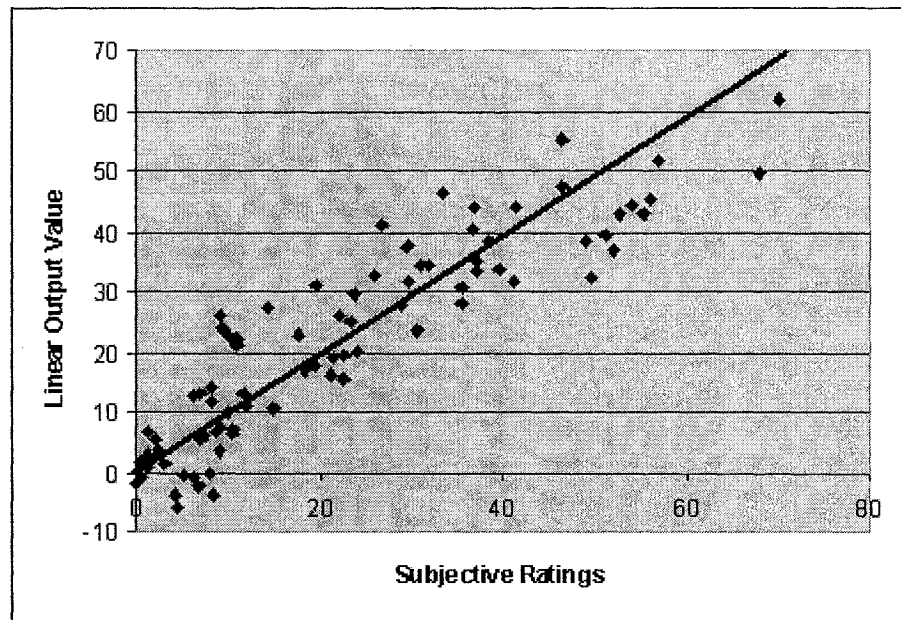


Figure 3.19 Scatter Plot of LMA versus Subjective Ratings (4QP's)

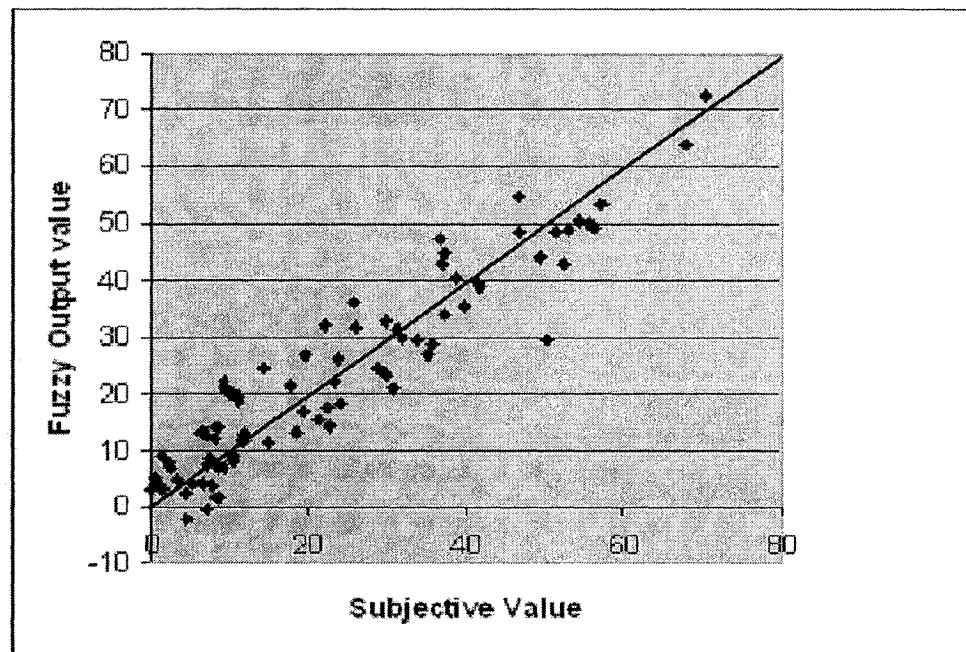


Figure 3.20 Scatter Plot of the 4QPFA versus Subjective Ratings (4 QP's)

It is found that the scatter points in Figure 3.20 are located closer to the plot diagonals than the scatter points in Figure 3.19. It also shows that the 4QPFA is better than the LMA [28,34].

The raw simulation result of the eight quality primitives fuzzy algorithm(8QPFA) is given in the Table 3.19 and Table 3.20 for the training set and the non training set. Linear mapping algorithm (LMA) [28, 34] using the same QP's is also computed to show the goodness of the fuzzy algorithm. Table 3.21 is the summary of the simulation result. It is found that the MSE between 8QPFA and the subjective ratings is smaller than that between the LMA and the subjective rating. It shows that the 8QPFA is better than the LMA [28, 34].

Training Video Clips	Q factor	Subjective Rating	8QPFA (8QP)	LMA (8QP)
Autumn Leaves	12	2.2	4.1845	4.3736
	16	12	9.0987	11.5744
	24	17.8	20.2331	24.482
	32	39.7	38.1773	35.5958
	36	51.3	49.8696	41.8894
	40	54.3	56.8701	48.2408
Sailboat	48	57.2	57.4045	58.9729
	12	5.3	5.3884	-2.206
	16	9.3	6.7871	2.7207
	24	6.5	7.3777	11.5224
	32	10.8	13.0693	19.3997
	36	14.6	18.1541	24.324
Flower Garden	40	26.1	23.0626	29.6095
	48	25.8	28.0018	38.2547
	12	8.5	3.339	-5.4109
	16	8	6.5972	-0.7929
	24	10.7	10.9969	7.35
	32	22.7	17.9779	15.0359
Mobile& Calendar	36	24.1	24.1261	19.1655
	40	30.7	30.0574	22.9071
	48	35.1	35.2774	30.1231
	12	0	3.3499	-2.8568
	16	0.5	5.0817	1.9635
	24	11.7	7.7008	12.3313
Table Tennis	32	9.2	17.0621	23.6066
	36	29.8	23.3861	29.0787
	40	29.5	29.187	34.6952
	48	33.6	35.0575	44.3675
	12	1.5	6.0566	6.9827
	16	7.2	5.6256	14.8113
Bette Pas Bette	24	10.2	8.9075	23.7991
	32	19.6	23.8671	30.6664
	36	31.1	29.588	33.8259
	40	38.7	35.1114	37.238
	48	36.9	38.8647	43.7267
	12	1.3	1.0272	-0.0015
Ferris Wheel	16	8.7	3.8314	6.6413
	24	18.6	14.9848	17.4895
	32	22.2	27.2824	26.0907
	36	41.3	37.4188	31.5818
	40	49.2	46.8579	37.9953
	48	46.6	50.5478	48.5115
Susie	12	4.6	3.3716	-4.6019
	16	7	6.5832	0.5693
	24	9.3	14.4215	12.0286
	32	21.5	23.7938	22.8263
	36	35.5	29.6917	30.8429
	40	37.3	34.6614	37.1135
	48	41.6	41.856	48.8837
	12	7.2	11.6023	6.7404
	16	15.1	12.6634	13.1043
	24	11.1	13.4727	23.1897
	32	31.8	29.8991	33.5437
	36	53	53.1015	40.5564
	40	68	69.3366	47.2837
	48	70.4	69.4786	59.3911

Table 3.19 the Simulation Result calculated from the training set with the 8QPFA (8QP)

Non-Training Video Clips	Q factor	Subjective Rating	4QPFA (8QP's)	LMA (8QP)
Birches	12	4.5	0.8321	-4.8141
	16	6.5	2.8528	-0.8808
	24	7.5	6.9313	8.1023
	32	8.3	11.6952	15.0802
	36	8.4	15.3504	18.2395
	40	21.3	17.7176	20.834
	48	22.6	20.0292	24.5762
Football	12	1.5	4.8938	-3.7922
	16	10.2	5.5873	4.0378
	24	9.4	11.2253	20.1083
	32	37.3	35.4351	32.6444
	36	36.8	42.4986	38.4074
	40	56.3	50.6203	43.5657
	48	46.6	47.2122	53.8365
Horseback Riding	12	3.2	7.3183	0.7109
	16	10.6	6.5934	7.144
	24	19.4	12.8653	21.3871
	32	28.8	36.6583	32.928
	36	50	44.9739	38.0482
	40	52.3	54.1415	42.8424
	48	55.5	53.3365	48.872
Tempete	12	0.6	4.2075	-1.7478
	16	2.5	5.9332	3.6403
	24	12.1	9.7928	13.9053
	32	11.1	17.3235	22.4199
	36	23.3	22.2933	26.5452
	40	23.7	26.6652	30.7391
	48	29.8	29.3363	38.093

Table 3.20 the Simulation Result calculated from the non training set with the 8QPFA (8QP)

	The Mean Square Error between the 8QPFA (8QP's) and the subjective rating	The Mean Square Error between the LMA (8QP's) and the subjective rating
Training Set	9.6753	57.75
Non Training Set	16.4522	46.08
Total	11.9353	53.86

Table 3.21 Simulation Result of 8QPFA (8 QP's)

Figure 3.22 is the figure of the scatter plot of output of 8QPFA versus the subjective rating values and Figure 3.21 is the scatter plot of the output of LMA and subjective rating values.

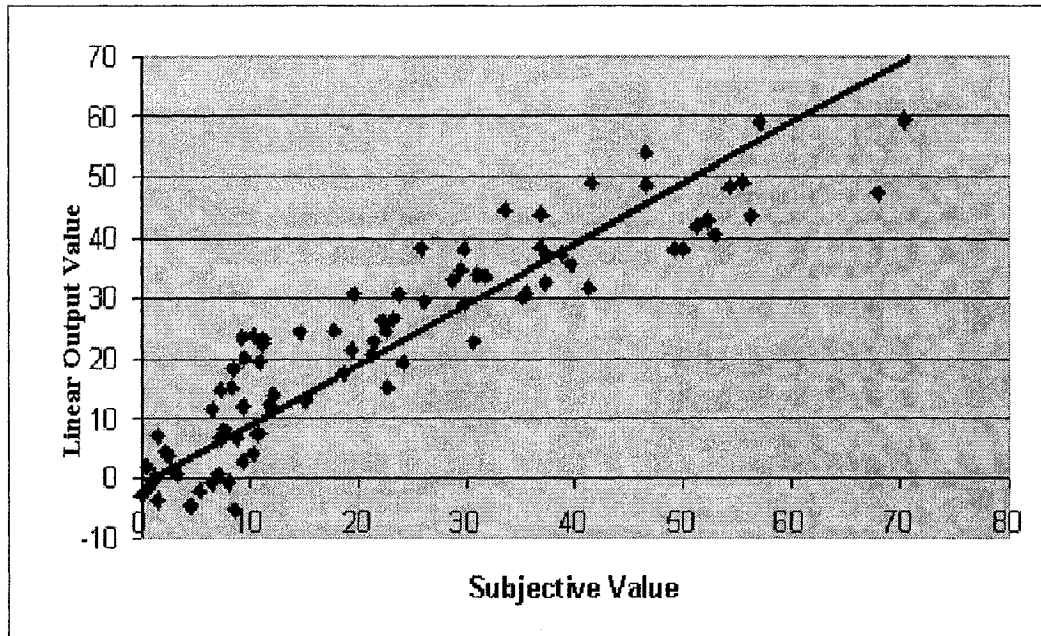


Figure 3.21 Scatter Plot of LMA versus Subjective Ratings(8 QP's)

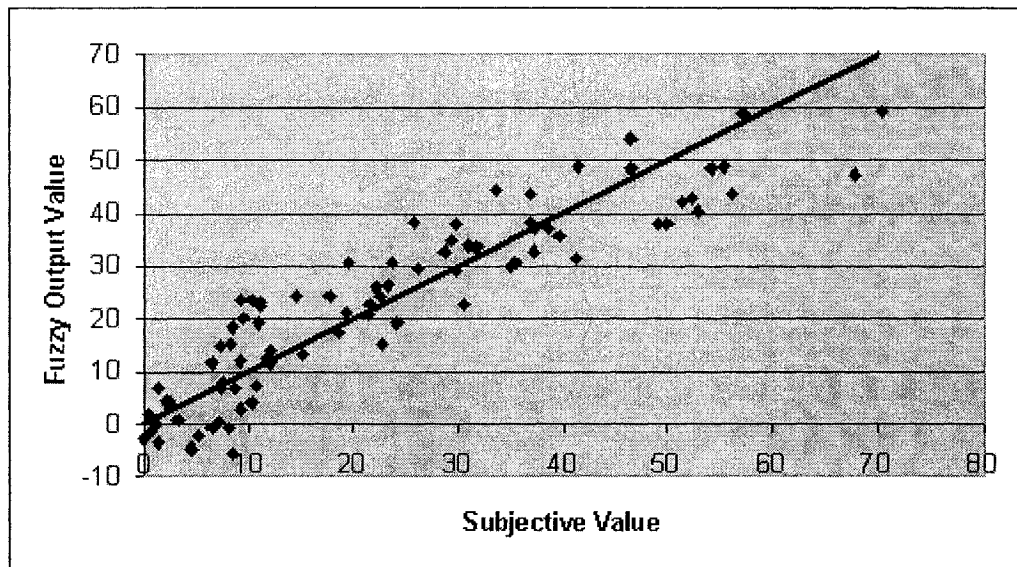


Figure 3.22 Scatter Plot of the 8QPFA( 8 QP's) versus Subjective Ratings

It is found that the scatter points in Figure 3.22 are located closer to the plot diagonals than the scatter points in Figure 3.21. It also shows that the EQPFA is better than the LMA [28] [34].

The raw simulation result of the twelve quality primitives fuzzy algorithm(12QPFA) using the hybrid method is given in the Table 3.22 and Table 3.23 for the training set and the non training set. Linear mapping algorithm (LMA) [28,34] using the same QP's is also computed to show the goodness of the fuzzy algorithm. Table 3.24 is the summary of the simulation result. It is found that the MSE between 12QPFA and the subjective ratings is smaller than that between the LMA and the subjective rating. It shows that the 12QPFA is better than the LMA [28, 34].

Training Video Clips	Q factor	Subjective Rating	12QPFA (12QP)	LMA (12QP)
Autumn Leaves	12	2.2	7.8255	11.5040
	16	12	12.1811	16.7227
	24	17.8	28.4783	29.6898
	32	39.7	38.9244	41.8809
	36	51.3	43.9097	46.9826
	40	54.3	52.7322	51.5138
	48	57.2	57.2384	58.9800
Sailboat	12	5.3	3.1216	-1.6627
	16	9.3	5.5869	2.7038
	24	6.5	7.5879	11.0468
	32	10.8	13.6683	18.7515
	36	14.6	18.2809	22.4247
	40	26.1	23.709	26.0695
	48	25.8	35.4031	32.8722
Flower Garden	12	8.5	4.8441	-0.2408
	16	8	6.0679	3.8463
	24	10.7	8.9127	11.5291
	32	22.7	14.8936	18.5281
	36	24.1	18.0945	21.5186
	40	30.7	21.4338	24.0976
	48	35.1	28.5154	29.0795
Mobile& Calendar	12	0	-1.1058	-5.0653
	16	0.5	3.3357	-1.5595
	24	11.7	6.7503	9.3824
	32	9.2	15.6	20.4805
	36	29.8	23.1835	26.5958
	40	29.5	32.5367	32.1289
	48	33.6	37.5343	41.8217
Table Tennis	12	1.5	6.1356	4.8147
	16	7.2	8.7932	11.4815
	24	10.2	16.9088	19.7422
	32	19.6	23.9749	25.3623
	36	31.1	26.1832	27.4751
	40	38.7	28.2813	28.7975
	48	36.9	38.715	35.3493
Bette Pas Bette	12	1.3	5.7832	2.7647
	16	8.7	6.9388	8.2044
	24	18.6	15.3291	18.6505
	32	22.2	26.8777	27.7223
	36	41.3	34.9993	32.4139
	40	49.2	44.9149	37.4120
	48	46.6	44.615	45.6393
Ferris Wheel	12	4.6	2.8922	-2.2595
	16	7	5.5649	2.0927
	24	9.3	8.5299	11.9968
	32	21.5	21.365	23.9401
	36	35.5	34.6361	32.0651
	40	37.3	44.3256	36.7534
	48	41.6	45.1905	45.3153
Susie	12	7.2	5.7871	5.0305
	16	15.1	7.3558	11.1501
	24	11.1	18.7924	23.2648
	32	31.8	42.6367	37.0182
	36	53	45.6124	47.0864
	40	68	68.0258	54.5937
	48	70.4	70.4115	71.2715

Table 3.22 the Simulation Result calculated from the training set with the 12QPFA (12 QP)

Non-Training Video Clips	Q factor	Subjective Rating	12QPFA (12QP's)	LMA (12QP)
Birches	12	4.5	4.6609	0.6085
	16	6.5	5.9707	4.820
	24	7.5	7.8898	12.2379
	32	8.3	12.4894	19.3377
	36	8.4	15.0284	22.4184
	40	21.3	19.4503	26.2325
Football	48	22.6	26.0362	31.8766
	12	1.5	5.6935	3.6096
	16	10.2	8.3652	12.2203
	24	9.4	28.5663	28.9717
	32	37.3	37.8468	42.2153
	36	36.8	47.2301	47.6044
Horseback Riding	40	56.3	57.323	52.7453
	48	46.6	46.5654	62.3571
	12	3.2	5.8289	4.4153
	16	10.6	7.1998	10.5637
	24	19.4	16.0357	21.3566
	32	28.8	27.612	30.3969
Tempete	36	50	31.8481	33.6525
	40	52.3	36.5532	36.6113
	48	55.5	46.2412	41.7020
	12	0.6	3.9773	-1.2340
	16	2.5	5.9089	3.7167
	24	12.1	9.0026	13.2991
	32	11.1	16.513	21.5229
	36	23.3	21.309	25.0077
	40	23.7	26.2993	28.3137
	48	29.8	38.8357	34.7498

Table 3.23 the Simulation Result calculated from the non-training set with the 12QPFA (12QP)

	The Mean Square Error between the 12QPFA (12QP) and the subjective rating	The Mean Square Error between the LMA (12QP) and the subjective rating
Training Set	25.44	38.48
No Training Set	50.87	65.59
Total	33.92	47.52

Table 3.24 Simulation Results of 12QPFA (12 QP's)

Figure 3.24 is the figure of the scatter plot of the output of TQPFA versus the subjective rating values and Figure 3.23 is the scatter plot of the output of LMA and subjective rating values.

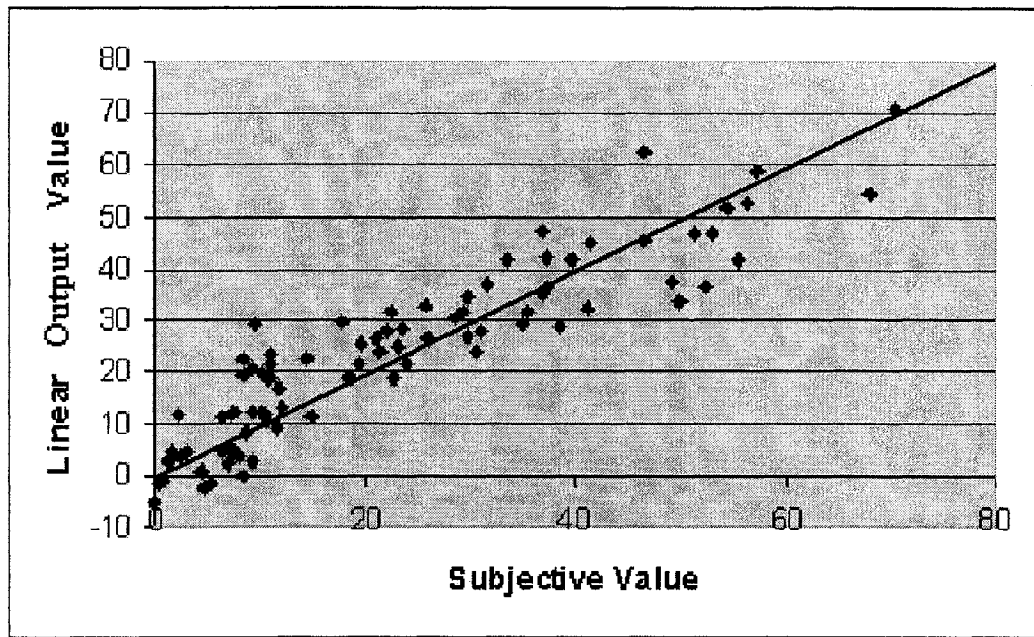


Figure 3.23 Scatter Plot of the LMA versus Subjective Ratings (12 QP's)

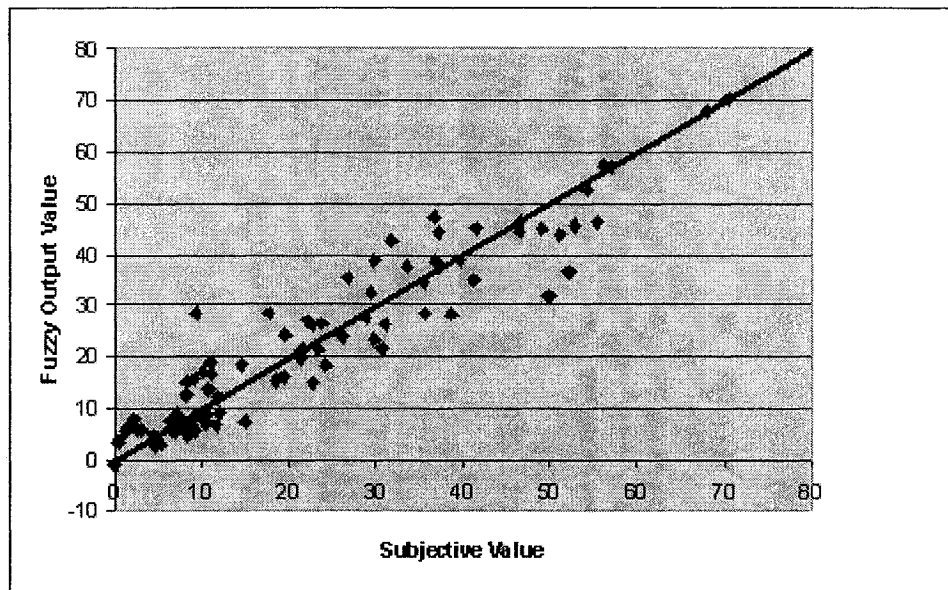


Figure 3.24 Scatter Plot of the 12QPFA versus Subjective Ratings (12 QP's)

It is found that the scatter points in Figure 3.24 are located closer to the plot diagonals than the scatter points in Figure 3.23. It also shows that the 12QPFA is better than the LMA [28, 34].

### 3.6 Conclusion

In this chapter, the Takagi-Sugeno Fuzzy Scheme is used to approximate the complex nonlinear function that simulates the subjective ratings. From the Section 3.5, it is found that the fuzzy algorithms (4QPFA and 8QPFA and 12QPFA) are more accurate in the prediction of the subjective ratings than linear mapping method(LMA)[28][34]. But from the Section 3.2, it is found that the tuning method for the parameters of the fuzzy algorithms is very complex and time consuming. Especially, the initial parameters of the fuzzy algorithms must be chosen arbitrarily. It is main the drawback of the fuzzy algorithms. So further research should be done to provide the solution to this problem.

## Chapter 4

### Flashing Block Artifact

#### 4.1 Introduction

Spatial QP's which represent spatial artifacts of video sequences have been extensively researched [28, 31, 32] but QP's which represent temporal artifacts have seldom been mentioned. In this chapter, a temporal QP which represents the temporal artifacts (flashing block) of the video sequences is described. An automatic measure of the Flashing Block artifacts is introduced, i.e. a Flashing Block QP is introduced. This new QP is then integrated into a fuzzy logic based QS similar to the one in the previous chapter.

This chapter is organized as follows. In Section 4.2 the Block Flashing artifact is described. In Section 4.3, 'Flashing Blocks' Detection Method(FBDM) is proposed. The method is made up of three parts-the temporal method and the spatial frequency method and the over dark and over bright method which are then combined. In Section 4.4 a method to calculate a QP which measures the Block Flashing artifact is described. As in Chapter 3, in Section 4.5, the T-S fuzzy scheme [36] is then used to create a QS which this time includes the new Block Flashing QP. Section 4.6 gives simulation results. Conclusions are found in Section 4.7.

## 4.2 The Definition of the ‘Flashing Blocks’ Artifact

It is well known that lossy compression methods introduce video artifacts. In artifact based quality measurement papers, most ([28, 31, 32]) write about spatial artifacts such as Blocking and Ringing. In this section, a temporal artifact of video sequences – ‘Flashing Blocks’ will be described. It is found that some blocks in the frames of still videos visibly flash when the compression ratio is high. This temporal artifact is called ‘Flashing Blocks’. To illustrate this, one ‘Flashing Blocks’ and one ‘non Flashing Blocks’ in the flat domain and one ‘non Flashing Blocks’ in the non-flat domain are chosen. Their spatial locations are shown in the following figure:

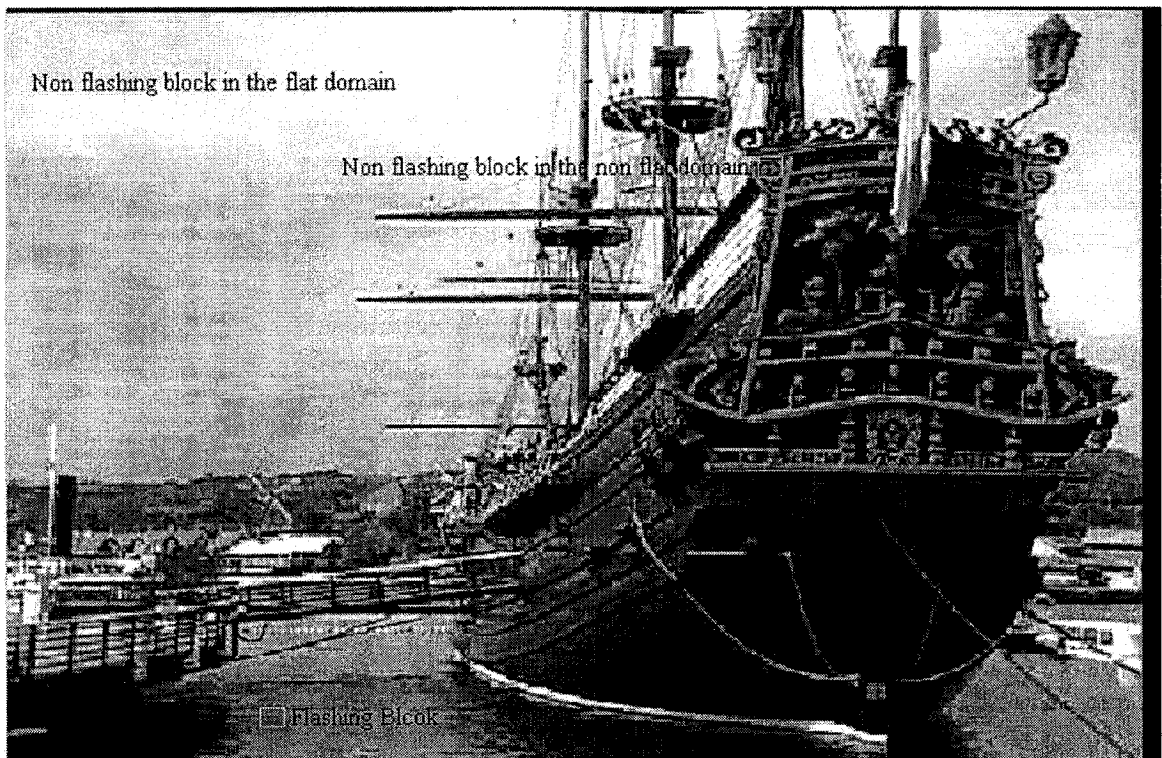


Figure 4.1 Spatial Locations of ‘Flashing Blocks’ and ‘non Flashing Blocks’

The decoded video sequences-sailboat.24.mpeg2 will be used to show the 'Flashing Blocks'. As mentioned in Chapter 2, these clips contain 300 frames. Each clip is (480×720). The decoded video sequences are divided into S-T regions of 8×8×300. There are 5400 S-T regions.

Figure.4.2 is the temporal plot of 'Flashing Blocks'. 300 8×8 blocks in the same 'Flashing Blocks' spatial locations and in different frames of the decoded videos are plotted together to show the temporal change of this 'Flashing Blocks'. As shown in the Figure.4.2, the blocks are plotted in raster order.

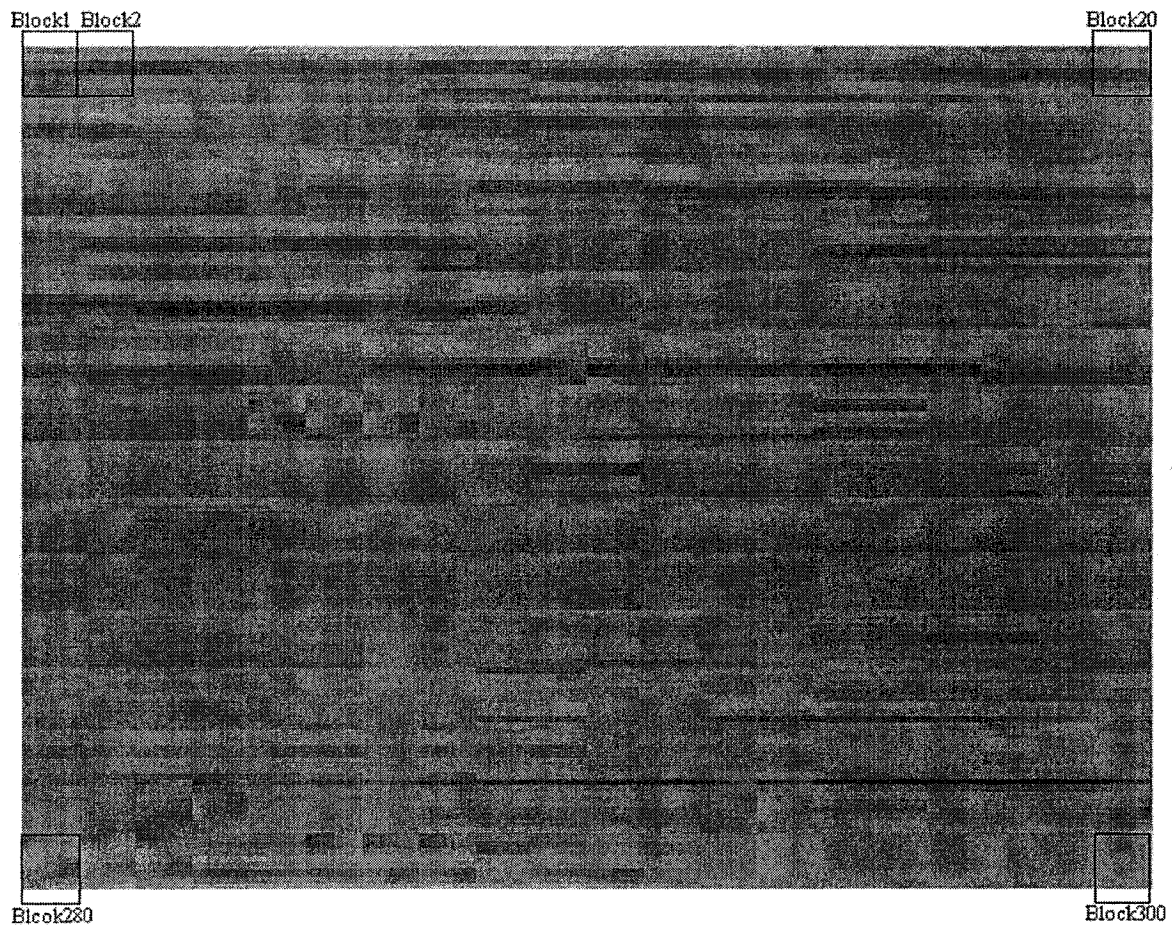


Figure 4.2 the Temporal Plot of the 'Flashing Blocks' (Spatial Domain)

Figure 4.3 is the temporal plots of ‘Flashing Blocks’ in the DCT domain. Each DCT coefficient of the compressed blocks shown in Figure 4.2 are plotted together to show the temporal change of the ‘Flashing Blocks’. In the DCT-Based plots, black means DCT coefficients are zero and white means DCT coefficients are non zeros.

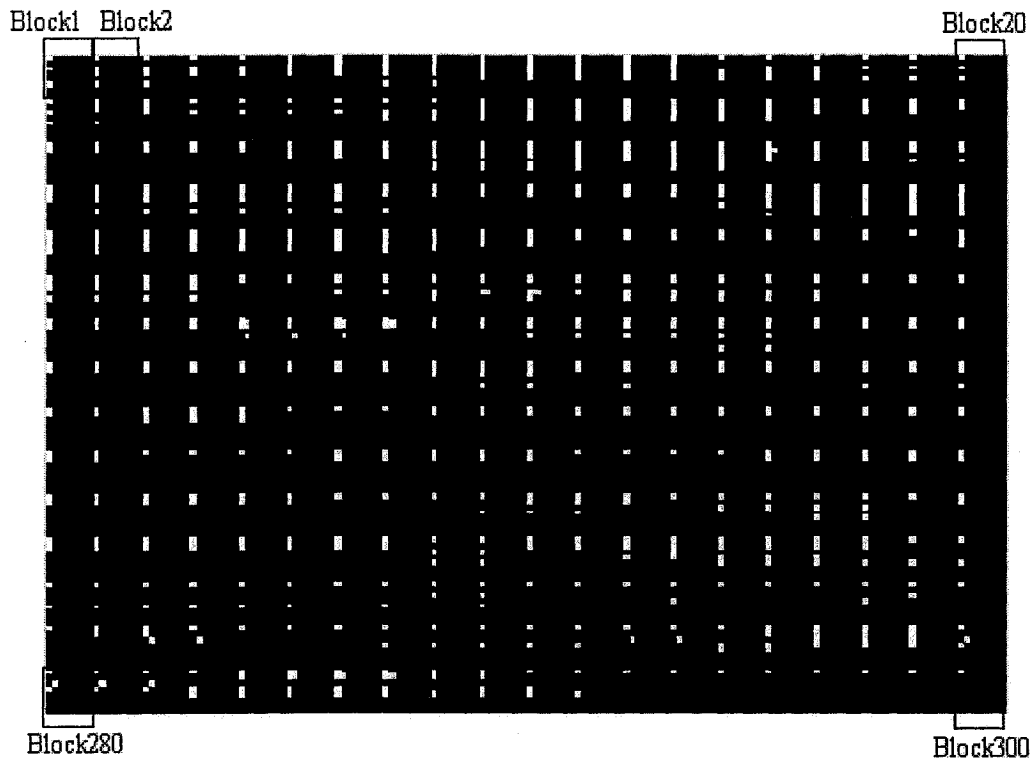


Figure 4.3 the Temporal Plot of ‘Flashing Blocks’ (DCT Domain)

848	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
57	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
-78	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Table 4.1 2D DCT coefficients of the block in the first frame in Figure 4.3

In Table 4.1, the 64 coefficients of the ‘Flashing Blocks’ from the first frame are shown. The coefficient 848 located in the first row and the first column is a DC coefficient, while the other coefficients are AC coefficients.

Figure 4.4 is the plots of ‘non Flashing Blocks’ location in the spatial flat domain. 300 data blocks from the compressed video are plotted together to show the temporal change. The blocks are plotted in raster order.

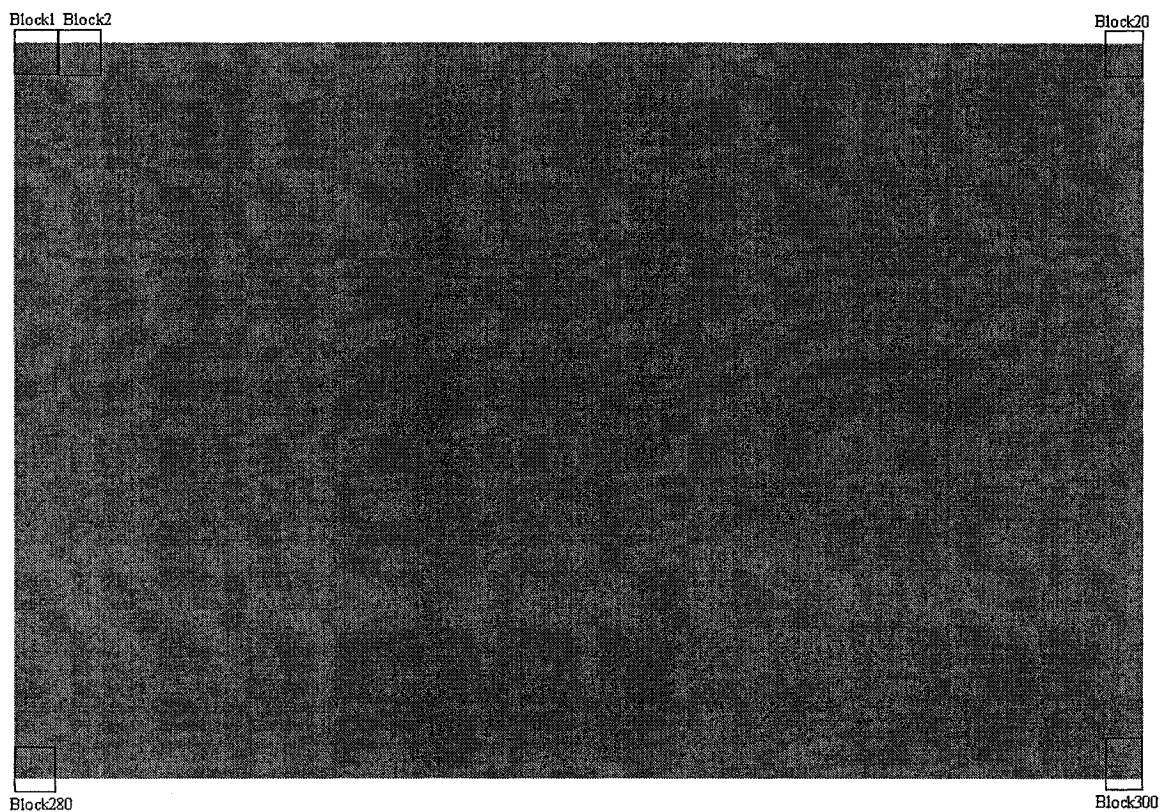


Figure 4.4 the Temporal Plot the ‘non Flashing Blocks’ (Spatial flat Domain)

Figure 4.5 is the temporal plot of the ‘non Flashing Blocks’ spatial location in the DCT domain. Each DCT coefficient of the blocks in Figure 4.4 are plotted together to show the temporal change of the ‘non Flashing Blocks’. In the DCT-Based plots, black means that the DCT coefficients are zero and white means that they are non zero.

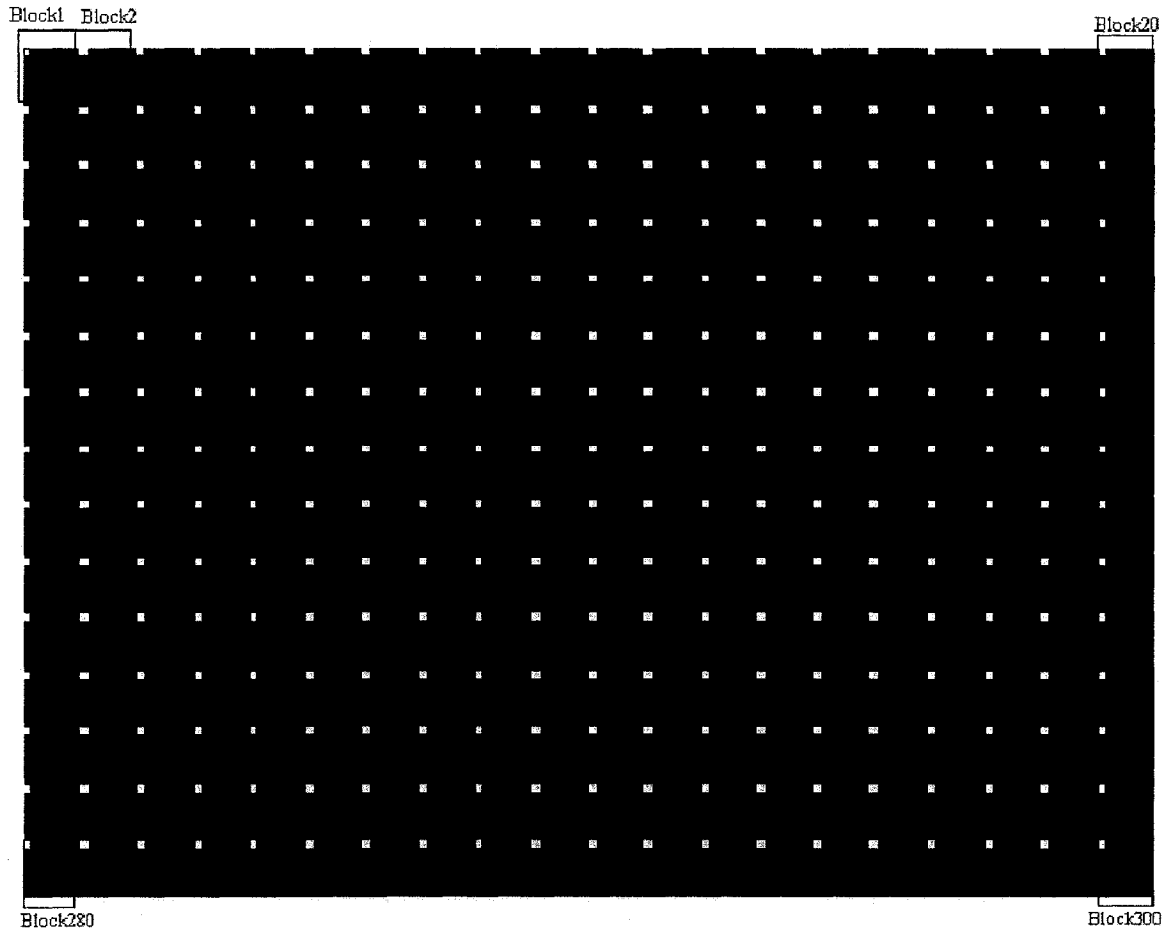


Figure 4.5 the Temporal Plot of the 'non Flashing Blocks'(DCT Domain)

1624	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 4.2 2D DCT coefficients of the block in the first frame in Figure 4.5

Table 4.2 shows 64 coefficients from the block in the first frame. The coefficient 1624 located in the first row and the first column is a DC coefficient, while the other coefficients are AC coefficients.

Figure 4.6 is the plots of a ‘non Flashing Blocks’ location in the spatial domain. This is for a non-flat block. 300 data blocks from the compressed video are plotted together to show changes along the temporal axis. The blocks are plotted in raster order.

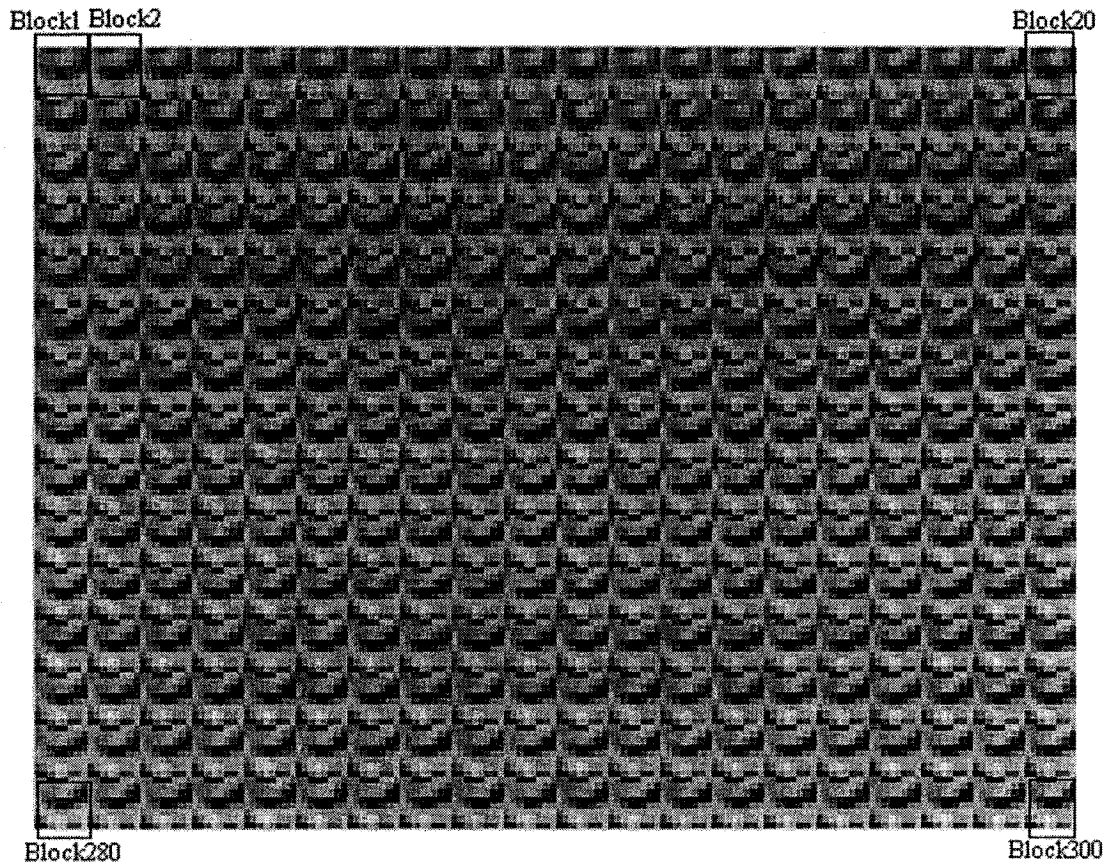


Figure 4.6 the Temporal Plot of the ‘non flashing block’ (Spatial non-flat Domain)

Figure 4.7 is the temporal plot of the same ‘non flashing blocks’ location in the DCT domain. Each DCT coefficient of the blocks in Figure 4.6 are plotted together to show the change along the temporal axis. In the DCT-Based plots, black means that the DCT coefficients are zero and white means that they are non zero.

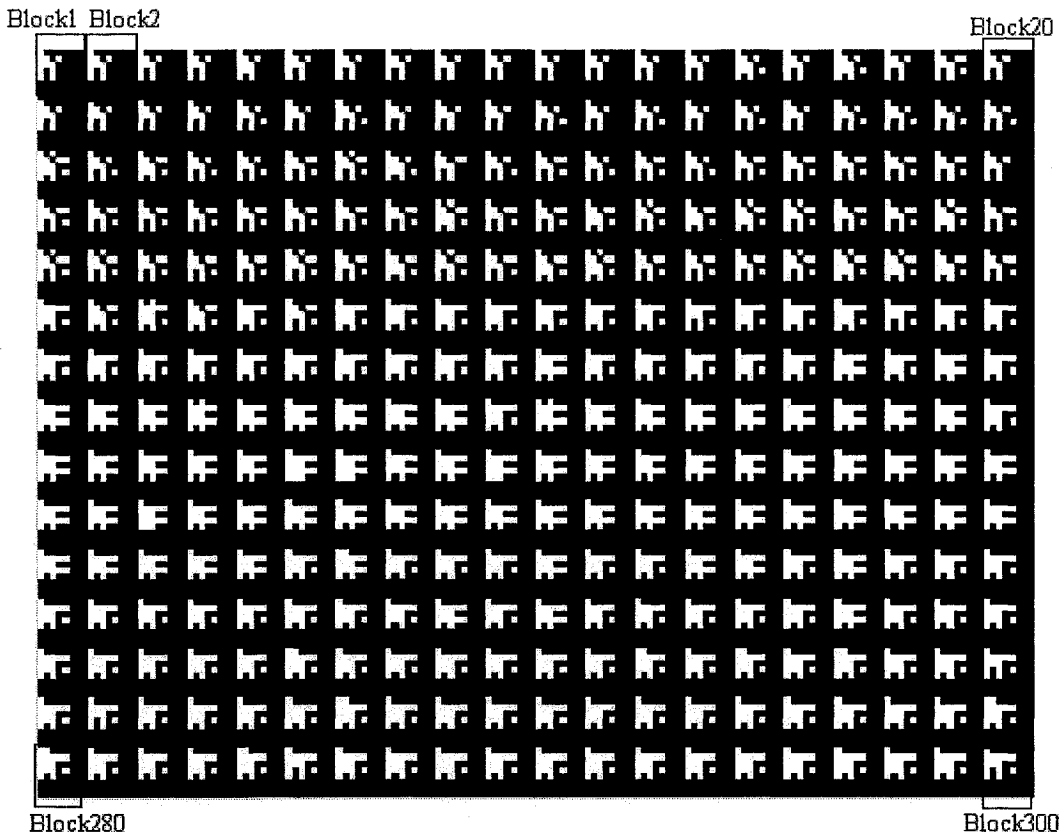


Figure.4.7 the Temporal Plot of the ‘non Flashing Blocks’ (DCT Domain)

1624	0	0	0	0	0	0	0
-96	48	0	144	0	0	0	0
57	-66	-78	0	0	0	0	0
66	0	78	0	0	0	0	0
-132	0	81	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 4.3 2D DCT coefficients of the block in the first frame in Figure 4.7

Table 4.3 shows 64 coefficients from the block in the first frame. The coefficient 1624 located in the first row and the first column is a DC coefficient, while the other coefficients are AC coefficients.

Figures 4.2 and 4.3 indicate that differences along the temporal axis are present in ‘Flashing Blocks’. This makes sense as block flashing is a temporal artifact. On the other hand, in Figures 4.4 and 4.5 which is a ‘non-Flashing Blocks’ there is little or no variation along the temporal axis. Figures 4.6 and 4.7 show a different kind of ‘non-

Flashing Blocks’. Here there is some variation along the temporal axis however the spatial detail masks the block flashing artifact. Note that a high spatial detail block with little or no variation along the temporal axis also results in a ‘non-Flashing Blocks’. These observations will be used to develop a ‘Flashing Blocks’ QP in the next sections.

### 4.3 ‘Flashing Blocks’ Detection Method

In this section, ‘Flashing Blocks’ Detection Method (FBDM) will be introduced. According to the conclusions obtained from the last section, it is found that the ‘Flashing Blocks’ are temporal artifact and locate in the spatial flat domain because spatial detail masks the block flashing artifact. In [40], it is also found for the spatial domain that HVS is sensitive to the artifacts in the middle luminance range but is not sensitive the artifacts in the Over Dark and Over Bright domain. So ‘Flashing Blocks’ Detection Method can be divided into three parts-The Temporal Method (TM) and The Over Dark and Over Bright Method (ODOBM) and The Spatial Frequency Method (SFM). The outputs of these methods are combined by the logic operation “and”. This FBDM can be expressed as the following figure:

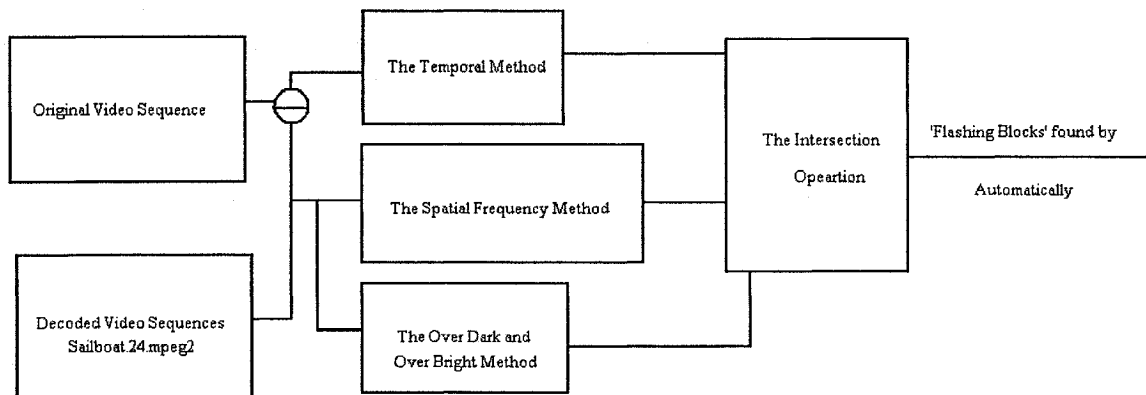


Figure 4.8 ‘Flashing Blocks’ Detection Method

Here three methods are used to identify ‘Flashing Blocks’. Blocks that are labeled as ‘Flashing Blocks’ by all three methods will be labeled as ‘Flashing Blocks’ by the overall method. If the  $D$  is the set of ‘Flashing Blocks’ selected by the overall method and the ‘Flashing Blocks’ found by the TM is  $A$  and the ‘Flashing Blocks’ found by the SFM is  $B$  and the ‘Flashing Blocks’ found by the ODOBM is  $C$ , then:

$$D=A \cap B \cap C \quad (4.1)$$

In the next sections, the TM, and SFM and ODOBM are introduced.

#### **4.3.1 the Temporal Method**

In this section, a TM will be proposed to help locate the ‘Flashing Blocks’. The proposed method has the following steps:

- Get the difference sequence,  $e(x,y,n)$ , between the original video sequence,  $fo(x,y,n)$ , and degraded video sequence,  $fd(x,y,n)$ . That is, form  $e(x,y,n)=fo(x,y,n)-fd(x,y,n)$ . Here  $x$  is the spatial vertical location in the video sequence where  $x=1,\dots,480$ ;  $y$  is the spatial horizontal location in the video sequence where  $y=1,\dots,720$  and  $n$  is the temporal location in the video sequence where  $n=1,\dots,300$ .
- Divide  $e(x,y,n)$  into S-T regions ( $8 \times 8 \times 300$ ).
- For each S-T region there are three hundred  $8 \times 8$  spatial blocks. Take the 2D spatial DCT on each of these blocks.
- For each S-T region label the 300 DC coefficients of each block  $d_1, d_2, d_3, \dots, d_{300}$ .

- Take a length 300 DCT (along the temporal direction) of the DC coefficients  $d_1, d_2, d_3, \dots, d_{300}$ . The result is denoted  $t_1, t_2, t_3, \dots, t_{300}$ . Here  $t_1$  represents the DC coefficient of the entire S-T region.
- Get the temporal AC energy (Re) of the 300 DC coefficients of each block in a

$$\text{S-T region in the temporary domain. } Re = \sum_{i=2}^{300} (t_i)^2$$

If  $Re < Te$ , this S-T region is thought as a potential 'Flashing Blocks'. The temporal threshold ( $Te$ ) is set in the Section 4.3.4.

#### 4.3.2 The Spatial Frequency Method

Because the 'Flashing Blocks' artifact can be masked in areas of spatial detail, the Spatial Frequency Method (SFM) mentioned in Section 2.4 attempts to find S-T regions that have low spatial detail content. It is as follows:

1. Get the sobel filtered image ( $fd1$  and  $fd2$ ) of the first frame in the degraded video sequences using the (Finite Impulse Response Linear Time Invariant) sobel filters.

Their impulse responses are:

$$\text{Vertical Sobel Filter } \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad \text{Horizon Sobel Filter } \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

2. Get the sobel magnitude image  $fdm$ :

$$fdm(x, y) = \sqrt{fd1(x, y)^2 + fd2(x, y)^2} \quad (4.2)$$

3. If the sobel magnitude of every pixel in a block  $< Tl$ , this block is labeled as a potential 'Flashing Blocks'. The value of  $Tl$  is empirically set.

### 4.3.3 The Over Dark and Over Bright Method:

‘Flashing Blocks’ is most visible in the range of middle luminance. The Over Dark and Over Bright Method(OBODM) has the following steps:

1. Get the degraded video sequence,  $fd(x,y,n)$ . Here  $x$  is the spatial vertical location in the video sequence where  $x=1,\dots,480$ ;  $y$  is the spatial horizontal location in the video sequence where  $y=1,\dots,720$  and  $n$  is the temporal location in the video sequence where  $n=1,\dots,300$ .
2. Divide  $fd(x,y,n)$  into S-T regions ( $8\times 8\times 300$ ).
3. For each S-T region there are three hundred  $8\times 8$  spatial blocks. Take the 2D spatial DCT on each of these blocks.
4. For each S-T region label the 300 DC coefficients of each block  $m_1, m_2, m_3, \dots, m_{300}$ .
5. Take a length 300 DCT (along the temporal direction) of the DC coefficients  $m_1, m_2, m_3, \dots, m_{300}$ . The result is denoted  $n_1, n_2, n_3, \dots, n_{300}$ . Here  $n_1$  represents the DC coefficient of the entire S-T region.

If  $ll < n_1 < lh$ , this S-T region is labeled as a potential ‘Flashing Blocks’. Here  $ll$  and  $lh$  are empirically set.

### 4.3.4 Selection of the Temporal Threshold T1

The process of selecting the temporal threshold for the Temporal Method is given in this section. The experimental video sequence is the decoded video sequence of sailboat.24.mpeg2. These contain 300 frames and each clip has spatial dimensions  $480\times 720$ .

In setting the temporal threshold the thresholds of the SFM and the ODOBM are held constant. The parameter  $T_e$  in the TM is varied and the results of the overall ‘Flashing Blocks’ Detection Method(FBDM) are examined. In this examination the concept of miss and false alarm are used from detection theory[39]. A missed block is an S-T region that truly has the ‘Flashing Blocks’ artifact but that FBDM does not label as a ‘Flashing Blocks’. A false alarm block is a S-T region that does not have the ‘Flashing Blocks’ artifact but that FBDM labels as a ‘Flashing Blocks’. To use these concepts, it is necessary to obtain a “true” labeling of each S-T region as either a ‘Flashing Blocks’ or a ‘non Flashing Blocks’. To this end, two individuals (graduate students including this author) viewed the decoded video sequence independently. S-T regions that both observers labeled as ‘Flashing Blocks’ are taken as true ‘Flashing Blocks’. All other S-T regions are labeled as ‘non-Flashing Blocks’. The true ‘Flashing Blocks’ are shown outlined in red in Figure 4.9.

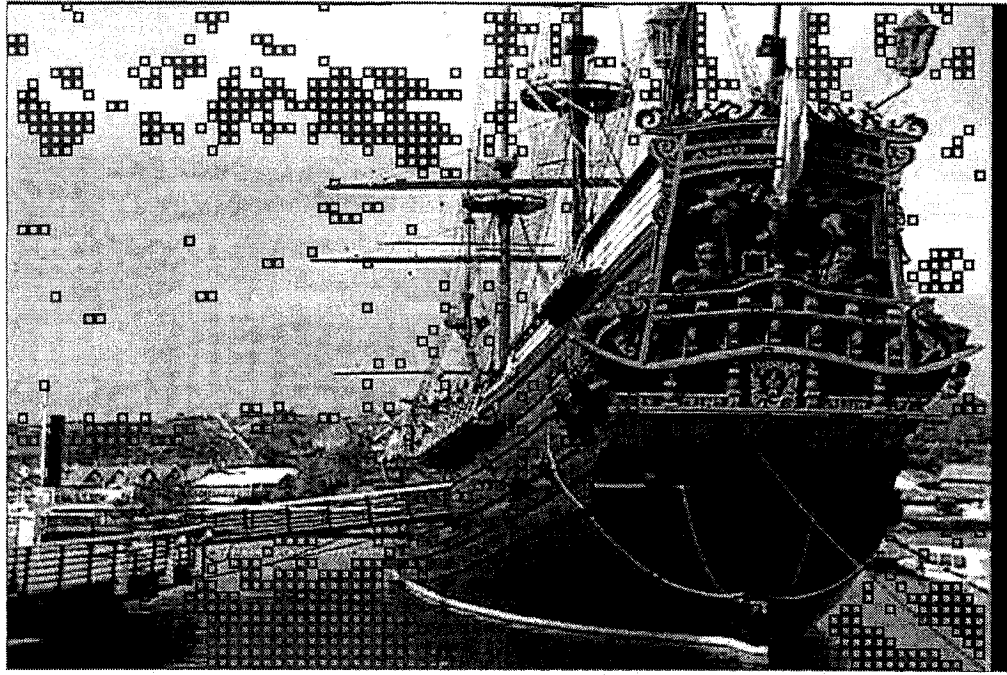


Figure 4.9 Spatial Locations of the 'Flashing Blocks' found by eyes

The true 'Flashing Blocks' are called 'Flashing Blocks' found by eyes. The following standard quantities are defined:

$$\text{Miss Rate} = \frac{\text{number of missed flashing blocks}}{\text{number of 'flashing blocks found by eye'}}$$
 (4.3)

$$\text{False Alarm Rate} = \frac{\text{number of false alarm flashing blocks}}{\text{number of total blocks-number of 'flashing blocks found by eye'}}$$
 (4.4)

By using different temporal thresholds, the false alarm rates and the miss rates are computed, shown in Table 4.4 and plotted in Figure 4.10.

<b>Temporal Threshold Te</b>	<b>The number of 'Flashing Blocks' found by FBDM</b>	<b>The number of 'Flashing Blocks' found by Eye</b>	<b>The number of False Alarm blocks</b>	<b>The number of Miss blocks</b>	<b>False Alarm Rate</b>	<b>Miss Rate</b>
<b>150</b>	750	741	211	190	0.0453	0.256
<b>130</b>	920	741	336	167	0.072	0.23
<b>120</b>	1020	741	447	153	0.096	0.21
<b>115</b>	1150	741	556	138	0.12	0.19
<b>105</b>	1276	741	658	123	0.14	0.166
<b>95</b>	1398	741	762	105	0.16	0.14
<b>85</b>	1506	741	854	90	0.18	0.12
<b>75</b>	1588	741	923	75	0.19	0.1
<b>65</b>	1682	741	1002	60	0.22	0.08
<b>50</b>	1784	741	1083	40	0.23	0.05

Table 4.4 'Flashing Blocks' found by eye and FBDM (the threshold for the SFM is  $T_l=400$  and the threshold for the ODOB is  $l_l=50$  and  $l_h=220$ )

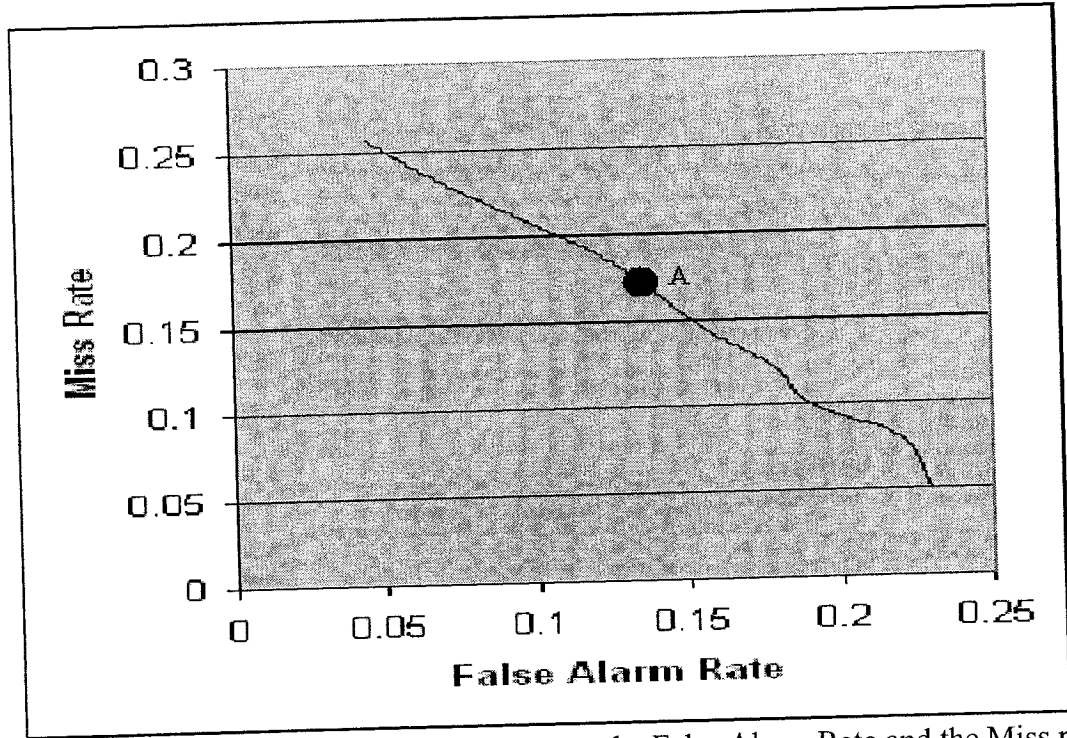


Figure 4.10 the relationship curve between the False Alarm Rate and the Miss rate (the threshold for the SFM is  $T_l=400$  and the threshold for the ODOB is  $l_l=50$  and  $l_h=220$ )

The point marked A ( $T_e$ ) is selected as giving a good trade-off between false alarm and miss rates and will be used in simulations in the remainder of this chapter.

#### 4.4 Method to compute the 'Flashing Blocks' QP

After locating the 'Flashing Blocks' using FBDM, a QP denoted T must be developed.

The procedure of generating T could be expressed using Figure 4.11:

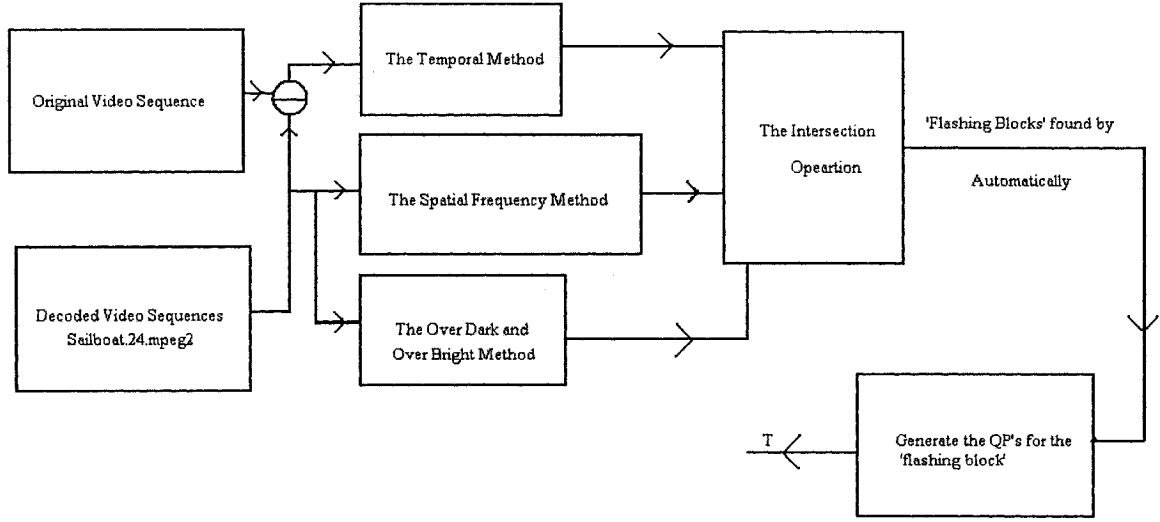


Figure 4.11 the Procedure for generating the QP (T) for the 'Flashing Blocks' artifact  
The QP for the 'Flashing Blocks' artifact is computed as follows:

$$T = \frac{\text{Number of 'Flashing Blocks' found by automatically}}{\text{Number of total blocks}} \quad (4.8)$$

The QP T is used to represent the temporal 'Flashing Blocks' artifact of the video sequence. The QP (T) combined with the spatial QP's should be used as the input to the fuzzy system that estimates the Mean Score Value of subjective observation.

#### 4.5. Fuzzy Model Using Temporal Artifacts as the Input

In this section, a nonlinear mapping will be used to combine the spatial and temporal QP's of the video sequences to simulate the perception of human beings. The spatial QP's are the same as the QP's that were used in Chapter 3. It can be divided into two parts –8 QP's and 12 QP's. Because this nonlinear mapping is very complex, a fuzzy algorithm will be used to approximate the nonlinear mappings

#### 4.5.1 Fuzzy Models with 4+1 QP's

In this section, four spatial QP's ( $x_1, x_2, x_3, x_4$ ) defined in the Section 3.2 and one temporal QP obtained from the last section (T) will be used as the input to the T-S fuzzy system used for non linear mapping whose goal is to reflect the Mean Opinion Score of the subjective evaluation. Define  $\mathbf{x}=[x_1 \ x_2 \ x_3 \ x_4 \ T]$ . Similarly to Chapter 3, the proposed fuzzy algorithm is as follows:

If the  $x_1$  is big and  $x_2$  is small and  $x_3$  is small and  $x_4$  is small and T is big, then

Join<sub>1</sub>( $\mathbf{x}$ )

If the  $x_1$  is small and  $x_2$  is big and  $x_3$  is big and  $x_4$  is big and T is small, then

Join<sub>2</sub>( $\mathbf{x}$ )

This fuzzy metric can be divided into two parts: the premise part (if section) and the consequent part (then section) and can be described mathematically as:

$$\text{Join}(\mathbf{x}) = \frac{\sum_{i=1}^2 \mathbf{B}_i(\mathbf{x}) \text{join}_i(\mathbf{x})}{\sum_{i=1}^2 \mathbf{B}_i(\mathbf{x})} = \sum_{i=1}^2 \text{Normal}_i(\mathbf{x}) \text{join}_i(\mathbf{x}) \quad (4.9)$$

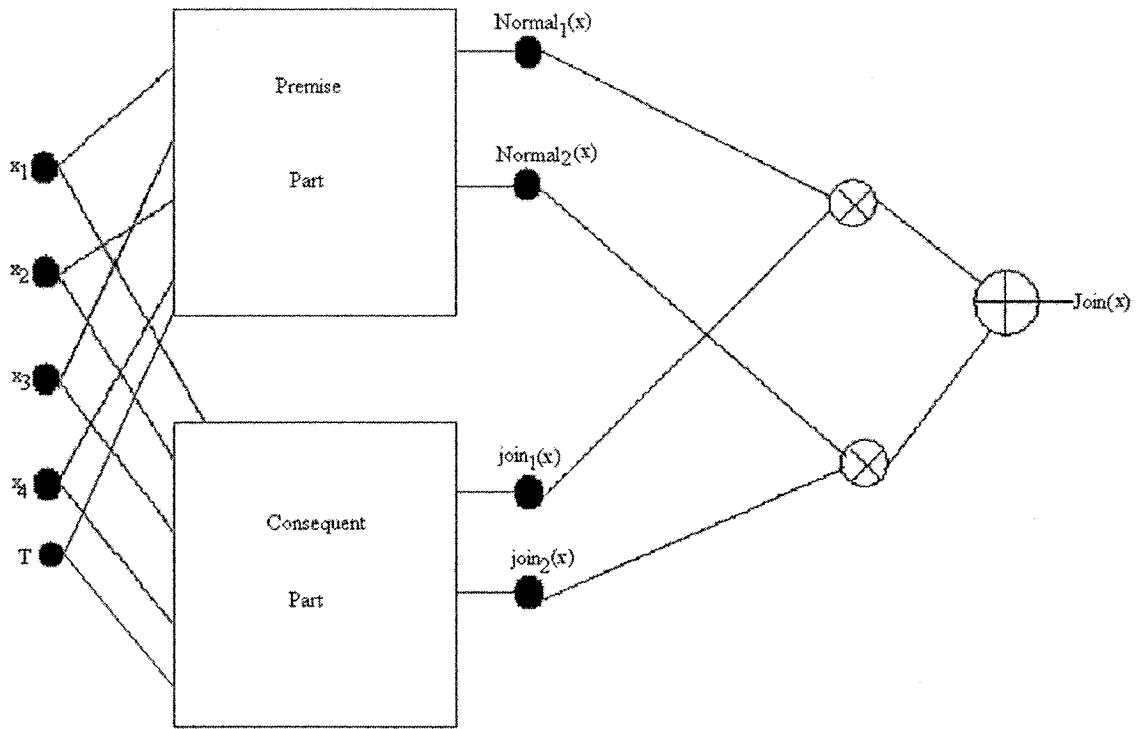


Figure 4.12 Figure of the Fuzzy Algorithm (4.9)

where  $join_i(x)$  is the output of the consequent part and  $normal_i(x)$  is the output of the premise part. The Sum of the product ( $join_i(x) \times Normal_i(x)$ ) of the output of the premise part and the consequent part is the output of the fuzzy system- $Join(x)$ . The detailed description of this fuzzy algorithm and the parameter tuning of this algorithm is similar to Section 3.2, readers can refer to it.

#### 4.5.2 Fuzzy Model with 8+1QP's

In this section, 8 spatial QP's defined in Table 2.4 ( $x_i$ ,  $i=1..8$ )(flat and non-flat) combining Wolf's and Bistawhi's method and one temporal QP's (T) will be used as the input to the fuzzy system. Define  $x=[x_1 \ x_2 \ x_3 \ x_4 \ x_5 \ x_6 \ x_7 \ x_8 \ T]$ . The proposed fuzzy algorithm is as follows:

If the  $x_1$  is small and  $x_2$  is small and  $x_3$  is small and  $x_4$  is small and  $x_5$  is small and  $x_6$  is small and  $x_7$  is small and  $x_8$  is small and  $T$  is small, Then

$Join_1(x)$

If the  $x_1$  is big and  $x_2$  is big and  $x_3$  is big and  $x_4$  is big and  $x_5$  is big and  $x_6$  is big and  $x_7$  is big and  $x_8$  is big and  $T$  is small, Then

$Join_2(x)$

This fuzzy metric can be divided into two parts: the premise part (if section) and the consequent part (then section) and can be described in mathematically as follow forms:

$$Join = \frac{\sum_{i=1}^2 B_i(x) join_i(x)}{\sum_{i=1}^2 B_i(x)} = \sum_{i=1}^2 Normal_i(x) join_i(x) \quad (4.10)$$

The following figure is used to express (4.10)

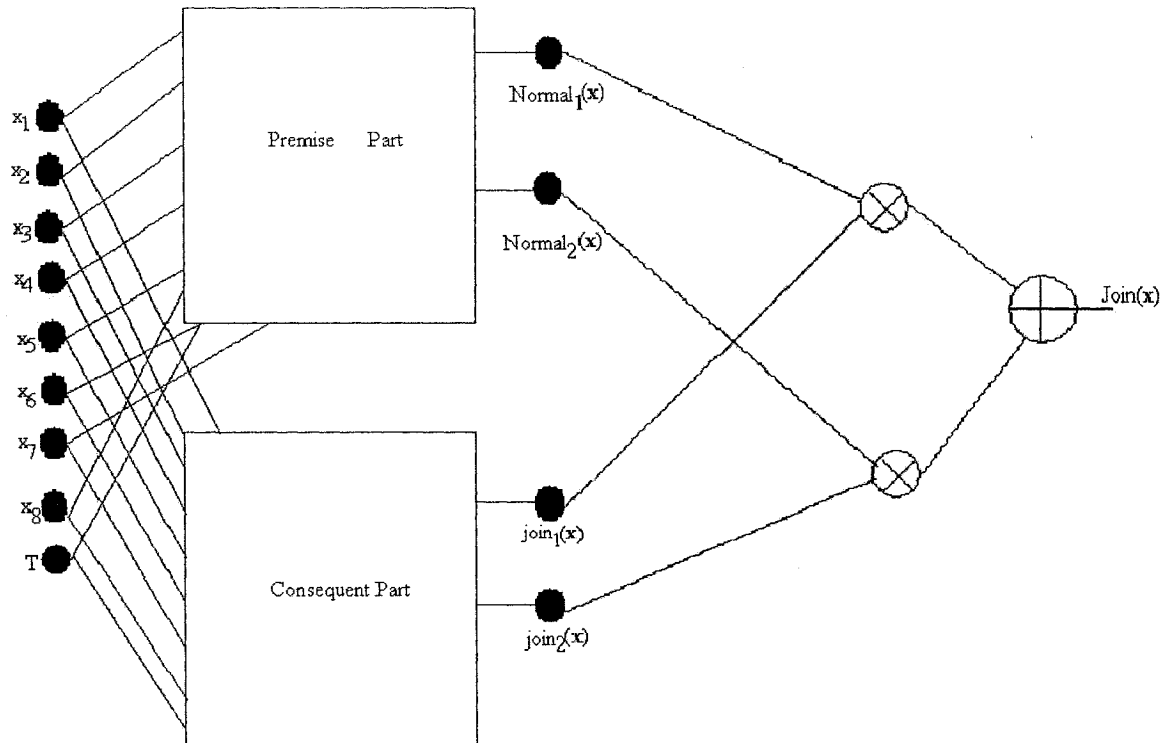


Figure 4.13 Figure of the Fuzzy Metric

where  $\text{join}_i(\mathbf{x})$  ( $i=1..2$ ) is the output of the consequent part and  $\text{Normal}_i(\mathbf{x})$  ( $i=1,2$ ) is the output of the premise part. The Sum of the product ( $\text{join}_i(\mathbf{x}) \times \text{Normal}_i(\mathbf{x})$ ) of the output of the premise part and the consequent part is the output of the fuzzy system- $\text{Join}(\mathbf{x})$ . The detailed description of this fuzzy algorithm and the parameter tuning of this algorithm is similar to Section 3.3 and readers can refer to it.

#### 4.5.3 Fuzzy Model with 12+1 QP's

As in Section 3.4, for the twelve QP's defined in Table.2.5 which are represented as  $x_i$ ,  $i=1..12$  and one temporal QP (T), the algorithm combining linear mapping and fuzzy algorithm is used. This algorithm is divided into two parts: At first, the linear mapping method is used to get the input of the fuzzy system. Then the fuzzy method is used to get the final result. This method will give a simpler algorithm comparing with other algorithms similar to Equations 4.9 and 4.10. Assume the output of the linear mapping is M, then the following steps can be obtained:

1.

$$M = \sum_{i=1}^{12} \text{REG}_i x_i + \text{REG}_{13} T;$$

$\text{REG}_i$   $i=1..13$  is linear parameters produced by linear regression.

2.

If M is very small

Then  $\text{join}_1(M)$

If M is small

Then  $\text{join}_2(M)$

If M is big

Then join<sub>3</sub>(M)

If M is very big

Then join<sub>4</sub>(M)

The second part (fuzzy algorithm) can be expressed in mathematically as follows:

$$\text{Join}(M) = \frac{\sum_{i=1}^4 \text{MF}_i(M) \text{join}_i(M)}{\sum_{i=1}^4 \text{MF}_i(M)} = \sum_{i=1}^4 \text{Normal}_i(M) \text{join}_i(M) \quad (4.11)$$

The following figure can be used to express this hybrid algorithm:

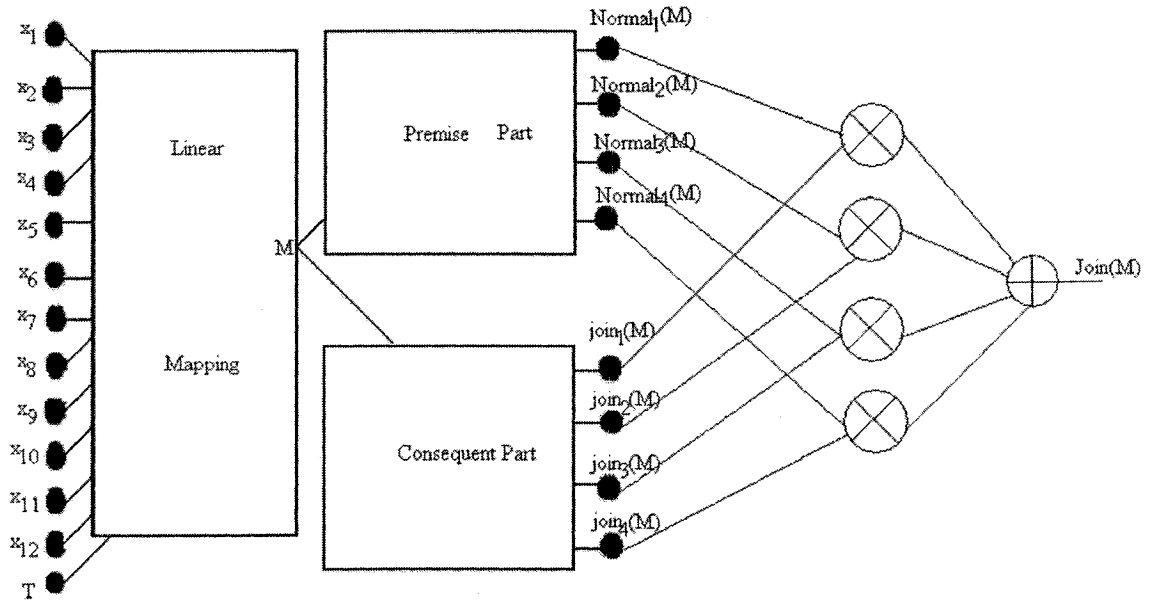


Figure 4.14 Figure of the Hybrid Algorithm

where  $\text{join}_i(M)$  ( $i=1,2$ ) is the output of the consequent part and  $\text{Normal}_i(M)$  ( $i=1,2$ ) is the output of the premise part.. The Sum of the product ( $\text{join}_i(M) \times \text{Normal}_i(M)$ ) of the output of the premise part and the consequent part is the output of the fuzzy system- $\text{Join}(M)$ . The detailed description of this fuzzy algorithm and the parameter tuning of this algorithm is similar to Section 3.4, and readers can refer to it.

#### 4.6 Simulation Result of Fuzzy Algorithms

In this section, the simulation results of the fuzzy algorithms that include the temporal QP (T) will be shown. The fuzzy algorithms that have the same spatial QP's but do not include T will also be shown as a reference. The fuzzy algorithm that includes four spatial QP's and T is called 4QPFAT. It has the same spatial QP's as the fuzzy algorithm (4QPFA) defined in Section 3.5. The fuzzy algorithm that includes eight spatial QP's and T is called 8QPFAT. It has the same spatial QP's as the fuzzy algorithm (8QPFA) defined in Section 3.5. For the fuzzy algorithm that includes twelve spatial QP's and T is called 12QPFAT. It has the same spatial QP's as the fuzzy algorithm (12QPFA) defined in Section 3.5. The figure of merit is the mean square error between the results predicted by the fuzzy algorithms and the subjective rating of human beings. The training set and non training set of the video sequences are the same as those mentioned in Section 3.5.

Table 4.5 summarizes the simulation results for 4QPFAT and 4QPFA. It is found that the MSE between 4QPFAT and the subjective ratings is smaller than that between the 4QPFA and the subjective rating. This indicates that the temporal QP (T) contributes to the enhancement of Quality Scores.

	4QPFAT	4QPFA
Training Set	22.8	33.14
Non Training Set	24.5	44.59
Total	23.6	36.96

Table 4.5 Mean Square Error (MSE) between the given Quality Score (4QPFAT and 4QPFA) and the subjective rating values

Figure 4.15 is the scatter plot of 4QPFAT versus the subjective rating values, while Figure 4.16 is the scatter plot of 4QPFA and subjective rating values.

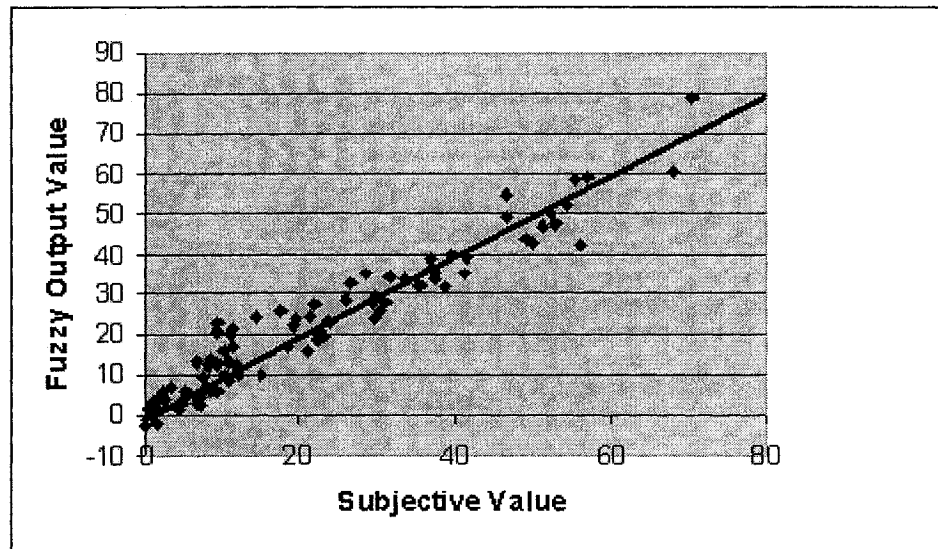


Figure 4.15 Scatter Plot of the Output of the 4QPFAT versus Subjective Rating Values

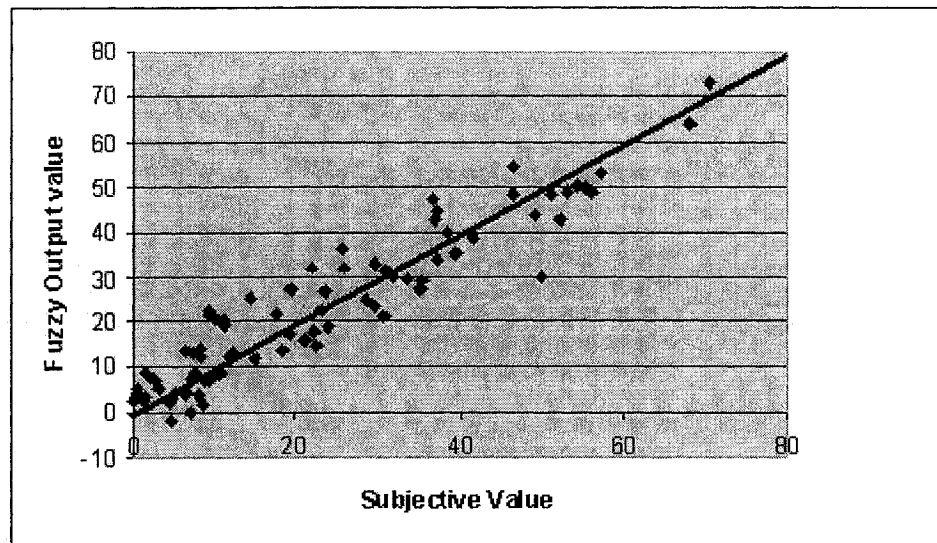


Figure 4.16 Scatter Plot of the Output of the 4QPFA versus Subjective Rating Values

It is found that the scatter points in Figure 4.15 are located closer to the plot diagonals than the scatter points in Figure 4.16. This further indicates that T contributes to the enhancement of the Quality Score.

Table 4.6 summarizes the simulation results for 8QPFAT and 8QPFA. It is found that the MSE between 8QPFAT and the subjective ratings is smaller than that between the FQPFA and the subjective rating. This indicates that the temporal QP's (T) contributes to the enhancement of Quality Scores.

	8QPFAT	8QPFA
Training Set	9.17	9.68
Non Training Set	13.85	16.46
Total	10.74	11.94

Table.4.6 Mean Square Error (MSE) between the given Quality Score (8QPFAT and 8QPFA) and the subjective rating values

Figure 4.17 is the scatter plot of 8QPFAT versus the subjective rating values, while Figure 4.18 is the scatter plot of 8QPFA and subjective rating values.

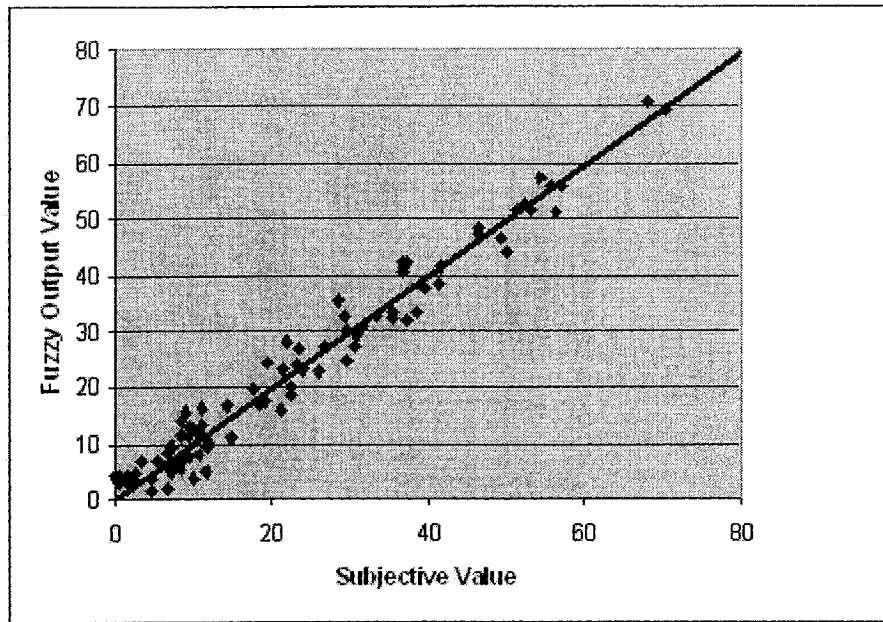


Figure 4.17 Scatter Plot of the Output of the 8QPAT versus Subjective Rating Values

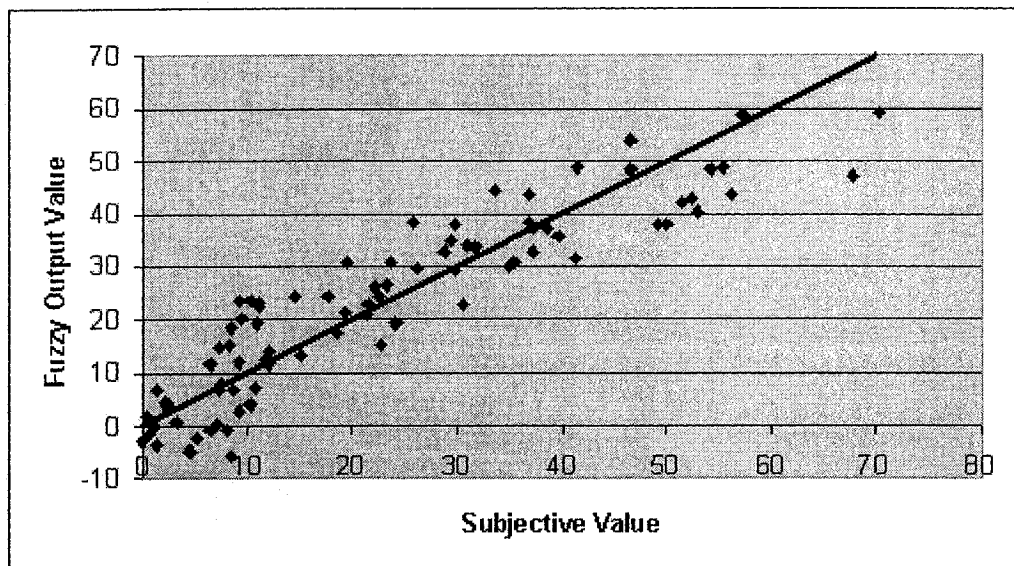


Figure 4.18 Scatter Plot of the Output of the 8QPFA versus Subjective Rating Values

It is found that the scatter points in Figure 4.17 are located closer to the plot diagonals than the scatter points in Figure 4.18. This further indicates that T contributes to the enhancement of the Quality Score.

Table 4.7 summarizes the simulation results for 12QPFAT and 12QPFA. It is found that the MSE between 12QPFAT and the subjective ratings is smaller than that between the 12QPFA and the subjective rating. This indicates that the temporal QP (T) contributes to the enhancement of Quality Scores.

	12QPFAT	12QPFA
Training Set	27.59	25.44
Non Training Set	46.17	50.87
Total	33.79	33.92

Table 4.7 Mean Square Error (MSE) between the given Quality Score (12QPFAT and 12QPFA) and the subjective rating values

Figure 4.19 is the scatter plot of 12QPFAT versus the subjective rating values and Figure 4.20 is the scatter plot of 12QPFA versus the subjective rating values.

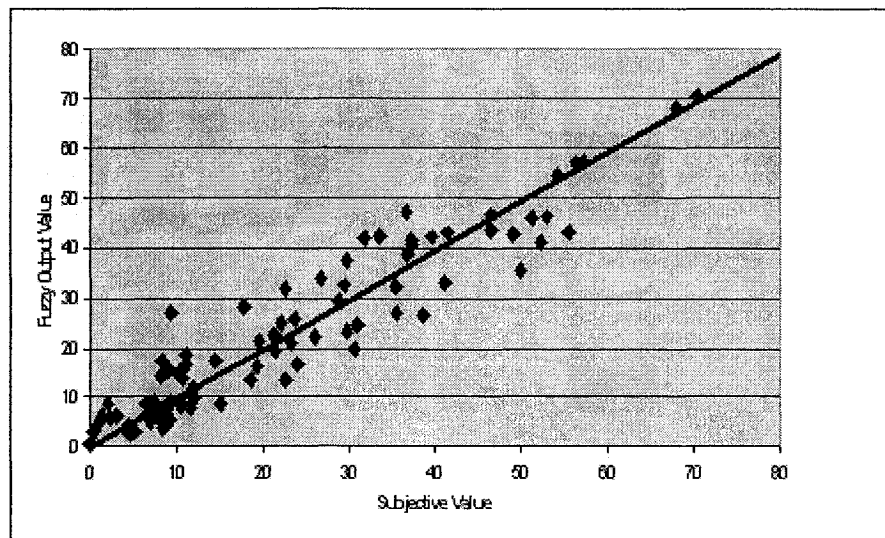


Figure 4.19 Scatter Plot of the Output of the 12QPFAT versus Subjective Rating Values

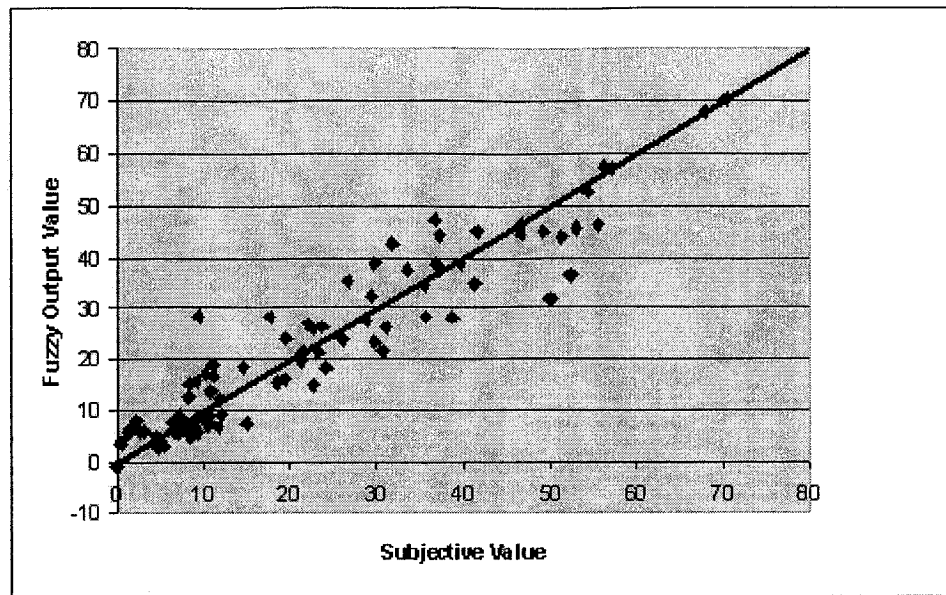


Figure 4.20 Scatter Plot of the Output of the 12QPFA versus Subjective Rating Values

It is found that the scatter points in Figure 4.19 are located closer to the plot diagonals than the scatter points in Figure 4.20. This further indicates that T contributes to the enhancement of the Quality Score.

#### 4.7 Conclusions:

In this chapter, a new temporal artifact-“flashing block” is introduced. Intuitively the addition of a temporal Quality Primitive to measure this artifact should enhance Quality Scores. This is borne out by the simulation results presented in this chapter.

## Chapter 5

### Conclusions and Further Research

#### 5.1 Contributions

This thesis presents a method using Fuzzy Logic to automatically estimate the subjective ratings that human users would assign to video which has been compressed. The framework followed is to create Quality Primitives that estimate particular artifacts commonly created by compression and then creating an equation or mapping to translate them into an appropriate Quality Score. The main contributions of this work are as follows:

- **Using Fuzzy Logic to do the Nonlinear Mapping of Quality Primitives to a Quality Score**

Simulation results show that this strategy is successful.

- **Define a Quality Primitive to measure the ‘Flashing Blocks’ artifact**

The ‘Flashing Blocks’ is an artifact that differs from other three artifacts previously defined (i.e., Blocking, Ringing and Blurring). It is a temporal artifact found in video rather than a spatial artifact like the others.

#### 5.2 Conclusions

Given the simulation results it is apparent that a linear mapping of the Quality Primitives is not sufficient to capture the complexity of the Human Visual System. The improvement found by using a mapping based on Fuzzy Logic indicates that non-linear mappings offer potential to improve performance.

Similarly, the improvement in performance of Quality Scores that use a Quality Primitive that is temporal in nature (rather than solely spatial) indicates that temporal artifacts play an important part in human perception. For compressed video the Block Flashing artifact examined here was clearly important. In the context of degradations other than those caused by compression, temporal artifacts and Quality Primitives that measure them, should also be considered.

### **5.3 Further Research**

In order to achieve better impairments detection and measurement, and to reach a video quality rating that is highly correlated with subjective measurements results, the following tasks are recommended in the future:

- Other frameworks for generating Non-Linear Mappings should be explored.
- Because the ‘Flashing Blocks’ is a new temporal artifact, more research work need to be done to refine the methods to locate ‘Flashing Blocks’ and to reduce the difference between the ‘Flashing Blocks’ found by automatically and the ‘Flashing Blocks’ found by human beings.
- Extension of this framework to degradations other than those caused by compressions should be explored.
- Quality Primitives for temporal artifacts other than ‘Flashing Blocks’ (both with compression degradation and other types of degradation) should be investigated.
- The test sequences should not be limited on MPEG-2, but rather be extended to MPEG-1 and MPEG-4 video sequences. Here, the automatic metric may directly measure MPEG-1 and MPEG-4 video without any modification.

## Bibliography

- [1] S. Winkler, "Issues in vision modeling for perceptual video quality assessment," Signal Processing, vol. 78, no. 2, pp. 231-252, 1999.
- [2] Michael W.Marcellin, "An overview of JPEG-2000", Data Compression Method, 2000.
- [3] K.T.Tan et al.: "An objective measurement tool for MPEG video quality", Signal Processing, 1998.
- [4] Yao Wang and Ya-Qing Zhang, "Video Processing and Communications," Prentice Hall, August 2000.
- [5] "Objective Perceptual Video Quality Measurement Using a JND-Based Full Reference Technique", ANSI T1.TR.74-2001, American National Standards Institute, New York, 2001.
- [6] S. Winker, "Issues in Vision Modeling for Perceptual Video Quality Assessment", Signal Processing, October 1999.
- [7]Judith Dijk, Piet W., "A quantitative measure for the perception of sharpening and smoothing in images", 5<sup>th</sup> Annual Conference of the advanced School for Computing and Imaging(ASCI), 1999.
- [8] Ismail Avcibas, "Steganalysis Using Image Quality Metrics", IEEE Transactions on Image Processing, Vol.12. No.2, February 2003.
- [9] Bernd Girod, "A Multistage Perceptual Quality Assessment for Compressed Digital Angiogram Images", IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 20, NO. 12, DECEMBER 2001.

- [10] Ahmet M. Eskicioglu, "Quality Measurement for Monochrome Compressed Images in the past 25 Years", in Proceedings of Optical Engineering, May 2000.
- [11] R.C.Gonzalez and R.E.Woods, "Digital Image Processing", Addison-Wesley, 2002.
- [12] Andrew B.Waston, "Digital Images and Human Vision", Human Vision and Electronics Imageing II, 1997.
- [13] S.Winkler, E.D,Gelasca, T.Ebrahimi, "Toward Perceptual Metrics for video Watermark Evaluation", in Proceedings of SPIE, 2003.
- [14] Thomas.D.Kite, "Image Quality Assessment Based on a Degradation Model", IEEE Transactions on Image Processings, Apr, 2000.
- [15] V.R.Algazi, Y.Kato, M.Miyahara, and K.Kotani, "Object-based Intra-frame Wavelet Video Codings", in Proceedings of the SPIE, Application of Digital Image Processing, 2001.
- [16] E. Chiu, S. Atkins, and J. Vaisey, "Wavelet-based space-frequency compression of ultrasound images", IEEE Transactions on Information Technology in Biomedicine, vol. 5, pp. 300-310, 2001.
- [17] F.X.J.Lukas and Z.L.Budrikis, "DVQ: A digital video quality metric based on human vision", Journal of Electronics Imaging, 10(1), 2001
- [18] Paul M.Farrelle, "Recursive Block Coding for Image Data Compression", Springer-Verlag, 2003.
- [19] E. Chiu, S. Atkins, and J. Vaisey, "MPEG-4 rate control for multiple video objects", IEEE Transactions on Circuit and Systems for Video Technology, vol.1, pp.186-199, December 1999.

- [20] S.Winker, "A perceptual distortion metric for digital color video", in proceeding of SPIE, vol.3644, pp.175-184.CA., 1999.
- [21] Marcelo J. Weinberger and Gadiel Serrousi, "Principles and Standardization into JPEG-LS", IEEE Transactions on Image Processing, Vol.9, No.8, Aug, 2000.
- [22] Anil K. Jain, "Fundamentals of Digital Image Processing", Prentice Hall, 1989.
- [23] Alan V, Oppenheim, Ronald W, Schafer "Discrete-time signal processing", Prentice Hall, 1998
- [24] Ahmet M. Eskicioglu and Paul S.Fisher, "A multidimensional image quality measure using Singular Value Decomposition", in Proceedings of Picture Coding Symposium, September 2001.
- [25] Peter K. Kaiser, Robert M. Boynton, "Human Color Vision", Optical Society of America, May 1996.
- [26] Z. Wang, L. Lu, and A. C. Bovik , "Video quality assessment based on structural distortion measurement", Signal Processing: Image Communication, special issue on Objective video quality metrics. February 2004
- [27] Patrick C. Teo and David J.Heeger, "A Perceptual Distortion Metric for Digital Color Images", in Proceedings of IEEE, 1998.
- [28] S. Wolf, and M. Pinson, "Spatial-temporal distortion Metric for in-service quality monitoring of any digital video system," in Proc. SPIE I.S.V.V.D.C., Boston, MA, September 1999.
- [29] S. Winkler, "A perceptual distortion metric for digital color video", Human Vision and Electronic Imaging IV, (1999).

- [30] A. B. Watson, "Toward a perceptual quality metric", Part of the IS&T /SPIE conference on Human Vision and Electronic Imaging III, San Jose, California. January 1998.
- [31] VQEG home page: [http:// www.its.bldrdoc.gov/vqeq/](http://www.its.bldrdoc.gov/vqeq/)
- [32] Final report from the video quality experts group on the validation of objective models of video quality assessment, March 2000
- [33] Wajih W. Bishtawi, "Objective measurement of subjective image quality", Master's thesis, Department of Electrical and Computer Engineering, Concordia University, 1996
- [34] S.Wolf and M.Pinson: "Video Quality Measurement Techniques", NTIA Report 02-392, 2002.
- [35] P. Witold, G. Fernando, "An Introduction to Fuzzy Sets: analysis and design", 1998.
- [36] Jyh-Shing Roger Jang, "ANFIS: Adaptive-Network-Based Fuzzy Inference System", IEEE Transactions on systems, man, and cybernetics, 1993.
- [37] FR-TV Full Reference Television Phase II Subjective Test Plan, September 2002.
- [38] Zheru Chi, Hong Yan and Tuan Pham, "Fuzzy Algorithms: with application to image processing and pattern regonition", 1996
- [39] D. Kazakos, "Detection and estimation," Computer Science Press, 1990.
- [40] 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, December 1995