DEVELOPMENT OF QUALITY-AWARE VIDEO SYSTEMS AND NMR SPECTRUM REGISTRATION

by

BASAVARAJ HIREMATH

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Electrical Engineering

THE UNIVERSITY OF TEXAS AT ARLINGTON

August 2007

ACKNOWLEDGEMENTS

It gives me immense pleasure to acknowledge few people who have been with me during my work and have been a constant source of inspiration. I would like to thank Dr. Zhou Wang, my supervisor, instructor, and my mentor at The University of Texas at Arlington for his immense support, constant guidance and resourceful inspiration. I am obliged to him for his belief in me to carry over the work he initiated. I would like to thank Dr. Seoung Bum Kim for his guidance and support during the implementation of our work. His inputs and feedbacks have been very critical in this work.

I would like to thank Dr. K R Rao for his interest in the research work and being part of the thesis committee. I would like to take this opportunity to thank Qiang Li for his support in implementing my work.

I am grateful to all my teachers, family members and friends for being part of my life and supporting me throughout.

Lastly, I would like to thank my parents for being my source of inspiration, for their love and support, and for the freedom they have provided me during my course of study.

June 18, 2007

ABSTRACT

DEVELOPMENT OF QUALITY-AWARE VIDEO SYSTEMS
AND NMR SPECTRUM REGISTRATION

Publication No. _____

BASAVARAJ HIREMATH, M. S

The University of Texas at Arlington, 2007

Supervising Professor: Zhou Wang

The work introduces a novel concept of *Quality-Aware Video* (QAV) System and demonstrates its successful implementation. During the course, it develops algorithms for reduced reference video quality assessment and for data hiding. The idea here is to extract quality defining features of the video sequence and embed them in the original video without causing any perceptual changes to obtain QAV. Such a QAV can then be exposed to distortions and adverse attacks that affect the perceptual quality of the video. At the receiving end, the algorithm extracts the quality features of the distorted video, decode the original video quality features, and estimates the current quality. The beauty of QAV is that they carry the original quality features along with them and hence their quality can be assessed anywhere on the fly. The algorithm developed does not assume any prior information about the attacks which means that the quality assessment is independent of the attack and shows that the algorithm has the potentials to generalize for various attacks. Our second work extends the existing idea of Bayesian estimation for registration to higher magnitude differences. This is achieved by employing pyramid (multi-resolution) approach to the existing algorithm. The advantage of pyramids is at the reduced scale we have better alignment of prominent features (peaks) of the spectrum

and as we move to higher levels, the finer details are taken care of. The results have demonstrated an improvement in the existing algorithm.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Quality-Aware Video Systems

"Quality in a product or service is not what the supplier puts in. It is what the customer gets out and is willing to pay for....Customers pay only for what is of use to them and gives them value...."

- Peter F Drucker

Consider a simple video broadcasting system as shown in figure 1.1. It shows an original video frame that is being broadcasted along with the video frames that are being received at different destinations. We can see that the quality of the videos received

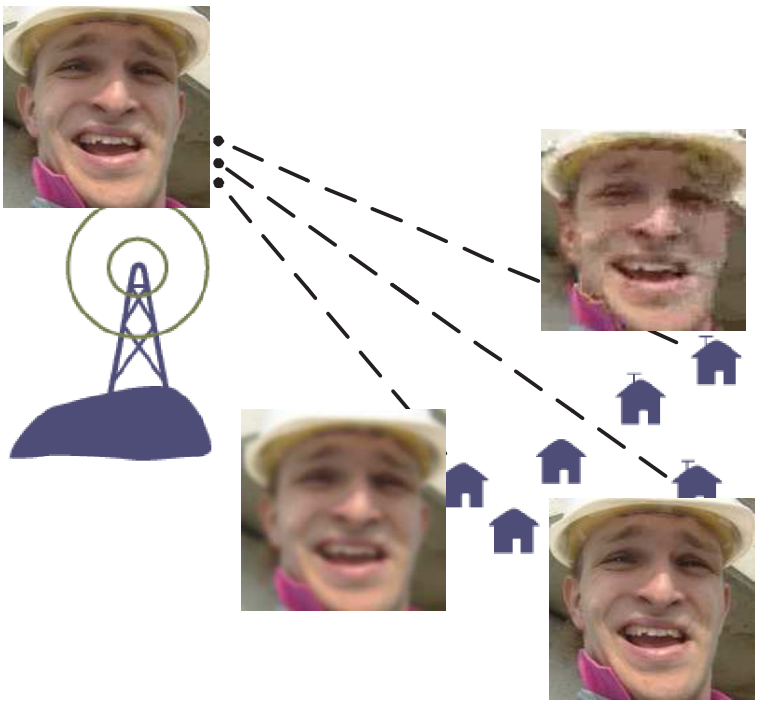

Figure 1.1. Illustration of broadcasting system.

is not same for all. The quality varies. Now consider pay-per-view system attached to

this broadcasting system. Since the same video for the same duration is being viewed by all viewers, all the destinations are billed with the same amount. However it is obvious that the quality of the video received is not the same and thus it is unfair to be billed by the same amount. Therefore there needs to be a system that considers the quality of the video being received by the viewer and bills accordingly. Such a system should be able to assess the quality of the video on the fly , *i.e.* the system should be dynamic. The system needs to have the knowledge of the video being transmitted from the service station and assess the quality of the received video accordingly. It is useful to have the knowledge of the original video because in cases where the original video is itself degraded due to various reasons, the viewers should not expect high quality video.

The example discussed above is one of the applications of the work to be presented *"Quality-Aware Videos"*. With the latest developments in technology, visual communication has grown and is more easily accessible. It has also become important to know the quality of data being transmitted and received. This has triggered a need for a system that is part of the communication system and dynamically monitor the quality. Quality-Aware video is one such way implementing the system that can serve the purpose.

1.2 NMR Spectrum Registration

"Another yet the same."

- Alexander Pope

The need for signal registration or signal alignment can be simply explained as "the comparison of two quantities is valid only if they are raised to the same platform and similar physical conditions...". Variations in physical conditions are inevitable either they are human or machine born. It is important to have a buffering system that compensates for these unwanted differences and provides a platform for comparison. Our work on high resolution Nuclear Magnetic Resonance (NMR) spectrum alignment is an attempt to provide solutions to such a system. Automatic diagnosis or computerized diagnosis has been an area of interest for many researchers now. It is gaining importance with the

advancements in the technology and growth in demand of efficient and faster diagnosis. NMR spectral analysis has recently become one of the major means for detection and recognition of metabolic changes of disease state, physiological alteration, and natural biological variation [1]. Since NMR data is exposed to human and machine born varying conditions, for better analysis of the spectra, there is a need for alignment or registration methods that buffers these varying conditions.

In the next chapter we will discuss the general quality assessment techniques and then advance to more complex methods, elements and development of Quality-Aware Video systems and illustrate the implementation with some experimental results of the QAV system. In chapter 3 we will discuss pyrmaid based NMR spectrum registration.

CHAPTER 2

QUALITY-AWARE VIDEO

2.1 Background

In any visual communication, along with the need to ensure successful exchange of data, it is necessary to monitor the quality of the content being exchanged. Quality monitoring also becomes necessary to analyze the performance of the communication techniques, channels, modes, and communication standards being used. In some cases (pay-per-view), quality needs to be monitored to have a better and proper estimate of the content being viewed by the customers. These factors have triggered a major research area of visual quality assessment. With the developments in the quality assessment techniques, the applications have been growing. Thus, the need for quality assessment and the techniques for the same have become complimentary to each other. Development in one triggers the other.

Quality is subjective and purposeful. These have made quality assessment an interesting and challenging task. Subjective quality assessment has been studied for years and interesting models have been developed. But the need for an objective quality assessment technique has been emphasized every now and then. This section will first look at the growth of quality assessment techniques in the simple and basic form of visual content, the images. Then extend the concept to more complex form, the videos. Our work is an effort to contribute to the latter part, quality assessment of videos.

2.1.1 Image Quality Assessment

One of the widely used objective quality assessment is Mean Square Error (MSE).

$$MSE = \frac{1}{M \times N} \sum_{i,j} (X_{i,j} - \overline{X}_{i,j})^2 \tag{2.1}$$

where X (Original image) and $\overline{X}$ (distorted image) are assumed to be of size M x N. Its straight-forward and simple implementation made it widely accepted and has found many application [2], [3], [4]. However, MSE gives a numeric estimation of the differences in the images without considering the visual perceptive quality of the images [5]. Moreover, it highlights the need of the original image, which leads to another classification of quality assessment techniques that will be discussed later.

Another frequently used measure is Peak Signal-to-Noise Ratio (PSNR) given by

$$PSNR = 10log_{10}\frac{L^2}{MSE} \tag{2.2}$$

where L is the range of allowable pixel intensities of an image. For $8bits/pixel$ grey scale image, $L = 2^8 - 1 = 255$.

Occasionally both of the above measures fail to correlate with the perceived quality of the images. This failure motivated the development of new quality assessment methods proposed in [5], [6], and [7]. In these proposed methods, Structural Similarity (SSIM) [7] has proved to be consistent and reliable.

The SSIM was based on the fact that natural images are highly structured. Their pixels exhibit strong dependencies, especially when they are spatially proximate, and these dependencies carry important information about the structures of the objects in the visual scene [7]. The SSIM index is defined as,

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{2.3}$$

where C_1 and C_2 are small constants to avoid instability, x and y are two local image patches, μ_x and μ_y are the mean intensities of x and y respectively, σ_x and σ_y are the standard deviations of x and y respectively, and σ_{xy} is the correlation coefficient between x and y [7].

If we look at the formulation of the methods discussed above, they require access to the original image source for the computation of the measure. This allows us to classify quality assessment methods based on the need of reference (original) image. All

the methods discussed above can be classified as full reference (FR) quality assessment methods. The advantage of having access to the original image can be used to built a better and more reliable quality metric. The metric can be made to be consistent with the subjective quality evaluations. But the need of access to the original image limits its applications only to test rooms, making it impractical for real time scenarios.

No reference (NR) quality assessment also called blind quality assessment methods require absolutely no access to the original source. Such a method is an ideal one for any system, but zero access to original source makes it difficult to implement such algorithms. There are several NR based methods in the literature that are based on the following constraints [8], [9], [10], [11], [12] and [13]; (1) they are limited to specific kinds of distortions, (2) prior information of distortion is assumed. These constraints limit their usage and a generalized solution is difficult to achieve. Moreover, in real time scenarios the knowledge of distortions is not available unless a specific system is always dealt with.

The non-applicability of FR methods in real time and failure of NR methods to generalize inspires us to look at the solutions that would lie in between these two methods, that is a reduced reference (RR) quality assessment method. RR methods provide a tradeoff between FR and NR methods. As the name suggests, the method does not require full access to the original image but requires some features of the original image to assess the quality, which can be made available by various means. So these methods would rely on methods to extract quality features of the original image and methods to make them available to the quality assessing system. Such methods make quality assessment easier compared to NR methods but bring in an extra load of quality features to be transferred.

The standard deployment of a RR method requires the side information to be sent through an ancillary data channel [14]. The need of an additional channel restricts their applications.

2.2 Quality-Aware Images

A novel concept of *quality-aware images*(QAI) was introduced in [14]. The idea was simple and straight-forward. The whole idea was to sent the extracted quality features as hidden messages. The method was based on three algorithms; (1) an algorithm to extract quality features of the original image, (2) a data embedding algorithm to hide these features in the original image, (3) a quality measuring algorithm that would give a quality score based on the extracted original quality features and the quality features of distorted image.

The advantages of this approach are: [14]

- It uses an RR method that makes the image quality assessment task feasible(as compared to FR and NR methods).
- It does not affect the conventional usage of the image data because the data hiding process causes only invisible changes to the images.
- It does not require a separate data channel to transmit the side information.
- It allows the image data to be stored, converted and distributed using any existing or user-defined formats without losing the functionality of "quality awareness", provided the hidden messages are not corrupted during the lossy format conversion.
- It provides the users with a chance to partially "repair" the received distorted images by making use of the embedded features.

Since the purpose of the embedding algorithm is not to provide data security, the above said application broadens the non-security applications of data embedding algorithms. The work on QAI [14] discusses several applications and advantages of the method.

2.3 Quality-Aware Videos

The concept of QAI can be extended to videos. The idea here is to develop an algorithm that can extract reduced reference quality features of the host video, develop an algorithm that can embed these features into the video and decode them at the receiving end along with the method that can access the quality based on these features

and the features of the distorted video. Videos have the advantage of huge data as compared to images which can be well explored for the development of data hiding and decoding algorithms. However, the increase in the amount of data means the increase in the number of quality features which results in more data to embed and decode. Videos bring in several additional features along with them that allows flexibilities to experiment with several data hiding techniques. The data hiding algorithm should not affect the visual quality of the video, as the very purpose of QAV would be hampered. The other important algorithm in any QAV system is the extraction of the quality features of the original video for the implementation Reduced Reference Video Quality Assessment (RR VQA). Extracting features of a video that are sensitive to most of the natural attacks is a challenging task. Literature survey shows that RR VQA [15] is a less explored area and very few attempts have been made to implement RR VQA.

Our current work on *Quality-Aware Videos* is an extension of the work on QAI. The successful implementation of QAI and its various applications motivated us to experiment the concept on videos. Assuming a video sequence to be a set of pictures arranged sequentially, the concept of QAI can undoubtedly be extended to videos. Our work demonstrates the successful extension and implementation of the concept to videos.

2.4 Reduced Reference Video Quality Assessment

Here we briefly review the RR VQA method for quality assessment (for details see [16]). The RR VQA is based on a statistical model of *temporal motion smoothness* in the complex wavelet transform domain [16]. Here, for simplicity we consider 1D data which can later be extended to higher dimensions. Ideally a time varying image sequence (video) can be created from the static image $f(x)$ with rigid motion and constant variations of contrast and average intensity:

$$h(x,t) = a(t)f[x + u(t)] + b(t)\,. \tag{2.4}$$

where $u(t)$ indicates how the image positions move spatially as a function of time, $a(t)$ and $b(t)$ are both real and account for the time-varying contrast and luminance changes, respectively. Applying a continuous complex wavelet transform of the form on equation 2.4

$$F(s,p) = \int_{-\infty}^{\infty} f(x)\, w_{s,p}^{*}(x)\, dx\,. \tag{2.5}$$

we get,

$$H(s,p,t) \approx F(s,p)a(t)\, e^{j(\omega_c/s)u(t)}\,. \tag{2.6}$$

Here $b(t)$ is eliminated because of the bandpass nature of the wavelet filters. The approximation is valid when the movement $u(t)$ is small compared to the width of the slowly varying window $g(x)$. Observing $H(s,p,t)$ at consecutive time steps $t_0 + n\Delta t$ for $n = 0, 1, ..., N$, it can be found that the following temporal correlation function is useful to test the (N-1)-th order temporal motion smoothness:

$$L_N(s,p) = \sum_{n=0}^{N} (-1)^n \binom{N}{n} \log H(s,p,t+n\Delta t)\,. \tag{2.7}$$

It can be shown that when the motion is (N-1)-th order smooth, the imaginary part of $L_N(s,p)$ is zero. Real natural images are expected to depart from these assumptions. However, by looking at the statistics of the imaginary part of $L_N(s,p)$, it is possible quantify such departure and use it as an indication of the strength of temporal motion smoothness.

Given a video sequence, it is divided into groups of pictures (GOPs) and each image frame is decomposed independently into subbands using the complex version [17] of the steerable pyramid decomposition [18]. $Ł_2(s,p)$ is computed for all coefficients within the subbands. A histogram of the imaginary part of $Ł_2(s,p)$ is obtained and is shown in Fig. 2.1. The top row shows the histogram for a GOP of a original video sequence, and the second and the third rows represent the corresponding video GOPs distorted by Gaussian noise and Gaussian blur respectively. As we can see, for the distorted video there is departure in distribution. A high peak at zero is observed, demonstrating a

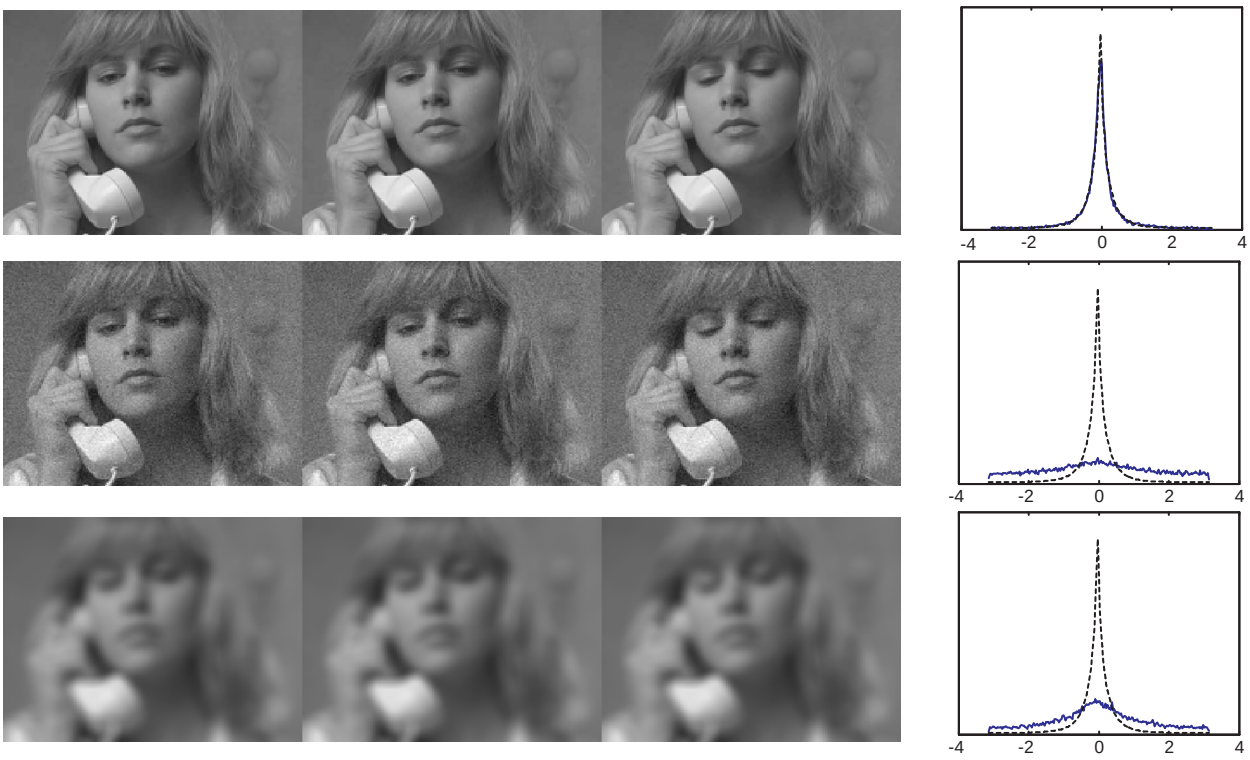

Figure 2.1. RR video features: demonstration.

strong prior of temporal motion smoothness. The histogram of each GOP, can be well fitted with a four-parameter function [16]:

$$P(\theta) = \frac{1}{Z}\left\{\exp\left[-\left(\frac{|\sin[(\theta-\theta_0)/2]|}{\alpha}\right)^{\beta}\right] + C\right\}, \tag{2.8}$$

where $\theta \in [-\pi, \pi]$, Z is a normalization constant, and the four parameters θ_0, α, β and C controls the center position, width, peakedness and the baseline of the function, respectively. The fitting process is optimized to minimize the Kullback-Leibler distance (KLD) [19] between the model and the observed distributions (denoted as p_m and p, respectively). The KLD between the model and the distorted video distributions, $d(p_m \parallel q)$ at the receiving end is estimated by

$$\hat{d}(p \parallel q) = d(p_m \parallel q) - d(p_m \parallel p)\,, \tag{2.9}$$

which can be easily shown to be a close approximation of $d(p \parallel q)$. The four fitting parameters θ_0, α, β and C, together with the KLD between the fitting model and the true

distribution [denoted as $d(p_m \parallel p)$], are included as RR features. These four parameters give us 16 scalar features for each GOP.

2.4.1 Fitting Models

Figure 2.2. Original video attacked by Gaussian blur.

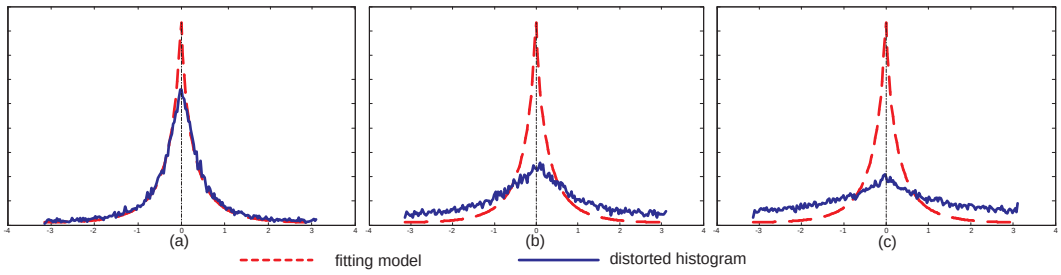

Figure 2.3. Curve fitting for the video in figure2.2.

Here we demonstrate the departure of the histogram from the fitting model. Figures 2.2 and 2.3 show the distorted videos and the histograms for the Gaussian blur. Figures 2.4 and 2.5 show the distorted videos and the histograms for the Gaussian noise. From all the four figures above, we can see that the histogram departs away from the original model as the strength of attack increases. Moreover, we can see from Figures 2.3 and 2.5 that the departures are different for the different kind of distortions.

Figure 2.4. Original video attacked by Gaussian noise.

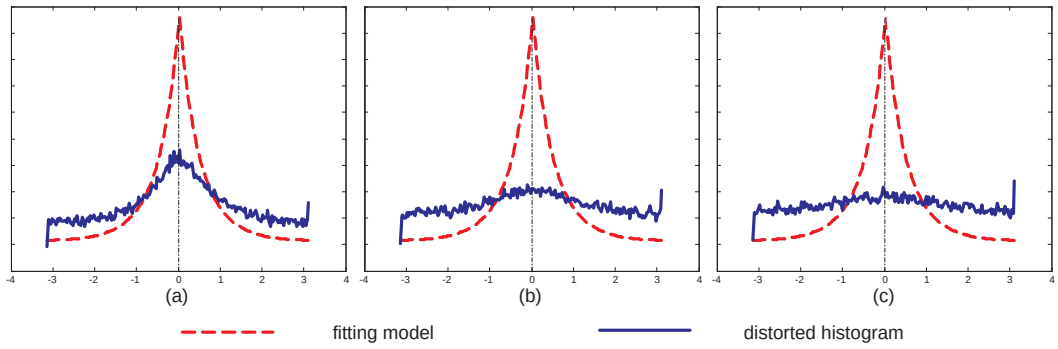

Figure 2.5. Curve fitting for the video in figure2.4.

2.5 Information Embedding

Data embeding and data decoding are two important processes in the implementation of QAV system. The successful implementation of QAV system depends on the efficiency of the algorithm to embed data without making perceptual changes to original video and decode the embedded data from distorted video. Two general requirements of an ideal data hiding algorithm are:

- It should be robust enough to sustain all natural distortions.
- It should not introduce any perceptual changes in the original video.

Any perceptual distortions introduced by the data hiding algorithm will hamper the purpose of the QAV system. It is expected that the quality score of the original video and the quality aware video to be similar. The data hiding algorithm employed in the

development of QAV system is based on 3D Discrete Cosine Transform (3D DCT) and Quantization Index Modulation (QIM).

2.5.1 Discrete Cosine Transform

Discrete Cosine Transform is an orthogonal transform and over the years DCT has found many applications in signal and image processing. The 1D DCT is given by,

$$X(k) = \omega(k) \sum_{n} x(n) \cos \frac{\pi(2n-1)(k-1)}{2N} \tag{2.10}$$

$$\omega(k) = \begin{cases} \frac{1}{\sqrt{N}} & k = 1 \\ \sqrt{\frac{2}{N}} & 2 \leq k \leq \text{N} \end{cases}$$

where $x(n)$ is the original sequence, $X(k)$ is the DCT of $x(n)$ and $k = 0, 1, 2...N$. The 2D DCT is given by

$$X(p,q) = \alpha_p \alpha_q \sum_{m} \sum_{n} x(m,n) \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N} \tag{2.11}$$

$$\alpha_p = \begin{cases} \frac{1}{\sqrt{M}} & p = 0 \\ \sqrt{\frac{2}{M}} & 1 \leq p \leq \text{M-1} \end{cases}$$

$$\alpha_q = \begin{cases} \frac{1}{\sqrt{N}} & q = 0 \\ \sqrt{\frac{2}{N}} & \text{q} \leq q \leq \text{N-1} \end{cases}$$

2.5.1.1 Energy Compaction

One of the important properties of DCT explored in our data hiding algorithm is its energy compaction. DCT transform has been proved to have very good energy compaction property. The transform forces the energy to be concentrated at the lower frequencies. The coefficients at the lower frequencies have higher magnitude as compared to those at mid and higher frequency ranges. This property has been well explored in compression techniques [20], [21], [22]. It is important to note here that the uncorrelated

data has its energy spread out, whereas the energy of the correlated data is packed into the low frequency region. DCT renders excellent energy compaction for correlated data. It is safe to assume here that the natural video frames are highly correlated and since we deal with 3D DCT, we will have a better energy compaction.

2.5.1.2 Separability

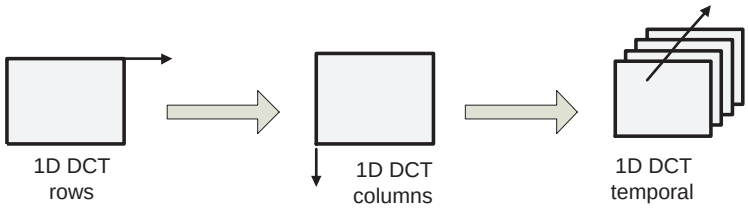

Figure 2.6. Implementation of 3D DCT.

The DCT in equation 2.11 can be expressed as in equation 2.12 using separability property of DCT,

$$X(p,q) = \alpha_p \alpha_q \sum_m x(m,n) \cos \frac{\pi(2m+1)p}{2M} \sum_n \cos \frac{\pi(2n+1)q}{2N} \tag{2.12}$$

This property becomes important as 3D DCT can be implemented by implementing 1D DCT operation on rows, columns and then temporally on a group frames. The idea is illustrated in figure. 2.6

Figure 2.7 shows few layers of the 3D DCT. As we can see that due to energy compaction, layer 1 has higher magnitude coefficients close to zero. As we move to the subsequent layers (layer1: top left, layer2: top right,...,layer8: bottom right), the magnitude reduces and the energy moves away from the zero point. After statistically analyzing the 3D DCT transforms, a group of DCT coefficients close to lower frequencies is selected in different layers of 3D DCT and then a key generated as the positions of these coefficients. These locations are used to embed data using the Quantization Index Modulation (QIM) method [23]. Coefficients close to lower frequencies are selected

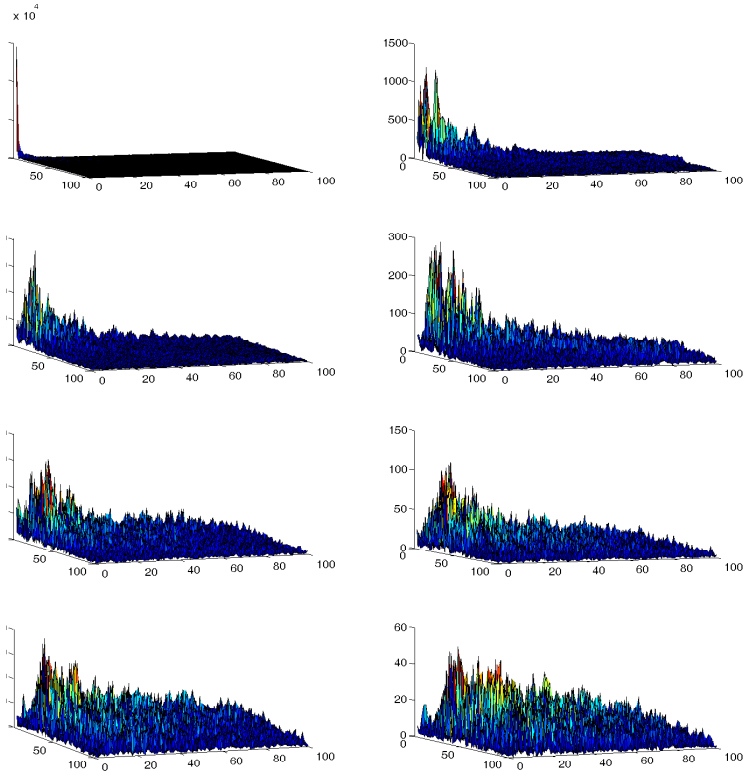

Figure 2.7. 3D DCT layers.

because of their high magnitude so that quantization of these values does not affect the perceptual quality of the video.

2.5.2 Quantization Index Modulation

Quantization Index Modulation (QIM) was introduced in [23] for digital watermarking. QIM involves embedding information by first modulating an index or a sequence of indices with the embedded information and then quantizing the host signal with the associated quantizer or a sequence of quantizers. QIM methods allow for blind decoding (decoding does not require the access to the reference image) of the embedded information.

A single bit of information ($m \in 0, 1$) can be embedded into a selected coefficient by the rule,

$$c_q = Q(c + d(m)) - d(m) \equiv Q^m(c). \tag{2.13}$$

where c_q is the marked coefficient, $Q(\cdot)$ is the base quantization operator with step size $\triangle$, and $d(m)$ is a dithering operator given by

$$d(m) = \begin{cases} -\Delta/4, & \text{if } m = 0 \\ \Delta/4, & \text{if } m = 1 \end{cases} \tag{2.14}$$

Figure 2.8 illustrates the process of embedding a single bit in the coefficient. The

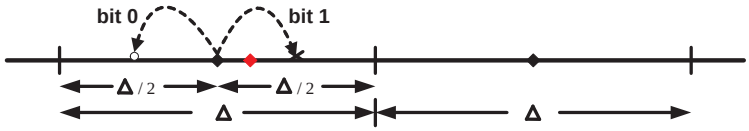

Figure 2.8. QIM for a single bit..

value of $\triangle$ can be tuned such that there is a tradeoff between the robustness of the data hiding algorithm, data embedding rate and imperceptibility. Figure 2.9 shows the

141490	**6625**	**11575**	**170**	**3269**	**-2113**	**-2485**	**-1216**
18922	**-8920**	**-4659**	**-890**	**-2810**	3465	596	-821
9504	**6235**	**-7869**	**-2851**	**-4151**	-3234	783	1329
2206	**6327**	**-5820**	**-2313**	**7631**	-2943	-1447	104
3091	**-8772**	**-3842**	**-1319**	5653	6583	-249	1959
-1795	**1114**	**-1367**	**1740**	-2658	934	273	955
-1946	**1061**	**3477**	-238	-1545	-812	-189	-3087
-354	**1859**	5058	-997	-349	-2137	-1527	1025
-2126	1211	17	722	-3292	1555	-647	
50	-2387	-159	-2008	-364	183	1820	

DCT coefficients before QIM

141490	**6563**	**11513**	**188**	**3263**	**-2138**	**-2588**	**-1163**
18938	**-8813**	**-4688**	**-863**	**-2888**	3465	596	-821
9488	**6338**	**-7763**	**-2813**	**-4238**	-3234	783	1329
2213	**6263**	**-5888**	**-2288**	**7688**	-2943	-1447	104
3113	**-8663**	**-3938**	**-1388**	5663	6583	-249	1959
-1838	**1013**	**-1388**	**1763**	-2658	934	273	955
-1913	**1088**	**3413**	-238	-1545	-812	-189	-3087
-354	**1838**	5058	-997	-349	-2137	-1527	1025
-2126	1211	17	722	-3292	1555	-647	
50	-2387	-159	-2008	-364	183	1820	

DCT coefficients after QIM

Figure 2.9. DCT coefficients before and after QIM.

9	34	197	255	247	254	254	254
9	32	196	255	247	254	254	254
8	29	201	255	165	132	137	133
8	31	208	239	121	168	255	254
9	35	210	234	130	137	132	168
9	37	196	255	247	172	132	128
9	38	201	255	243	255	223	124
9	34	209	235	130	133	144	124
8	41	203	255	247	254	254	254
15	36	106	167	197	208	214	218

pixel values before QIM

9.421	34.45	197.5	255.6	247.7	254.9	255	255.2
9.406	32.44	196.5	255.6	247.7	254.9	255	255.2
8.374	29.41	201.5	255.6	165.7	132.8	138	134.2
8.328	31.36	208.4	239.5	121.6	168.8	255.9	255.1
9.267	35.3	210.4	234.5	130.6	137.7	132.9	169
9.192	37.22	196.3	255.4	247.5	172.6	132.8	129
9.104	38.14	201.2	255.3	243.4	255.6	223.7	124.9
9.004	34.04	209.1	235.2	130.3	133.5	144.6	124.8
7.892	40.93	203	255.1	247.2	254.3	254.5	254.7
14.77	35.8	105.9	167	197.1	208.2	214.4	218.5

pixel values after QIM

Figure 2.10. pixel values before and after QIM.

DCT coefficients before and after the QIM and Figure 2.10 shows the pixel values before and after the QIM. We can see that there is no much difference between the original pixel values and the quantized pixel values. Thus QIM has the advantage of embedding data without making significant visible changes in the host and hence well suited for our application.

Once embedded, the data can be easily retrieved from the same but distorted coefficients using the minimum distance criterion,

$$\hat{m}(c_d) = \arg\min_{m\in\{0,1\}} \|c_d - Q^m(c_d)\| . \tag{2.15}$$

where $\hat{m}(c_d)$ is the decoded data bit, c_d is the distorted coefficient. The simplicity and robustness of QIM algorithm made it a favorable method in our application.

2.6 Quality-Aware Video System

Similar to quality assessment of images, the algorithms for the quality assessment in videos can also be classified into full reference (FR), no reference (NR) and reduced reference (RR) methods. The limitations of FR and NR exposed in images are visible in videos too. Literature survey shows several FR and NR implementations.

FR metrics are expected to provide more accurate quality measurements as they have access to the complete source. It is also expected that they give better results to all kinds of distortions and provide better generalization. As discussed in the section 2.1.1 of chapter 2 the existing NR methods, which do not assume any knowledge about the reference video require a prior knowledge of the distortion under consideration. They are limited to specific types of distortions [8], [9], [24], [25], [26], [27], [28], [29], [30], e.g., blocking artifacts created in block-based video compression. As it is true to images, in real time scenario the knowledge of distortion is not available to the quality assessment systems which limits the applications of NR methods. [25], [31], [10] and [13] proposed NR quality assessment methods based on digital watermarking. A pseudo random bit sequence or a watermark image is hidden inside the original video. The bit error rate or the degradation of the watermark image measured at the receiver side is then used as an indication of the quality degradation of the original video. Strictly speaking, these methods are not VQA methods because no extracted features about either the reference or the distorted video are actually used in the quality evaluation process. In essence, they do not assess the quality of the video but the quality degradation of the watermark embedded in the video.

2.6.1 Basic Structure of QAV Systems

Similar to a communication system, a basic QAV system will have a transmitter end and a receiver end. The transmitting end will extract the quality features and embed them. The receiving end will decode the original quality features, extract the quality

features of the distorted video and pass them to the quality assessing system, which in turn gives a quality measure.

The original video is passed through a transmitting block shown in Figure 2.11. It consists of a RR based quality feature extraction algorithm, and an encoder algorithm that will embed the extracted features into the original video. The data embedding is a key based algorithm which is shared with the receiving end.

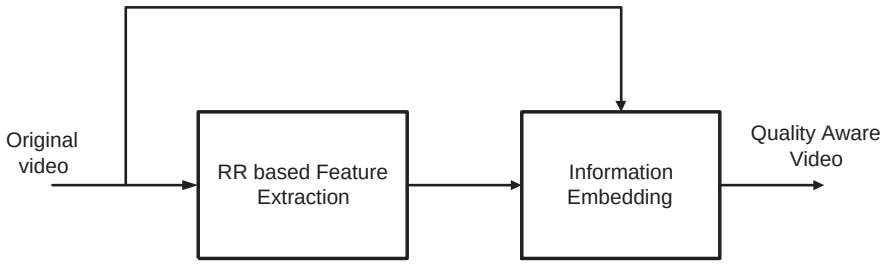

Figure 2.11. Transmitting end of QAV system.

The block diagram of the receiving end is as shown in figure.2.12. It receives a distorted QAV and gives a quality measure of the video. It consists of a RR based quality feature extraction algorithm similar to the one at the receiving end, a decoding algorithm that decodes the original video quality features from the distorted video, and a quality assessment algorithm which analyzes the quality features of both the videos and gives a quality score. Combining the methods discussed in previous sections, the

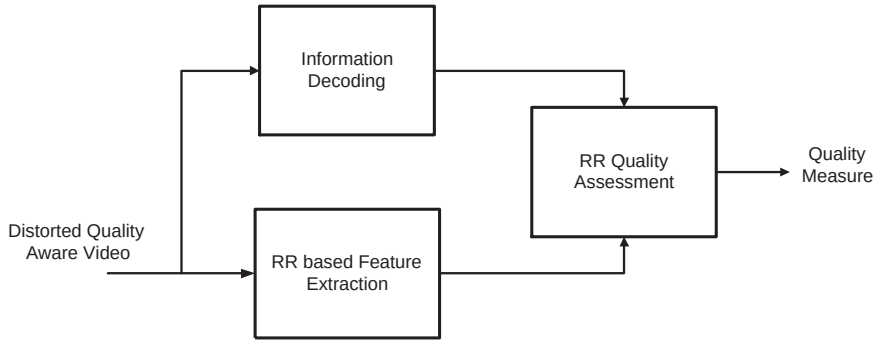

Figure 2.12. Receiving end of QAV system.

functionality of the QAV system can be summarized as follows. At the encoder end,

- Use RR VQA to extract quality features of the original video.
- Perform 3D DCT on the original video.
- Use QIM method to embed the extracted quality features in transform domain.
- Perform inverse 3D DCT to obtain QAV.

At the decoder end,

- Use RR VQA to extract quality features of the distorted QAV.
- Perform 3D DCT on the distorted QAV.
- Use QIM decoding algorithm (minimum distance criterion) to decode the original quality features.
- Use Quality Assessment algorithm to analyze the features of the original and the current video and compute a quality score.

To improve the robustness of the data embedding algorithm, error correcting codes can be used. We have employed BCH (15, 2, 7) [32] code, which can correct up to 2 errors in every 7 bits decoded.

2.6.2 Implementing QAV System

The original video is divided into groups of pictures (GOPs) with 8 frames in each. The 16 scalar features obtained in section 2.4 are then encoded using a 7 bit binary representation to obtain a total of 112 (16 x 7) bits of information. Now these 112 bits are to be embedded in the first 8 frames of the video. To improve the robustness, we introduce BCH error correcting codes as in [14]. The BCH used here is (15, 2, 7) that can correct 2 bits for every 7 bits and the encoded bit length is 15. After BCH coding, the total length of bits to be embedded is 240 (16 x 15). To embed 240 bits in a GOP, we need to select 240 DCT coefficients. It is important to select higher magnitude coefficients as we do not want to make any perceptible changes in the original video.

3D DCT is performed using the separability property of DCT. First a 2D DCT is performed on all the 8 frames (in a GOP) individually. Then 1D DCT is implemented

temporally by selecting the 8 corresponding elements one from each layer of 2D DCT to obtain 3D DCT. By doing this we obtain 8 layers of 3D DCT with the coefficients in the first layer closer to the origin representing lower frequency components and have higher magnitude (due to energy compaction as shown in figure 2.7). This is true because natural videos are highly correlated both spatially and temporally. As we move down the layers, the energy (higher magnitude coefficients) shifts away from the origin along with the considerable drop in the energy. This is similar to the phenomenon observed in 2D DCT of the images. Understanding of this makes it easy to select coefficients to embed data bits. Since 240 bits are to be implemented, and we do not want to make any perceptible changes to the video, we distribute the selection of coefficients into first few layers. This location of bits is the key to be used both in the encoding and decoding algorithms. As expected, the position of the selected coefficients shifts away from the origin as we move down the subsequent layers.

Once the key is determined, we use QIM to embed a single data bit in a single coefficient using Equation 2.13. The value of stepsize $\triangle$ is selected such that it does not make considerable change in coefficient so that the original video is altered but small enough to sustain distortions on the video. Once all the bits are embedded, we perform inverse 3D DCT first by performing inverse 1D DCT temporally and then inverse 2D DCT on the individual layers to obtain the QAV frames. This process is repeated to all the GOPs of the video to obtain QAV. The process is illustrated in Figure 2.13. As seen in the figure, we have implemented error correcting code (BCH code) to improve the robustness of the data hiding algorithm.

At the receiving end, we again use RR VQA to obtain the quality features of each GOP of the received distorted QAV. We perform 3D DCT and extract the embedded 240 bits from each GOP. After performing BCH decoding we are left with 112 bits of the original video quality features which are converted into 16 scalar RR features of the GOP. This is illustrated in Figure 2.14. These two sets of 16 scalar RR features are then passed to a quality assessment algorithm that analyzes the features and gives a distortion

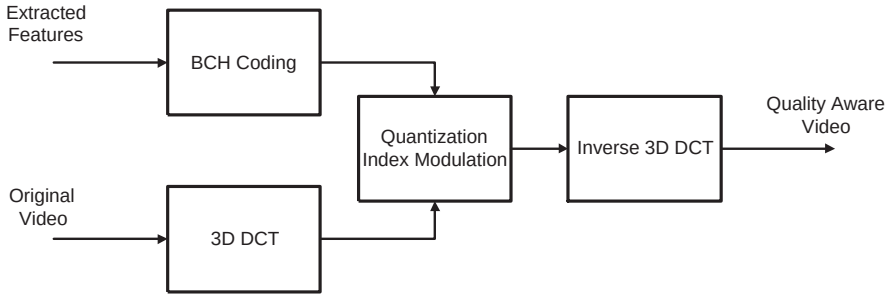

Figure 2.13. Illustration of data hiding algorithm.

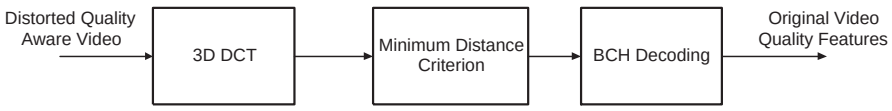

Figure 2.14. Illustration of data decoding algorithm.

measure. With the help of the original RR features, a fitting model is obtained. Then Kullback-Leibler distance (KLD) is computed between the model and the distorted QAV distributions. The distortion score is then obtained using the method discussed in the section 2.4.

2.7 Simulation Results

The algorithms were implemented on MATLAB using signal and image processing toolboxes. The results are analyzed under the following sections; Effect of data hiding on original video; Robustness of the data hiding algorithm; and Distortion measures for distortions. In all the sections we have considered two kinds of distortions;

- Additive White Gaussian Noise(AWGN)
- Gaussian Blur

2.7.1 Effect of Data Hiding

As a general requirement of any data hiding algorithm, it is important that it does not affect the perceptual quality of the host; otherwise the very purpose of the QAV system is hampered. Therefore it is important to analyze the effect of the data hiding algorithms and tune them to have a desirable effect. In our work, due to the selection

Figure 2.15. The original and the QAV frames.

of QIM to embed data, tuning is obtained by simply varying the step size of $\triangle$ in the QIM equation. It is important to note that the same step size should be used even during decoding the bits. The effect of data hiding is studied by comparing the quality scores of the original video and the QAV. We desire that the distortion measure for the videos should be similar. If so, then we can safely state that the data hiding introduces negligible perceptual change to the original video.

Figure 2.15 shows few frames of the original video and the corresponding QAV frames. The top row represents the original video and the bottom row represents the QAV. The frames are selected such that they represent different frames in a GOP (frame1, frame4 and frame8). We do not see any perceptual differences in the frames. After analyzing the results we can say that our data hiding algorithm does not introduce visible perceptual changes in the original video and is appropriate for QAV systems.

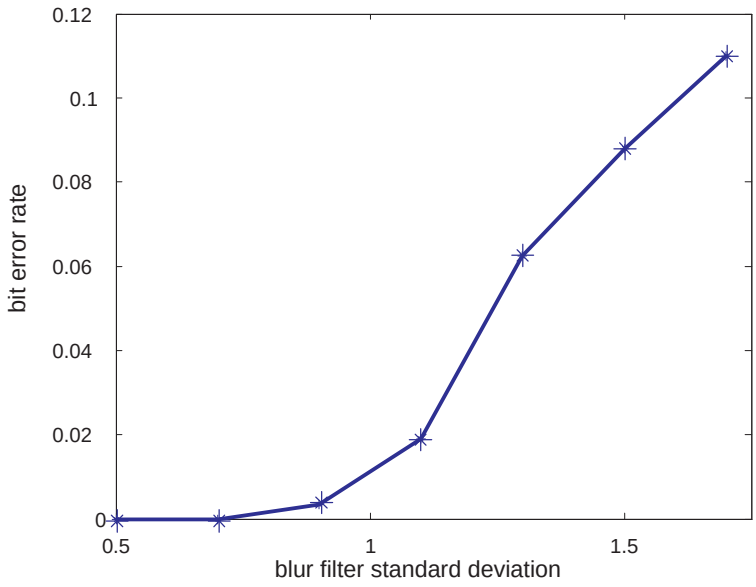

Figure 2.16. Bit error rate for Gaussian blur.

2.7.2 Robustness of Data Hiding Algorithm

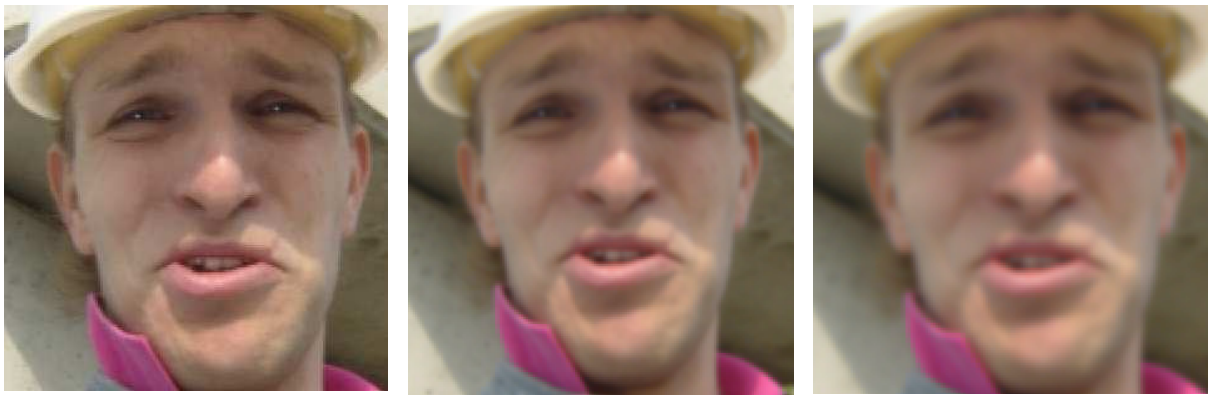

Figure 2.17. Video frames for Gaussian blur.

One more important feature of any data hiding algorithm is its ability to sustain different attacks. The robustness of the data hiding algorithm influences the applicability of the QAV system. We have studied the robustness of our algorithm and the results are shown in Figures 2.16, and 2.18. The corresponding distorted video frames are shown

in Figures 2.17[(a):σ=0.5, (b):σ=1.0, and (c):σ=1.5], and 2.19 [(a):var=20, (b):var=80, and (c):var=140]. The robustness of the algorithm is improved by employing BCH

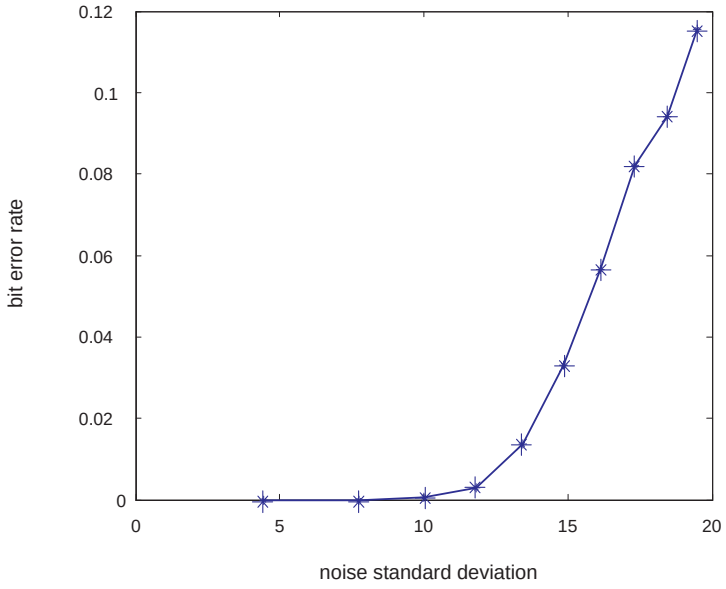

Figure 2.18. Bit error rate for Gaussian noise.

codes. The robustness can also be improved by increasing the step size $\triangle$ in the QIM. We select a appropriate value for $\triangle$ and then use BCH code as error correcting code to improve the robustness. By introducing BCH coding the number of bits to be embedded increases but this is not a critical issue as we are dealing with 3D DCT which operates on a GOP, which gives a large number of DCT coefficients with high magnitudes for QIM embedding. If a small perceptual change is allowed in the video, then the robustness can be improved by increasing the $\triangle$ value in QIM.

2.7.3 Distortion Measure

The purpose of implementing QAV system is to assess the quality of the video on the fly, so it becomes important to analyze the distortion measures of distorted videos. We expect that the distortion measure be monotonically increasing with the perceptual

Figure 2.19. Video frames for Gaussian noise.

Figure 2.20. QAV frames with Gaussian blur.

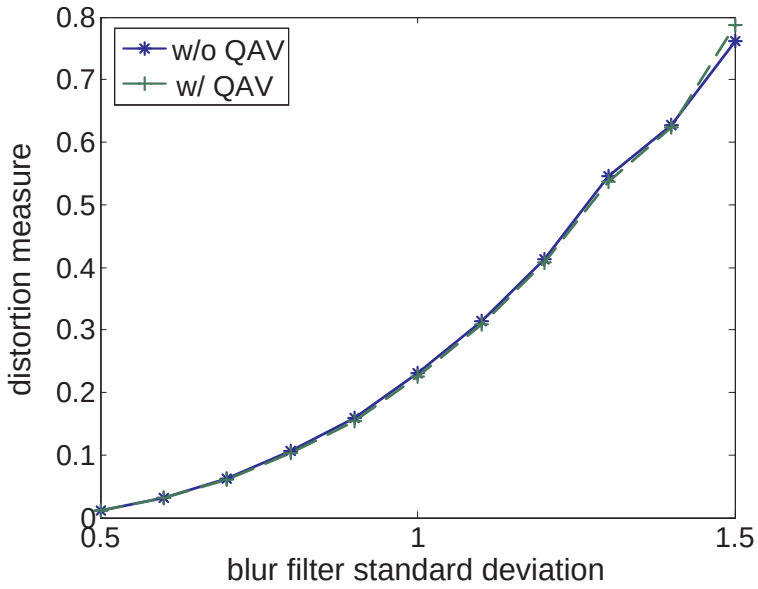

Figure 2.21. Distortion measure for Gaussian Blur distorted video.

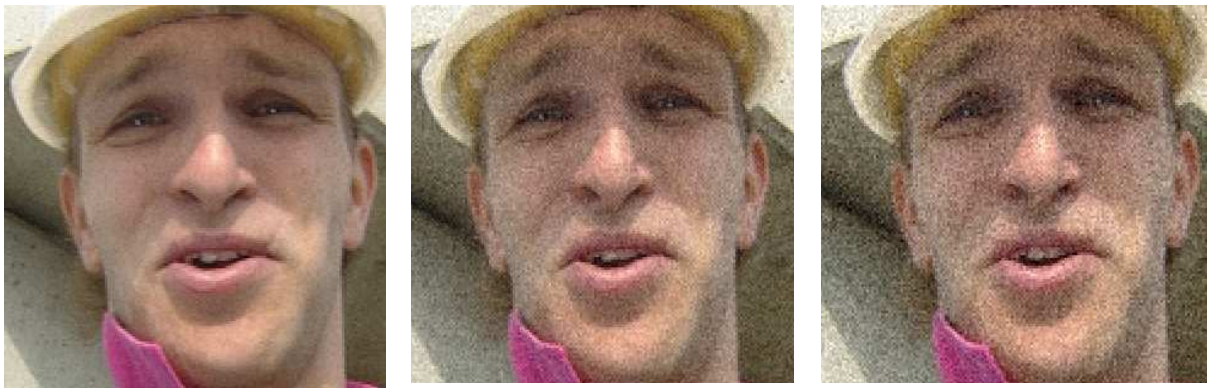

Figure 2.22. QAV frames with Gaussian noise.

distortion in the video and monotonically decreasing with the perceptual quality of the video. We also expect that distortion measures for different attacks follow the same trend. This makes sure that the increase in the strength of attacks increases the distortion measure and we see a trend in the reduction of the quality of the video. Figure 2.20 shows the QAV frames for the Gaussian blur and the distortion measure plot is shown in Figure 2.21.

The QAV frames for the Gaussian noise are shown in Figure 2.22 and Figure 2.23 shows distortion measure plot. The plots show two curves, one for the QAV and the other for the original video. The similarity suggests that there is negligible perceptual difference between the QAV and the original video. This is an important factor in deciding a data hiding algorithm. Distortion measures in both the attacks follow a same trend, suggesting that the distortion measure is dependent on strength of the attack. We can see the reduction in perceptual quality of the video with the increases in the distortion measure indicating the higher the distortion measure the lesser the quality.

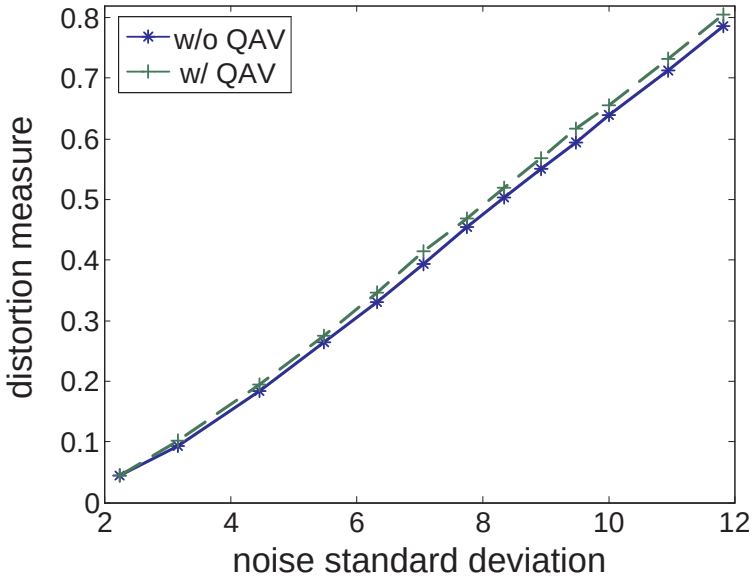

Figure 2.23. Distortion measure for AWGN distorted video.

CHAPTER 3

NMR SPECTRUM REGISTRATION

This work is an extension of the Bayesian spectrum alignment work in [1], which propelled us to extend the work. In the current work we apply a multi-scale or pyramid based approach to the existing Bayesian approach. In the next section, we look at the previous work and in the subsequent sections we introduce multi-scale approach and its application to the existing method.

3.1 Background

In [1], a new approach that simultaneously estimates the spectral shift and the baseline variations was introduced. It was based on Bayesian Least Square Estimation. The idea was to formulate the problem of registration in a Bayesian statistical framework with consideration of the effect of noise. A closed form solution obtained can be summarized as (see [1] for details):

Let $X(w)$ and $Y(w)$ be two spectral signals to be aligned, where w is the frequency index. If the two signals represent the same spectral structure but are shifted (in frequency and intensity) versions then we can write

$$y(w) = x(w + \triangle w) + \triangle a, \tag{3.1}$$

where $\triangle w$ and $\triangle a$ are spectral and intensity variations respectively. Applying a Taylor series expansion and ignoring the higher order terms, we can write Equation 3.1 as

$$\mathbf{y} = \mathbf{x} + \triangle \mathbf{x}' + \triangle a \mathbf{1}. \tag{3.2}$$

where $\mathbf{x} = [x(w_1), x(w_2), ..., x(w_N)]^T$, $\mathbf{y} = [y(w_1), y(w_2), ..., y(w_N)]^T$,
$\mathbf{x'} = [\frac{dx}{dw}|_(w_1)), \frac{dx}{dw}|_(w_2)), ..., \frac{dx}{dw}|_(w_N))]^T$, and $\mathbf{1}$ is an N-dimensional column vector of ones. In matrix format,

$$\mathbf{Ac} = \triangle \mathrm{x}. \tag{3.3}$$

where $\mathbf{A} = [\mathbf{x'}\ \mathbf{1}]$, $\triangle$x = y - x, and $\mathbf{c} = [\triangle w\ \triangle a]^T$ is a column vector containing the parameters we would like to estimate. The least square solution is then given by

$$\widehat{\mathbf{c}}_{LS} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T \triangle \mathrm{x}. \tag{3.4}$$

In certain cases, the matrix $(\mathbf{A}^T\mathbf{A})$ might be close to singular and inverting the matrix may be unstable.

The above problem was solved in [1] by introducing the noise effect in the formulation as,

$$g = \mathbf{Ac} - \triangle \mathrm{x}. \tag{3.5}$$

where g is modeled as a zero-mean Gaussian random vector. After certain mathematical assumptions and calculations, based on Bayes' rule

$$p(\mathbf{c}|g) \propto p(g|\mathbf{c})p(\mathbf{c}). \tag{3.6}$$

The Bayes least square (BLS) as well as the Bayes maximum a posterior (MAP) solution is given as,

$$\widehat{\mathbf{c}}_{BLS} = \widehat{\mathbf{c}}_{MAP} = (\mathbf{A}^T\mathbf{A} + \Lambda_n\Lambda_p^{-1})^{-1}\mathbf{A}^T \triangle \mathrm{x}. \tag{3.7}$$

where Λ_n is the noise variance and Λ_p is the covariance matrix.

The advantages of this approach are; the noise effect is considered; prior knowledge of the quantities being estimated can be included; and the addition of noise makes sure that the matrix $(\mathbf{A}^T\mathbf{A})$ is non-singular.

Since the algorithm discussed above is differentiation based, it works only if the variations in the source and the target are small. The algorithm breaks if the difference is large. For this approach to work, the obvious approach would be to reduce the differences

and then apply the Bayesian approach. By having multi-resolution or pyramids of the source and the target, the differences can be reduced considerably. Then by applying the Bayesian approach at different scales, a well registered signal can be obtained.

3.2 Pyramid based signal processing

The pyramid method for image processing was introduced in [33]. The idea is to have several copies of the image at increasing scale. The processing starts at the lowest scale, and is updated at subsequent scales until the original scale is reached. Their work discusses several applications of pyramid based processing. One application of interest is image registration. Here we try to explore the concepts and apply them to our problem of NMR spectrum registration.

3.2.1 Pyramids for signals

Consider a signal X of length N, the idea of the pyramid construction is to have multi resolution copies of X. The is illustrated in figure 3.1. The construction of the pyramids can be summarized as:

- Smooth the spectrum with a low pass filter.
- Reduce the size of spectrum such that $x'(n) = x(2n)$, $i.e.$,, downsample the signal by a factor of 2.
- Repeat the above steps until a maximal allowed number of levels or minimal allowed spectrum length is reached.

The pyramid constructed for two NMR spectra during simulation is as shown in Figure 3.1. The advantages of the pyramid approach are easily visible. As we move up the pyramid, the data gets shorted and smoother. We can see that the coarsest level has only the prominent features of the spectrum. The finer details are lost which makes alignment process more efficient. Moreover, the reduction in the spectral difference is also visible which is attributed to downsampling.

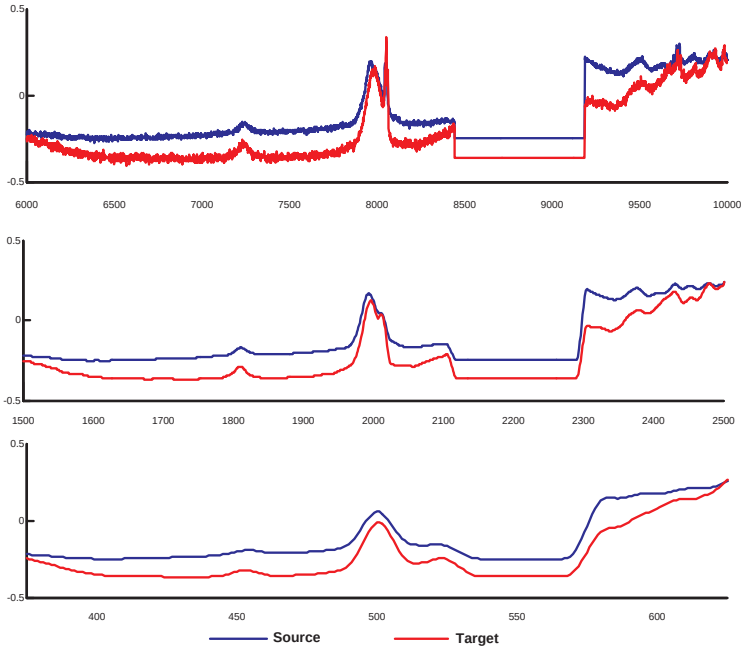

Figure 3.1. Pyramids for the source and target spectra.

3.3 Pyramid based Bayesian approach

The pyramid based Bayesian approach is shown in Figure 3.2 and it combines the Bayesian estimation discussed in section 3.1 of chapter 3. Let the two spectra under consideration be called the source and the target, where the source is the spectrum to be aligned to the target. The first step is to have the source and the target pyramids constructed to the required coarsest level. Let n be the number of levels in the pyramid with n being the coarsest level. Starting from n (coarsest level), alignment parameters are estimated. The estimated parameters are upsampled and the alignment is done at the $n-1$ level. Once the spectra is aligned, the spectra of the target at this level and the aligned spectra are used to estimate the new parameters. This process is repeated until we reach the finest level. Once the finest level is reached, we obtain the final alignment result.

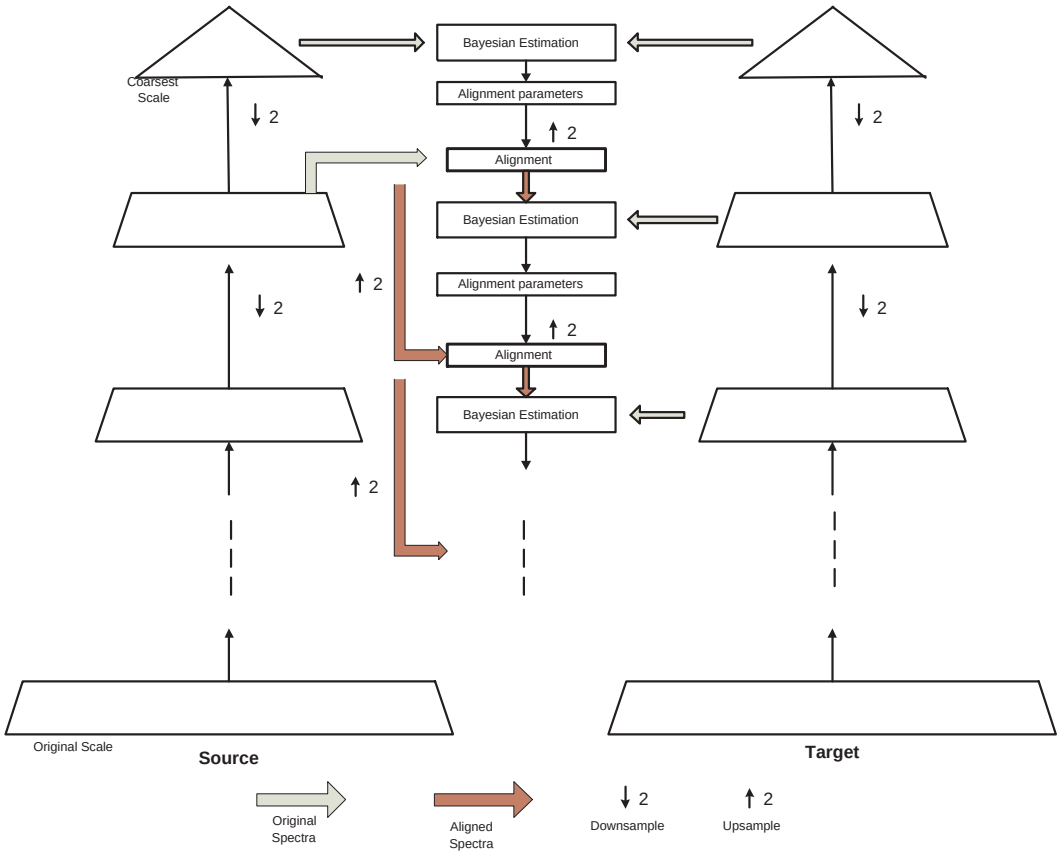

Figure 3.2. Illustration of pyramid based Bayesian approach.

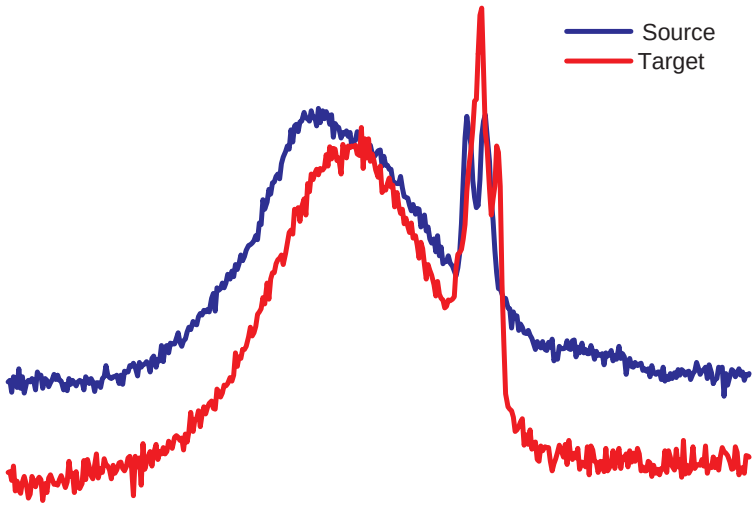

Figure 3.3. Section of the misaligned original source and target spectra.

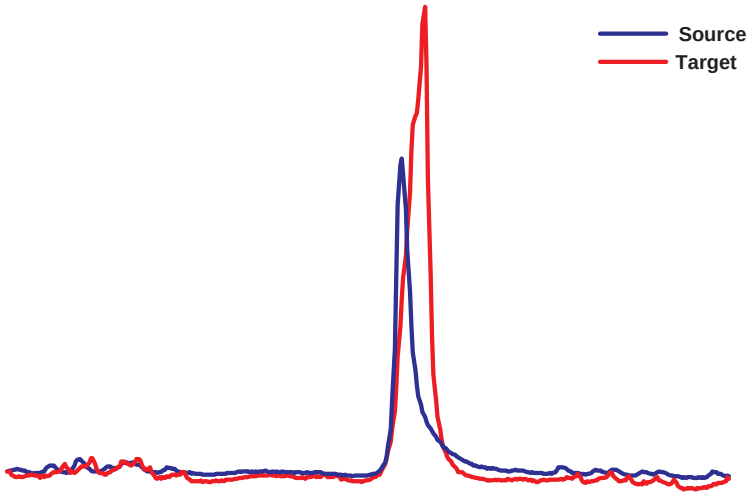

Figure 3.4. Section of the misaligned original source and target spectra.

3.4 Simulation Results

Figures 3.3 and 3.4 show the sections of the original source and target spectra which are misaligned. A pyramid structure similar to Figure 3.1 was constructed for both the spectra. The Bayesian estimation of parameters was done at the coarsest level and alignment at one upper level as explained before in the section 3.3. Figure 3.5 shows some of the intermediate results of the alignment. We can see that the peaks get aligned at the coarsest levels and it becomes easier to align the finer details as we move to the higher resolutions. Figures 3.6 and 3.7 show the alignment results of our algorithm. We can see the alignment in both the spectral and intensity variations.

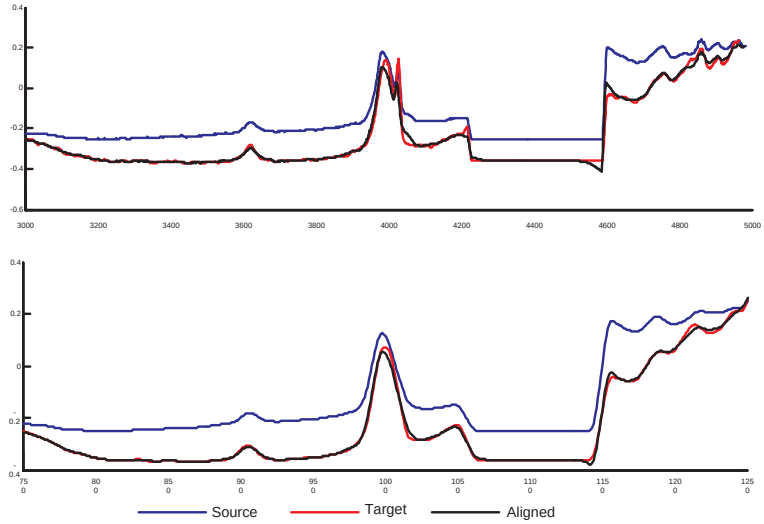

Figure 3.5. Intermediate results of pyramid based registration.

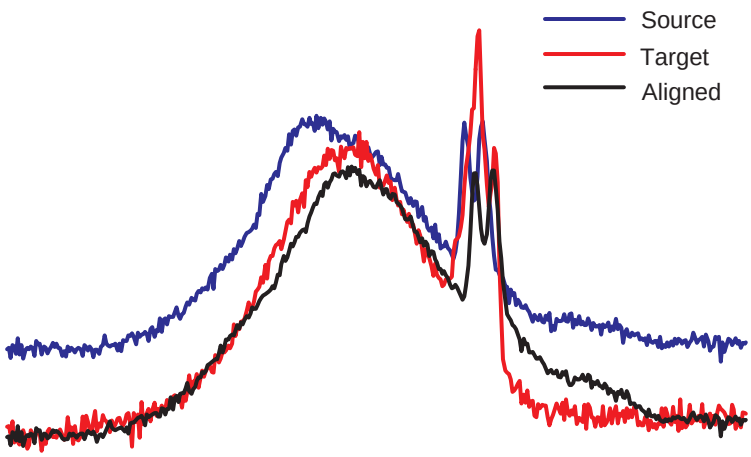

Figure 3.6. Aligned source with original source and target of figure 3.3.

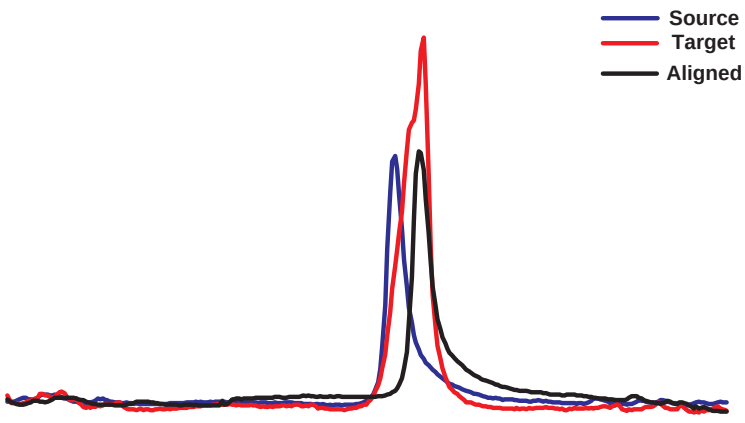

Figure 3.7. Aligned source with original source and target of figure 3.4.

CHAPTER 4

CONCLUSION

We have introduced a novel concept of *Quality-Aware Video* systems and demonstrated its implementation. The scope of the areas covered under this work is vast. It contributes in several directions: First, it introduces a novel concept of QAV system and provides a blueprint for its implementation. The concept can be well explored and different methods of implementing such a system can be developed. Contributions from areas of vision, perceptually quality and communications can enhance the efficiency of the system. Secondly, it develops a simple concept of extracting features of the original video. It emphasizes on the fact that knowing the quality of the original video is important to make any judgement on its distorted versions. The RR based extraction of the quality features is simple and straight-forward which at its core has a statistical model of temporal motion smoothness in the complex wavelet transform domain. Thirdly, it develops an algorithm to encode and decode data in the videos. With QIM at its core, the algorithm is able to operate without damaging the perceptual quality of the host video. The development of this algorithm widens the existing area of non-security applications of watermarking. With a key made available to all, the quality assessment can be done anywhere. And lastly, it talks about the dynamic system of quality assessment. As discussed in the chapter 1, the need for a dynamic system of quality assessment is fulfilled with QAV. The beauty of these RR based features is that they do not assume any prior knowledge of the distortion which means the quality of a QAV attacked by any natural distortion is assessable. This gives a hint of the potentials of the system to generalize to various attacks where many algorithms have failed to deliver. To the existing list, the advantages of the QIM can be added. The ability to provide varied tradeoffs between

the embedding rate, the perceptual quality and the robustness by the QIM can be well explored depending on the need and requirement of the system.

The proposed system also initiates some future work, in the development of robust data hiding techniques. The algorithm developed is sensitive to heavy compression algorithms which puts the hold on QAV system. With the developing technology, as more efficient and better compression algorithms are being built, it becomes necessary to develop a data hiding algorithm that overcomes these attacks. With such an algorithm, the scope of the QAV systems will widen. Alternatives for the method to extract RR features can also be studied. More RR based assessment methods could be tried and tested for better efficiency of the QAV systems. Applications of the QAV systems to error concealment, design of algorithms to partially repair the distorted QAV cannot be ruled out. The concept of "intelligent videos" wherein the videos not only carry the quality features but they sustain any attack on the communication network and correct itself during the course of transmission can be developed. This can be possible if the video knows its original quality and constantly monitor its quality on the fly. A small departure in the quality should trigger a correction algorithm before the error grows big.

The proposed NMR spectrum registration method overcomes the limitations of the existing differentiation based Bayesian estimation. It demonstrated that the existing Bayesian estimation approach implemented for small variations can be extended to larger differences. By involving the pyramid approach, it was possible to reduce the spectral differences which is critical in the application of the Bayesian estimation. The alignment of peaks at the low resolution makes it easier for Bayesian approach to work on the finer details at the higher resolutions. Moreover, it is shown that a reliable registration algorithm can be built even in the presence of the noise.

The existence of too many spikes in the spectrum limits the application scope of the algorithm. Presence of several "back-forth" differences in the spectra also affects the performance of the algorithm. In future, these can be overcome by employing preprocessing of the source and target spectra. A localized window method to globally estimate

the spectral and intensity shifts can be used as the part of preprocessing. Several options like, matching intensity difference at the coarsest level, splitting the spectrum in several bands before registering, knowledge of spiky areas in the spectrum can come handy and improve the performance of the system. In future, we propose to work on improving its performance and carry the concept forward to multi-dimensional data sets.

REFERENCES

[1] Z. Wang and S. B. Kim, "Automatic alignment of high-resolution nmr spectra using a bayesian estimation approach," in *International Conference on Pattern Recognition*, vol. 4. IEEE, 2006, pp. 667–670.

[2] P. William and M. Hoffman, "Error entropy and mean square error minimization for lossless image compression," in *International Conference on Image Processing*. IEEE, 2006.

[3] B. Denatale, G. Desoli, and D. Giusto, "Hierarchical image coding via mse-minimising bilinear approximation," *Electronics Letters*, vol. 27, pp. 2035–2037, Oct. 1991.

[4] F. Zampolo and R. Seara, "A comparison of image quality metric performances under practical conditions," in *International conference on image processing*, vol. 3. IEEE, 2005, pp. 1192–1195.

[5] Y. Lai, J. Li, and J. Kuo, "A wavelet approach to compressed image quality measurement," in *Signals, Systems and Computers*, vol. 2. IEEE, 1996, pp. 938–942.

[6] R. Langi, K. Soemintapura, T. Mengko, and W. Kinsner, "Multifractal measures of image quality," in *International conference on information, communications and signal processing*, vol. 2. IEEE, 1997, pp. 726–730.

[7] Z. Wang, A. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, pp. 601–612, June 2004.

[8] Z. Wang and A. C. Bovik, *Modern Image Quality Assessment*. Morgan and Claypool Publishers, 2006.

[9] L. Lu, Z. Wang, A. C. Bovik, and J. Kouloheris, "Full reference video quality assessment considering structural distortion and no reference quality evaluation of mpeg

video," in *International Conference on Multimedia and Expo.*, vol. 1. IEEE, 2002, pp. 61–64.

[10] P. Campisi, M. Carli, G. Giunta, and A. Neri, "Blind quality assessment system for multimedia communications using tracing watermarking," vol. 51, no. 4, pp. 996–1002, Apr. 2003.

[11] M. Masry, S. Hemami, and Y. Sermadevi, "A scalable wavelet-based video distortion metric and applications," *IEEE Transactions on circuits and systems for video technology*, vol. 16, no. 2, pp. 260–273, Feb. 2006.

[12] T. Mei, S. Hua, Z. Zhu, H. Zhou, and S. Li, "Home video visual quality assessment with spatiotemporal factors," *IEEE Transactions on circuits and systems for video technology*, vol. 17, no. 6, pp. 699–706, june 2006.

[13] M. Carli, M. C. Q. Farias, E. D. Gelasca, R. Tedesco, and A. Neri, "Quality assessment using data hiding on perceptually important areas," in *International Conference on Image Processing*, vol. 3. IEEE, 2005, pp. 1200–1203.

[14] Z. Wang, G. Wu, H. R. Sheikh, E. P. Simoncelli, E. Yang, and A. C. Bovik, "Quality-aware images," *IEEE Transactions on Image Processing*, vol. 15, pp. 1680–1689, June 2006.

[15] P. L. Callet, C. V. Gaudin, and D. Barba, "A convolutional neural network approach for objective video quality assessment," *IEEE Transactions on Neural Networks*, vol. 17, no. 5, pp. 1316–1327, sept 2006.

[16] B. Hiremath, Q. Li, and Z. Wang, "Quality-aware videos," in *International Conference on Image Processing*. IEEE, 2007.

[17] J. Portilla and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *Int'l Journal of Computer Vision*, vol. 40, no. 1, pp. 49–71, December 2000.

[18] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable multiscale transforms," *IEEE Trans Information Theory*, vol. 38, no. 2, pp. 587–607, March 1992, special Issue on Wavelets.

[19] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley-Interscience, 1991.

[20] W. B. Pennebaker and J. L. Mitchell, *JPEG: Still Image Data Compression Standard*. Springer, 1992.

[21] K. R. Rao, P. Yip, and V. Britanak, *Discrete Cosine Transform: Algorithms, Advantages, Applications*. Academic publishers, 2007.

[22] H. Wu and K. Rao, *Digital Video Image Quality and Perceptual Coding*. CRC, 2005.

[23] B. Chen and G. W. Wornell, "Quanitization index modulation: A class of provably good methods for digital watermarking and information embedding," *IEEE Transactions on Information Theory*, vol. 47, pp. 1423–1443, May 2001.

[24] F. Yang, S. Wan, Y. Chang, and H. R. Wu, "A novel objective no-reference metric for digital video quality assessment," in *IEEE Signal Processing Letters*, vol. 10. IEEE, 2005, pp. 685–688.

[25] F. Yang, X. Wang, Y. Chang, and S. Wan, "A no-reference method based on digital watermark," in *IEEE Proceedings on Personal, Indoor and Mobile Radio Communications*, vol. 3. IEEE, 2003, pp. 2707–2710.

[26] N. Montard and P. Bretillon, "Objective quality monitoring issues in digital broadcasting metworks," *IEEE Transactions on Broadcasting*, vol. 51, no. 3, pp. 269–275, sept 2005.

[27] C. Kuhmanch and C. Schremmer, "Empirical evaluation of layered video coding schemes," in *International Conference on Image Processing*, vol. 2. IEEE, 2001, pp. 1013–1016.

[28] R. A. Lotufo, W. D. F. DaSilva, A. X. Falcao, and A. C. F. Pessoa, "Morphological image segmentation applied to video quality assessment," in *International Symposium on Computer Graphics, Image Processing, and Vision*. IEEE, 1998, pp. 468–475.

[29] A. C. F. Pessoa, A. X. Falcao, A. E. F. Nishihara, and R. A. Lotufo, "Video quality assessment using objective parameters based on image segmentation," in *International Telecommunications Symposium*, vol. 2. IEEE, 1998, pp. 498–503.

[30] S. Olsson, M. Stroppiana, and J. Biana, "Objective methods for assessment of video quality: State of the art," *IEEE Transactions on Broadcasting*, vol. 43, no. 4, pp. 487–495, dec 1997.

[31] M. C. Q. Farias, M. Carli, and S. K. Mitra, "Objective video quality metric based on data hiding," *IEEE Transactions on Consumer Electronics*, vol. 51, no. 3, pp. 983–992, aug 2005.

[32] J. B. Anderson, *Source and Channel Coding: An Algorithmic Approach.* Kluwer Academic Publishers, 1991.

[33] E. Adelson, C. Anderson, J. Bergen, P. Burt, and J. Ogden, "Pyramid method in image processing," *RCA Engineer*, vol. 29, no. 6, pp. 33–41, Nov. 1984.

BIOGRAPHICAL STATEMENT

Basavaraj Hiremath was born in Karnataka, India, in 1981. He received his B.E. degree in Biomedical Engineering from Vishveshwariah Technological University (VTU), India, in 2003. From 2003 to 2005, he was with the department of Biomedical Engineering, KLE College of Engineering and Technology, as Lecturer. He taught Digital Image Processing and Microcontrollers during his tenure. His current research interest is in the area of Signal and Image processing. He is a student member of IEEE and permanent member of Indian Society of Technical Education (ISTE).