FEC FOR EFFICIENT VIDEO TRANSMISSION OVER CDMA

A THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF TECHNOLOGY

IN

ELECTRONICS SYSTEM AND COMMUNICATION

By

Ms. SUCHISMITA BEHERA

Department of Electrical Engineering

National institute of Technology

Rourkela-769008

2007

FEC FOR EFFICIENT VIDEO TRANSMISSION OVER CDMA

A THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF TECHNOLOGY

IN

ELECTRONICS SYSTEM AND COMMUNICATION

By

Ms. SUCHISMITA BEHERA

Under the Guidance of

Dr. SUPRAVA PATNAIK

Department of Electrical Engineering

National institute of Technology

Rourkela-769008

2007

National institute of Technology
Rourkela

CERTIFICATE

This is to certify that the thesis entitled "**FEC For efficient Video Transmission Over CDMA** " submitted by **Ms.Suchismita Behera**, in partial fulfillment of the requirements for the award of Master of Technology in the Department of Electrical Engineering, with specialization in '**Electronics System and Communication**' at National Institute of Technology, Rourkela (Deemed University) is an authentic work carried out by her under my supervision and guidance.

To the best of my knowledge, the matter embodied in the thesis has not been submitted to any other University/Institute for the award of any Degree or Diploma.

Date:

Dr. Suprava Patnaik
Asst. Professor
Department of Electrical Engineering
NATIONAL INSTITUTE OF TECHNOLOGY
Rourkela-769008

ACKNOWLEDGEMENTS

Suchismita Behera

M.Tech (Electronics System and Communication)

TABLE OF CONTENTS

CHAPTERS:

1. INTRODUCTION

2. BACKGROUND

3. ERROR CONTROL IN VIDEO COMMUNICATIONS AND BASICS OF CDMA

4. SENSITIVITY OF STREAM TO BIT ERRORS

5. FORWARD ERROR CORRECTION SCHEMES USING ERROR CORRECTING CODES

6. CONCLUSION

ABSTRACT

Video is one of the most important and also one of the most challenging types of traffic on communication networks. One of the challenges arises because communication networks can insert errors into video, and also compressed video is fragile in the presence of errors; that is a single error can propagate over a large portion of the video. This thesis describes how to protect the content of video during transmission by employing error correcting codes.

It deals with the concept of transmitting video packets with redundancy embedded in it for attempting error recovery and improving the quality at the receiver side. The effect of bit errors on the overall video quality has been studied. This thesis throws light on the importance of each bit in a video frame to the overall received quality. The effect of the important bits has been monitored by performing a "bit killing" operation, wherein the critical bits are flipped. We have adopted BCH codes to improve the quality of video. Two types of FEC schemes have been implemented. In the first FEC scheme, data is not prioritized and various segments in the video stream are given equal importance. BCH (63, 51, 2) is used for this case. In the second scheme, a stronger FEC is applied to the first frame (the first I frame) as the important data is found in it. Hence BCH (63, 24, 7) is used for the first frame and BCH (63, 57, 1) is used for the rest inter frames. The main aim of the algorithm is to strengthen the video stream against most possible error cases and also reduce the percentage of the overhead bits arising because of BCH coding.

The strength of the BCH is also varied to study the effect of signal-to-noise ratio on the BER. All discussions are furnished with simulation results. The primary target is the transmission of video through CDMA channels.

LIST OF FIGURES

LIST OF ACRONYMS

BCH	Bose, Chaudhuri and Hocquenghem
BER	Bit Error Rate
CDMA	Code Division Multiple Access
CIF	Common Intermediate Format
DCT	Discrete Cosine Transform
FEC	Forward Error Correction
IDCT	Inverse Discrete Cosine Transform
I Frame	Intra frame
ITU	International Telecommunication Union
MB	Macroblock
P frame	Inter frame
B frame	Bidirectional predictive frame
PSNR	Peak SNR
QCIF	Quarter Common Intermediate Format
SNR	Signal-to-Noise Ratio

CHAPTER 1

INTRODUCTION

Motivation

Objective

Thesis outline

1.1 INTRODUCTION

Guglielmo Marconi invented the wireless telegraph in 1896. In 1901, he sent telegraphic signals across the Atlantic Ocean from Cornwall to St. John's Newfoundland; a distance of 1800 miles. His invention allowed two parties to communicate by sending each other alphanumeric characters encoded in an analog signal. Over the last century, advances in wireless technologies have led to the radio, the television, the mobile telephone, and communication satellites. All types of information can now sent to almost every corner of the world. Recently, a great deal of attention has been focused on satellite communication, wireless networking, and cellular technology.

Historically, wireless data communications was principally the domain of large companies with specialized need. For example, large organizations need to stay in touch with their mobile sales force, or delivery services need to keep track of their vehicles and packages. However, this situation is steadily changing and wireless data communications is becoming as commonplace as its wired counterpart.

In recent years we have witnessed a persistent growth in wireless communication systems. Wireless access to the Internet may outstrip all other forms of access in the near future. It is likely that mobile users will expect similar levels of service quality as wire line users. One of the factors behind this growth was the dominance of digital wireless communication systems, which have superior bandwidths and can integrate voice and data communications. As a result, wireless communication system now offers higher quality and quantity of services.

The need for wireless data communications arises partially because of the need for mobile computing and partially because of the need for specialized applications, such as computerized dispatch services and mobile fleet management.

Mobile computing, which aims to migrate the computing world onto a mobile environment, is affected primarily by two components: portability and connectivity. Portability, i.e., the ability to untether computers from the conventional desktop environment, is getting increasingly feasible because with the continuous improvement in integration, miniaturization, and battery technology, the differences in performance and cost between desktop and portable computers are shrinking. Therefore, the processing power of desktop computing is becoming available to portable environments and this is highly desirable as far as productivity is concerned.

In the last decade, there has also been a rapid development in the digital video compression filed. Using new compression algorithms with high compression rates, digital video can provide superior quality with low bandwidths over its analog counterpart. With the addition of low and very-low bit rate coding algorithms, video transmission over wire-line channels is now a mature area.

Combination of advances in digital video compression and digital wireless communications resulted in a new service area called video over wireless. Potential applications such as HDTV and picture-phones attracted attentions, and now are under rapid development. Services related to video transmission over wireless channels are expected to grow exponentially in the near future.

Due to the proliferation of multimedia on WWW and the mergence of broadband wireless network, wireless video communication has received great interests from both industry and academy. However, the characteristics of wireless channel impose new problems to video transmission.

1.1 CURRENT STATE

Let us briefly summarize the current state of technologies on both wireless and video fields. On the wireless side, we have Digital Video Broadcasting (DVB) services (~ 20 Mbps) for broadcasting applications, Digital European Cordless Telephone (DECT) (~ 500 kbps) for short-range wireless communications, Global System for Mobile Communications (GSM) (~ 10 kbps) and in the near future 3G personal communication networks (~ 100 kbps).

On the video compression side we have HDTV format (1920x1080 pixels at 30 frames/sec) at ~20 Mbps, common intermediate format (CIF) (352x288 pixels ~ 10-20 frames/sec) at ~ 500 kbps and quarter common intermediate format (QCIF) (176x144 pixels ~ 10 frames/sec) at ~ 10 kbps. Although various data rates are possible, these correspond to DVB, DECT and GSM bandwidths, respectively

1.3 OVERVIEW OF DIGITAL VIDEO

First, we give a brief overview of digital video. Let us start with an analog video signal generated by an analog video camera. The analog video signal consists of a sequence of video

frames. The video frames are generated at a fixed frame rate (30 frames per second in the National Standards Committee (NTSC) format). For each video frame, the video camera scans the frame line by line(with 455 lines in NTSC). To obtain a digital video signal the analog video signal is passed to a digitizer. The digitizer samples and quantizes the analog video signal. Each sample corresponds to a picture element (pel). The most common digital frame formats are Common Intermediate Format (CIF) with 352x288 pixels, Source Intermediate Format (SIF) with 352x240 pixels, and Quarter Common Intermediate Format (QCIF) with 176x144 pixels. In all three frame formats, each video frame is divided into three components. These are the luminance components (Y), and the two chrominance components: hue (U) and intensity (saturation) (V). Since the human eye is less sensitive to the color information, than to the luminance information ,the chrominance components are sampled are sampled at a lower resolution. Typically each chrominance component is sampled at half the resolution of the luminance component in both the horizontal and vertical direction. (This is referred to as 4:1:1 chroma sub sampling.) In the QCIF frame format, for instance, there are 176x144 luminance samples , 88x72 hue samples, and 88x72 intensity samples in each video frame, when 4:1:1 chroma sub sampling is used. Finally, each sample is quantized; typically, 8 bits are used per sample.

As an aside we note that the YUV video format was introduced to make color TV signals backward compatible with black- and –white TV sets, which can only display the luminance (brightness) components. Computers monitors, on the other hand, use typically the RGB video format, which contains red, green, and blue components for each pel.

Before we discuss the specific features of MPEG-4 and H.263 we briefly outline some of their common aspects. Both encoding standards employ the Discrete Cosine Transform (DCT) to reduce the spatial redundancy in the individual video frames. Each video frame is divided into Macro Blocks (MBs). A macro block consists of 16x16 samples of the luminance component and the corresponding 8x8 samples of the two chrominance components. The 16x16 samples of the luminance component are divided into four blocks of 8x8 samples each. The DCT is applied to each of the six blocks (i.e., four luminance blocks and two chrominance blocks) in the macro block. For each block the resulting DCT coefficients are quantized using an 8x8 quantization matrix, which contains the quantization step size for each DCT coefficient. The quantization matrix is obtained by multiplying a base matrix by a quantization parameter. This quantization parameter is typically used to tune the video encoding. A larger quantization parameter results in

coarser quantization, which in turn results in a lower quality as well as a smaller size (in bit) of the encoded video frame. The quantized DCT coefficients are finally variable-length-coded, for a more compact representation.

Both, MPEG-4 and H.263 employ predictive encoding to reduce the temporal redundancy, that is, the temporal correlation between successive video frames. A given macroblock is either intracoded (i.e., without reference to another frame) or interceded (i.e., with reference to a preceding or succeeding frame). To intercede a given macroblock, a motion search is conducted to find the best matching 16x16 sample area in the preceding (or succeeding) frame. The difference between the macroblock and the best matching area is DCT coded, quantized, and variable-length-coded, and then transmitted along with a motion vector to the matching area.

1.2 MOTIVATION

With the increasing demand for multimedia services on mobile terminals, and the recent advances in mobile computing, video services are expected to be widely deployed. Over the last couple of years, delivering video services to users over wireless networks has become possible with advances in low bitrate video coding. Videoconferencing has been conjured to be a key mode of communication between people in the future. Multimedia allows people to share media rich information in a format that is comprised of text, audio, still images, and video. Collaborative systems are one another consumer of video conferencing / telephony systems. These systems require interaction among various users by way of multimedia communication, leading to a better information dispatch. Clearly, with the increase in bandwidth, the world is moving towards high bitrate communication. Video-on-demand applications are also gaining increased popularity. Thus, the world seeks visually enhanced video contingent with all resource limitations.

The advances in low bit-rate video coding technologies have led to the possibility of delivering video services to users through band-limited wireless networks. CDMA is the trend for the current and next generation wireless networks due to the advantages it offers over TDMA and FDMA. Code Division Multiple Access (CDMA), which is currently being used for cellular digital services, acts as the carrier protocol for most of those applications. The target device may be a mobile telephone (or) a laptop on a wireless network. ITU-T worked on a standard for such applications and developed H.263. H.263 is a compression standard developed to transmit video

using the phone network at data rates less than 64kbps. Adoption of the H.263 video compression standard for the wireless video applications seems ideal, since the bit-rate criterion is met.

Any path towards a better system has its own set of obstacles that hinder the overall performance. Owing to the applied compression, H.263 is very vulnerable to bit errors and CDMA does not guarantee such a quality. Fading on radio channels causes significant transmission errors. Hence extraneous effort is need to control error in the video streams. This is the true reason for incubation of this thesis work. There are various proposals for error control and ultimately improving video quality. This thesis deals predominantly with Forward Error Correction (FEC) based control.

1.3 RESEARCH OBJECTIVE AND SCOPE

For effective video communications, reduction of raw video data rates is only one of the necessary steps. Another equally important task is handling errors and losses in a communication network. In contrast with data communications, which are not usually subject to strict delay constraints and can therefore be handled using network protocols that use retransmission to ensure error free delivery, real- time video is delay sensitive and cannot easily make use of retransmission. The extensive use of predictive and variable-length coding in video coding renders compressed video especially vulnerable to transmission errors, and successful video communication in the presence of errors requires careful designs of the encoder, decoder and other system layers.

Forward error correction, which is regarded as the proactive mode of error control, is a popular strategy adopted in most system attributing to its low recovery time and latency. In FEC, irrespective of whether the channel has error at that particular instant, error protection is provided. This thesis research work is focused towards enhancing the error resilience of H.263 stream by way of FEC coding. The proposed solutions in this thesis can be applied to other error-prone systems too. But, the gravity of the wireless applications deemed it important to be considered as the primary target. Performance evaluation and comparison of the various schemes calls for appropriate performance metrics. In the case of video telephony, quality of the pictures is the main aspect under serious consideration. Quality refers to closeness of the reconstructed video to the original video transmitted. This integrates the factors of frame rate, continuity of

motion, chrominance levels and luminance levels. The main objective is to maximize the quality of the received video stream by employing of FEC coding. Different modifications to the original H.263 protocol are considered keeping in mind the reconstructed video quality.

Identifying all deterrents is the first step in the project. The anomalous cases, where the conventional forward error correction will not aid the quality, also needs to be identified. FEC coding brings in more overhead in the video stream thereby reducing the system throughput. This situation needs to be remedied by optimizing the required number of the parity bits. There is a high processing lag in performing the error correction operation. This needs to be resolved by determining when exactly the FEC is required. The challenge is to make best use of the overhead bits so as to achieve t the highest reconstructed video quality.

1.4 THESIS OUTLINE

The chapters in this thesis are organized as follows. In Chapter 2, the required background for understanding the research work is presented. All details on the adopted video compression standard H.263 ,motion estimation and motion compensation , block matching algorithms have been discussed.

Chapter 3 deals with the error control techniques in video communication. Many error control techniques such as layering, forward error correction, and retransmission have been discussed. BCH codes have also been discussed in this chapter which is used for forward error correction in this thesis. The concepts of quality estimation based on reconstructed video and transmission rate control are also discussed in quite some detail. This chapter also deals with the wireless channel characteristics. The matched filter detector is also discussed which is used in this thesis to estimate the received bits at the receiver side.

Chapter 4 deals with estimating sensitivity of a H.263 video stream to bit errors. In particular, it lists the predominant reasons for the decline in video quality owing to bit errors. Based on these a set of hypotheses is generated. This section outlines them and provides the experimental backing for the theories suggested. This section goes further and delves into sensitivity of each type of bit in a video stream. It finally summarizes the ideas to form the basis for the novel strategies.

Chapter 5 presents the simulation methodology and its results. Details on automating the simulation are provided. The results generated have been tabulated and further analyzed in this

section. The discussions are from higher perspective and compare the various schemes, basing peak signal to noise ratio as the quality determinant. This chapter also suggests limitation brought in by each scheme.

Finally, chapter 6 provides the conclusion of this thesis work. The insights generated from the research work have been highlighted in this section. Momentum for further improvement of video quality has been provided in the form of future work.

CHAPTER 2

BACKGROUND

Motion compensated hybrid coding

The H.263 coding standard

Motion estimation and motion compensation

This chapter discusses some of the important concepts that needed reading and that are also applied in this thesis work. As much detail is provided as will be required to understand the core of the research. Importance is given to the basics of H.263 video compression and the various error control techniques. The theory behind some of the optimization is explained in this section.

2.1 MOTION-COMPENSATED HYBRID CODING

2.1.1 General Principles

Most state-of-the-art low bit-rate video codecs are motion-compensated hybrid codecs, as illustrated in Fig.2.1. Two basic modes of operation can be selected, depending on the position of the switch . These two modes allow the video signal in the current frame to be encoded either directly (INTRA coding), or with reference to previously encoded and reconstructed frames (INTER coding). The INTER mode combines differential pulse code modulation (DPCM) along an estimated motion trajectory with intraframe encoding of the residual prediction error. Motion-compensated prediction is carried out by estimating the motion between successive frames, shifting the contents of a previously encoded, reconstructed frame accordingly, and transmitting the motion vector in addition to the prediction error residual as side information. The residual prediction error is usually small and requires fewer bits than directly encoding the original video signal. In all current compression standards, the discrete cosine transform (DCT) is employed for this purpose with a blocksize of 8x 8 pixels. The transform coefficients are quantized and typically encoded as a series of zero-runs and quantizer levels. Transform coefficients and motion vectors are entropy coded along with other side information resulting in variable-length code words, which are multiplexed to the video bit stream. For most of a video signal, the INTER mode is the preferred mode because of its superior coding efficiency. However, some changes in successive frames, for example, due to uncovered background, new objects appearing in the scene, or after a scene cut, cannot be predicted well, and subtracting the prediction might lead to a prediction error that requires more bits than the original video signal. Therefore, the second basic encoding mode besides INTER coding is the INTRA mode, in which no reference to previous frames is made and the picture is directly intraframe coded. Again, a variety of schemes can be used, but typically a blockwise 8x8 DCT coder is employed.
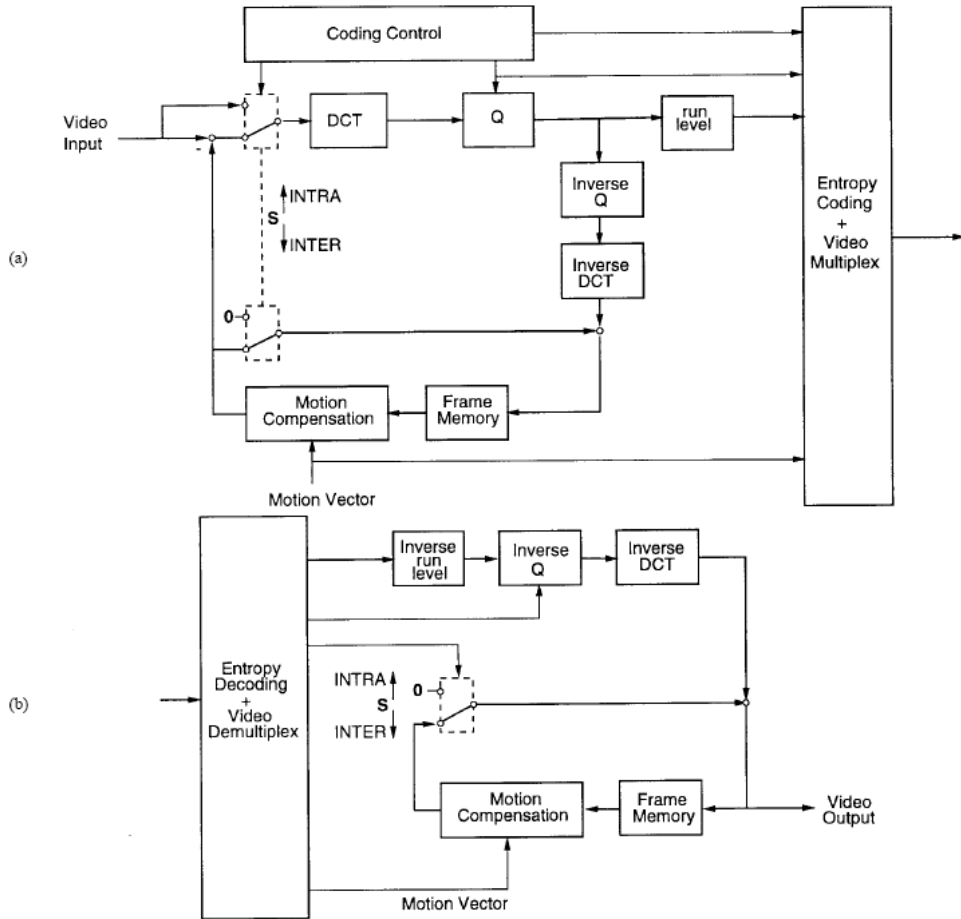
Figure.2.1

   (a) Motion compensated encoder

   (b) Decoder

2.2 THE H.263 CODING STANDARD

H.263 was designed for very low bitrate coding applications.H.263 is based on H.261 but it is significantly optimized for coding at low bit rates. H.263 is the video compression standard for low bitrate communications. It prescribes procedures to derive a 64 Kbps video stream from the raw motion samples (like QCIF). A hybrid of inter-picture prediction to utilize temporal redundancy and transform coding of the remaining signal to reduce spatial redundancy is adopted. Most of these procedures are found in MPEG encoding too. Both fixed length and variable length coding is used for the symbols to be transmitted [1] [11] [12].

9

The general principles discussed above are the basis for all video compression standards in use today, in particular the ISO standards MPEG-1, MPEG-2 , and MPEG- 4 , and the ITU-T Recommendations H.261, H.262 (identical with MPEG-2), and H.263 . We will use H.263 as an example throughout this paper. H.263 is the video compression standard for low bitrate communications. The picture resolution is often quarter common intermediate format (QCIF, 176x144 pixels), which is the most common input format at such low bit rates.

At QCIF resolution, each picture is divided into 11x9 macroblocks (MB's), which comprise 16x16 luminance samples, and two corresponding 8x8 blocks of chrominance samples. The luminance component of each MB is further subdivided into four 8x8 blocks, such that 8x8 DCT's can be applied to each block. Each MB is encoded either in INTRA or INTER mode. Motion compensation is carried out with half-pixel accuracy with one motion vector sent for each MB.

The encoding process is divided into the following steps: motion estimation, prediction calculation, DCT type estimation, subtraction of prediction from picture, DCT calculation, DCT coefficients quantization and generation of VLC data, picture interpolation, inverse DCT coefficients quantization and finally IDCT calculation and addition of predicted image[6].Due to the use of the motion compensated temporal interpolation technique, three different types of encoded images are produced, by using the following three different coding modes:

Intra mode: images encoded individually without using temporal prediction (without reference no any other picture);

Predicted mode **(P-pictures):** inter frame coded pictures using unidirectional MC prediction;

Bi-directional mode: which generates inter frame coded pictures using bi-directional motion compensated prediction. B-frames may be encoded using either forward prediction where reference is made to an image in the past, backward prediction where reference is made to a future image, and finally to an image in the past and one in the future.
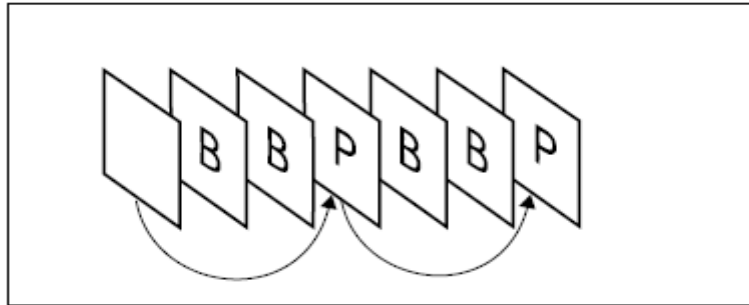
Forward prediction
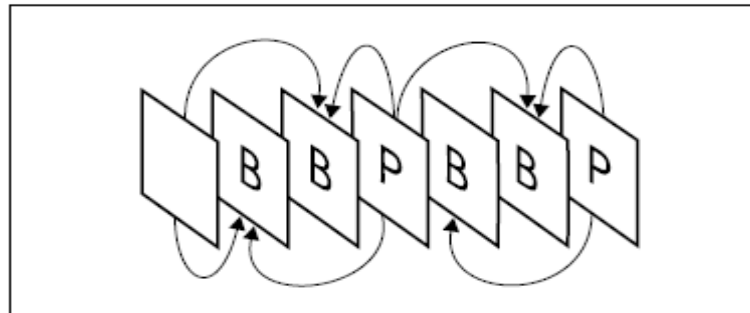
Backward Prediction

Figure 2.2

The I picture frame at the beginning of a GOP serves as a basic entry point to facilitate random seek or channel switching and also provides coding robustness to transmission error, but it is coded with only moderate compression to reduce the spatial redundancies. P picture frames are coded more efficiently using motion compensated prediction from a past I or P picture frame and generally used as a reference for further prediction. B picture frames provide the highest degree of compression, but require both past and future reference pictures for motion compensation. It should be mentioned that B pictures are never used as references for prediction.

The structure of the GOP needs to be specified. This specification indicates the number of frames in a GOP, the number of P frames in a GOP, the number of B frames in a GOP, and the interleaving of the I, P, and B frames in the GOP.

The GOP structure is specified using two integer parameters, N and M. N specifies the number of frames in the GOP. M specifies the length of the substructure in the GOP (i.e., after each I frame or P frame, there are M - 1 consecutive B frames before the next I or P frame). N

must be an integer multiple of M. The default GOP pattern is IBBPBBPBB (i.e., N =9, M = 3). To maximize coding efficiency, both the motion compensation information and the transformed prediction error are represented using variable length (Huffman) codes (VLC).

The figure 2.3 below depicts the syntax of H.263 frame. In H.263 only I and P pictures are considered and not there are no B pictures. These predictive-coded frames (P frames) are coded in terms of differences of the current image relative to the previous I or P frame. For coding efficiency, each picture is divided into macroblocks, where each macro block consists of four luminance blocks and two spatially aligned color difference blocks. Each block consists of 8 pixels **x** 8 lines of luminance or chrominance. One or more macroblock rows are combined into a group of blocks (GOB) to enable quick resynchronization after transmission errors. Each picture consists of 9 Group of Blocks (GOB) , which in turn contains 11 macroblocks(MB). Each macroblock contains 4 luminance components and 2 chrominance components.
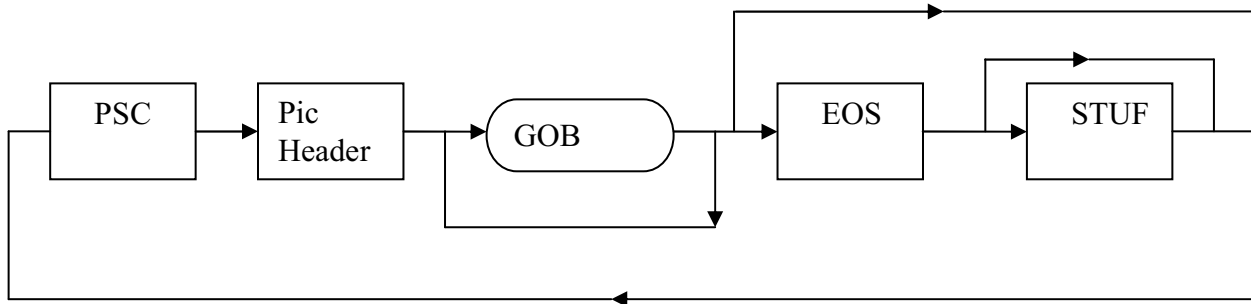
PICTURE LAYER

GROUP OF BLOCKS (GOB) LAYER
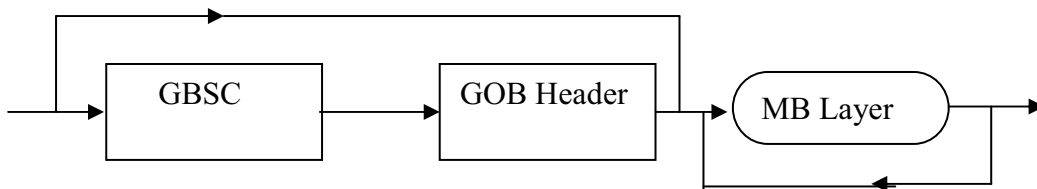
MACROBLOCK (MB) LAYER

BLOCK LAYER
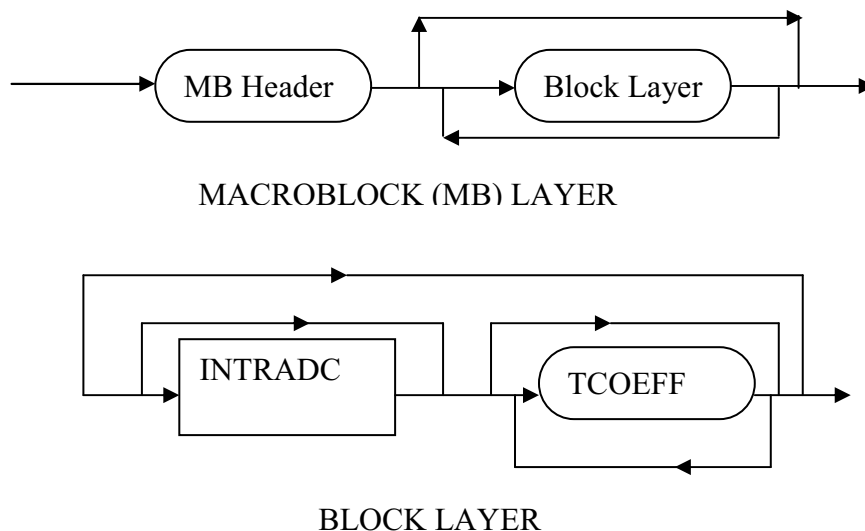
Fig.2.3: Simplified syntax diagram for the H.263 video bitstream. Each frame (also called picture) comprises of 9 Group of Blocks (GOB), which in turn contains 11 Macroblocks (MB). Each macroblock contains 4 luminance components and 2 chrominance components

2.2.1 Intraframe Coding

The spatial redundancy is exploited using the intraframe coding. The word spatial refers to space in a single picture and the goal of spatial compression is to minimize the duplication of data in each picture. Bit rate reduction in spatial compression is achieved by first transforming the video data from the space and time domain into the frequency domain using the discrete cosine transform (DCT) method and then applying quantization and variable length coding techniques to reduce the bit rate. To accomplish data reduction, we must first transform the video data into the frequency domain. This is where DCT (a trigonometrically formula derived from Fourier analysis theory) is used to transform the data in each block of 8x8 pixels into blocks of 8x8 frequency coefficients. In the frequency domain, most of the high energy (and therefore most noticeable) picture elements are represented by low frequencies at the top left corner of the block, and the less important details are revealed as higher frequencies towards the bottom right At this stage we have not yet discarded any bits.
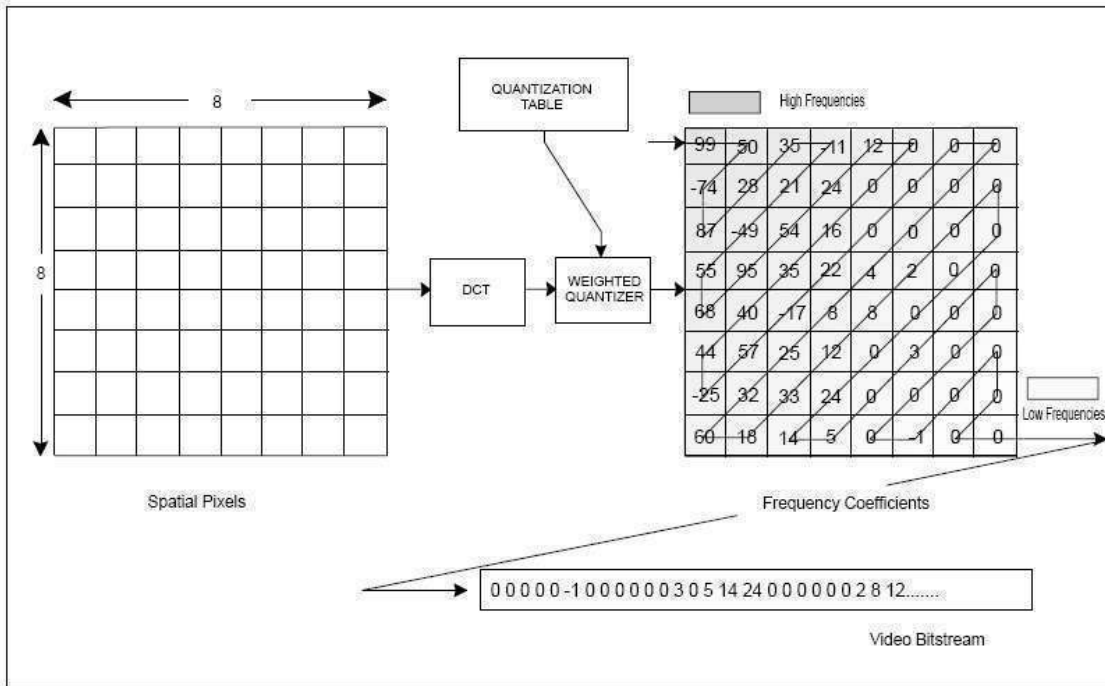
Fig.2.4 Zig-zag scanning

After DCT encoding, the data is subjected to a quantization process, weighted to reduce data in the high frequency areas, where the eye is less sensitive. We use more bits per pixel to quantize the important low-frequency coefficients and fewer bits per pixel for the high-frequency coefficients.

The DC components are normally quantized at 10 bits, because if we employ coarser quantization of very low frequencies, the blocks themselves can start to become visible in the pictures. We have now achieved the first step in spatial bit rate reduction.

To create the compressed video bit stream, the 64 frequency coefficients are scanned in a zig-zag fashion from top left to bottom right and, the high-frequency areas are represented by strings of zeros. Further data reduction can now be achieved by transmitting only the number of zeros instead of the usual values of the coefficients.

The last stage in the spatial compression process employs variable length coding (VLC). VLC assigns shorter code words for frequently occuring events and longer code words for less frequent events; it is also reversible.

2.2.2 Interframe Coding

Interframe processing is the key to exploit and reduce the temporal redundancy in digital video compression. Temporal redundancy results from a high degree of correlation between adjacent pictures. B and P pictures are based upon the Interframe coding techniques. The encoding process for P and B pictures is explained as follows.

Data representing macroblocks of pixel values for a picture to be encoded are fed to both the subtractor and the motion estimator. The motion estimator compares each of these new macroblocks with macroblocks in a previously stored reference picture or pictures. It finds the macroblock in the reference picture that most closely matches the new macroblock. The motion estimator then calculates a motion vector (mv) which represents the horizontal and vertical displacement from the macroblock being encoded to the matching macroblock-sized area in the reference picture. Note that the motion vectors have 1/2 pixel resolution achieved by linear interpolation between adjacent pixels. The motion estimator also reads this matching macroblock (known as a predicted macroblock) out of the reference picture memory and sends it to the subtractor which subtracts it, on a pixel by pixel basis, from the new macroblock entering the encoder. This forms a error prediction or residual signal that represents the difference between the predicted macroblock and the actual macroblock being encoded. This residual is often very small. The residual is transformed from the spatial domain by a 2 dimensional DCT. (The two dimensional DCT consists of separable vertical and horizontal one-dimensional DCTs.) The DCT coefficients of the residual are then quantized in a process that reduces the number of bits needed to represent each coefficient. Usually many coefficients are effectively quantized to 0.

The quantized DCT coefficients are Huffman run/level coded which further reduces the average number of bits per coefficient. This is combined with motion vector data and other side information (including an indication of I, P or B picture) and sent to the decoder. The following sections describe the problem of Motion Estimation and different proposed algorithms to solve this problem.

2.3 INTRODUCTION TO MOTION ESTIMATION AND COMPENSATION

The movement of a macroblock from the reference frame to the current frame is determined using a technique known as motion estimation . A search of a macroblock in the current frame is conducted over a portion (or all) of the reference frame. The best matching macroblock (under certain criteria) is selected, and a motion vector is obtained, as shown in Figure 2. A motion vector consists of horizontal and vertical components. This motion vector can be expressed in integer or half-pixel accuracy. Half-pixel accuracy corresponds to bilinear interpolation.

A predictive frame is constructed from the motion vectors obtained for all macroblocks in the frame. Macroblocks from the reference frame are replicated at the new locations indicated by the motion vectors. This technique is known as motion compensation. A predictive error frame (PEF) is calculated by taking the difference between the current and predicted frames, and intraframe encoded. Since the energy of the PEF is likely to be low, the amount of bits necessary for its encoding is small. Motion compensation is used both at the encoder and the decoder to produce a motion compensated version of the current frame. This motion compensated frame is reconstructed by using the predicted frame and the PEF [3] [4].

2.3.1 Block Matching Algorithm
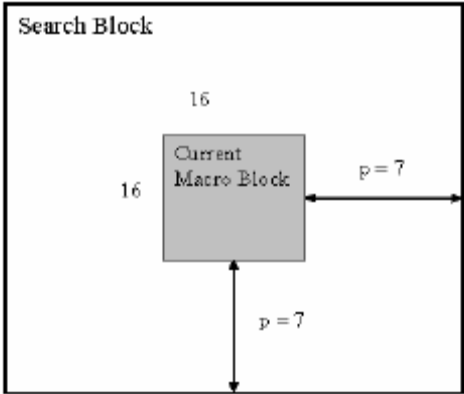
Fig.2.5: Block matching a macro block of side 16 pixels and a search parameter
  p of size 7 pixels.

The underlying supposition behind motion estimation is that the patterns corresponding to objects and background in a frame of video sequence move within the frame to form corresponding objects on the subsequent frame. The idea behind block matching is to divide the current frame into a matrix of 'macro blocks' that are then compared with corresponding block and its adjacent neighbors in the previous frame to create a vector that stipulates the movement of a macro block from one location to another in the previous frame. This movement calculated for all the macro blocks comprising a frame, constitutes the motion estimated in the current frame. The search area for a good macro block match is constrained up to p pixels on all fours sides of the corresponding macro block in previous frame. This 'p' is called as the search parameter. Larger motions require a larger p , and the larger the search parameter the more computationally expensive the process of motion estimation becomes. Usually the macro block is taken as a square of side 16 pixels, and the search parameter p is 7 pixels. The idea is represented in Fig 2.5. The matching of one macro block with another is based on the output of a cost function. The macro block that results in the least cost is the one that matches the closest to current block. There are various cost functions, of which the most popular and less computationally expensive is Mean Absolute Difference (MAD) given by equation (i).Another cost function is Mean Squared Error (MSE) given by equation (ii).

$$MAD = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \left| C_{ij} - R_{ij} \right| \tag{2.1}$$

$$MSE = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \left(C_{ij} - R_{ij} \right)^2 \tag{2.2}$$

where N is the side of the macro bock, C_{ij} and R_{ij} are the pixels being compared in current macro block and reference macroblock , respectively [3] [11].

2.3.2 Exhaustive Search (ES)

Given an image block in the anchor frame β_m , the motion estimation problem at hand is to determine a matching block β'_m in the target frame such that the error between these two blocks is minimized. The displacement vector $\mathbf{d}_m$ between the spatial positions of these two blocks is the MV of this block. Because the estimated MV for a block affects the prediction error in that block

only, one can estimate the MV for each block individually, by minimizing the prediction error accumulated over each block, which is:

$$E_m(d_m) = \sum_{x \in \beta_m} |\Psi_2(x + d_m) - \Psi_1(x)|^p \qquad (2.3)$$

One way to determine the $\mathbf{d}_m$ that minimizes this error is by using exhaustive search, and this method is called the *exhaustive block-matching algorithm* (EBMA). As illustrated in Figure2.6, the EBMA determines the optimal $\mathbf{d}_m$ for a given block β_m in the anchor frame by comparing it with all candidate blocks β'_m in the target frame within a predefined search region and finding the one with the minimum error. The displacement between the two blocks is the estimated MV.
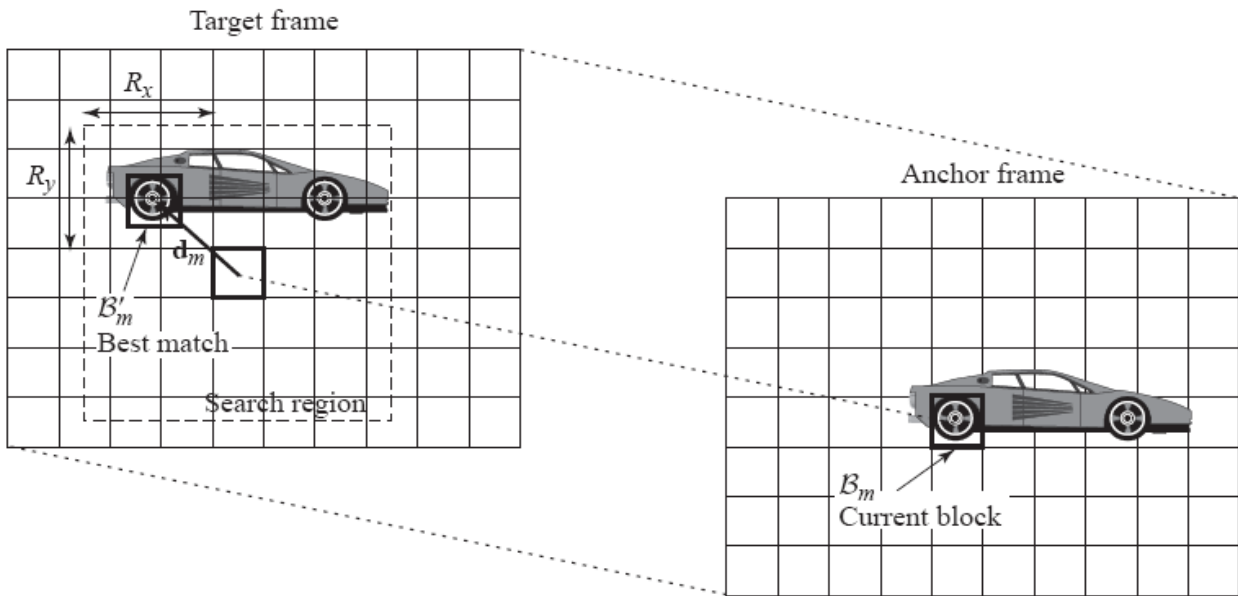
Figure 2.6 The search procedure of the exhaustive block- matching algorithm.

To reduce the computational load , the MAD error(p=1) is often used. The search region is usually symmetric with respect to the current block, up to R_x pixels to the left and right, and up to R_y pixels above and below, as illustrated in the figure. The estimation accuracy is determined by the search step size, which is the distance between the two nearby candidate blocks in the horizontal and vertical direction. In the simplest case, the step size is one pixel, which is known as *integer- pel accuracy search*.

Let the block size be N × N pixels, and the search size be ±R pixels in both horizontal and vertical directions (see Figure). With a step size of one pixel, the total number of candidate

matching blocks for each block in the anchor frame $(2R+1)^2$. Let an operation be defined as consisting of one subtraction, one absolute value computation, and one addition. The number of operations for calculating the MAD for each candidate estimate is N^2. The number of operations for estimating the MV for one block is then $(2R+1)^2N^2$. For an image of size M × N, there are $(M/N)^2$ blocks . The total number of operations for a complete frame is then $M^2(2R+1)^2$. It is interesting to note that the overall computational load is independent of the block size N.

This algorithm, also known as Full Search, is the most computationally expensive block matching algorithm of all. This algorithm calculates the cost function at each possible location in the search window. As a result of which it finds the best possible match and gives the highest PSNR amongst any block matching algorithm. Fast block matching algorithms try to achieve the same PSNR doing as little computation as possible. The obvious disadvantage of ES is that the larger the search window gets the more computations it requires.

2.3.3 Fractional accuracy EBMA

The stepsize for searching the corresponding block in the BMA need not be an integer. For more accurate motion representation, fractional-pel accuracy is needed. A problem with using fractional step sizes is that there may not be corresponding sample points in the target frame for certain sample points in the anchor frame-these samples must be interpolated from the available sample points. Bilinear interpolation is commonly used for this purpose. In general, to realize a stepsize of 1/K pixel, the target frame must be interpolated by a factor of K. An example of K=2 is shown in figure 2.7, which is known as *half-pel accuracy search*. For half pel EBMA, the following formulas are used for interpolating the points:

O [2x, 2y] =I[x, y]   (2.4)
O [2x+1, 2y] = (I[x, y] +I[x+1, y])/2   (2.5)
O [2x, 2y+1] = (I[x, y] +I[x, y+1])/2   (2.6)
O [2x+1, 2y+1] = (I[x, y] +I[x+1, y] +I[x, y+1] +I[x+1, y+1])/4   (2.7)
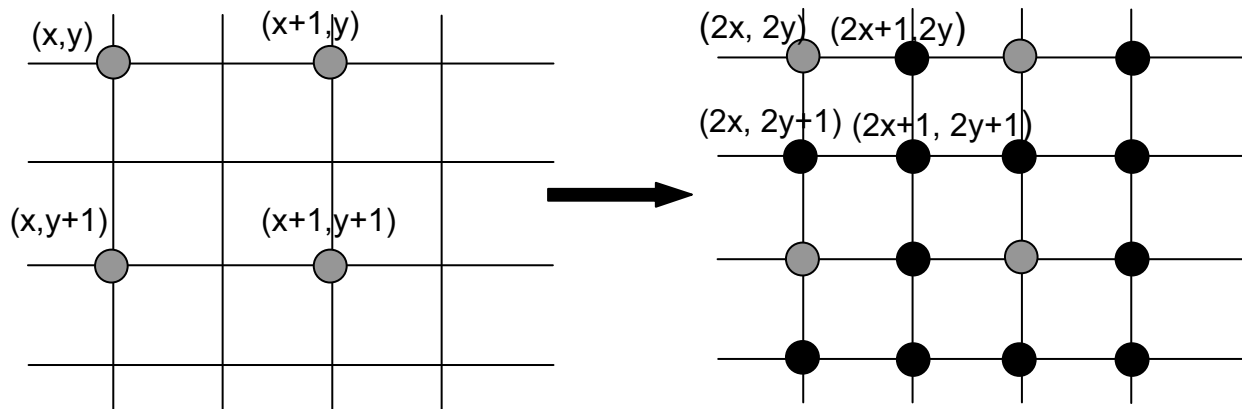
Figure 2.7: Half-pel accuracy block matching. Grey circles are samples existing in the original target frame, black circles are interpolated samples using the above formulas.

Obviously, with a fractional-pel stepsize, the complexity of the EBMA is further increased. For example, with half pel search, the number of search points is quadrupled over that using integer-pel accuracy. The overall complexity is more than quadrupled, considering the extra computation required for interpolating the target frame.

2.3.4 Fast Algorithms

As we have shown , the EBMA requires a very large amount of computation. To speed up the search, various fast algorithms for block matching has been deployed. The key to reducing the computation is reducing the number of search candidates. As previously described, for a search range of $\pm R$ and a stepsize of 1 pixel, the total number of candidates is $(2R+1)^2$ with EBMA. Various fast algorithms differ in the ways that they skip those candidates that are unlikely to have small errors[11].

Three-Step Search Method   As illustrated in Figure 2.8 , the search starts with a stepsize equal to or slightly larger than half of the maximum search range. In each step, nine search points are compared. They consist of the central point of the search square, and eight search points located on the search area boundaries. The stepsize is reduced by half after each step, and the search ends with a stepsize of 1 pel. At each new step, the search center is moved to the best matching point resulting from the previous step. Let R_0 represent the initial search stepsize; then there are at

most L=[log$_2$ R$_0$+1] search steps, where [x] represents the lower integer of x. If R$_0$ =R/2, then L=[log$_2$R]. At each search step, eight search points are searched, except in the very beginning, when nine points must be examined. Therefore, a total number of search points is 8L+1.
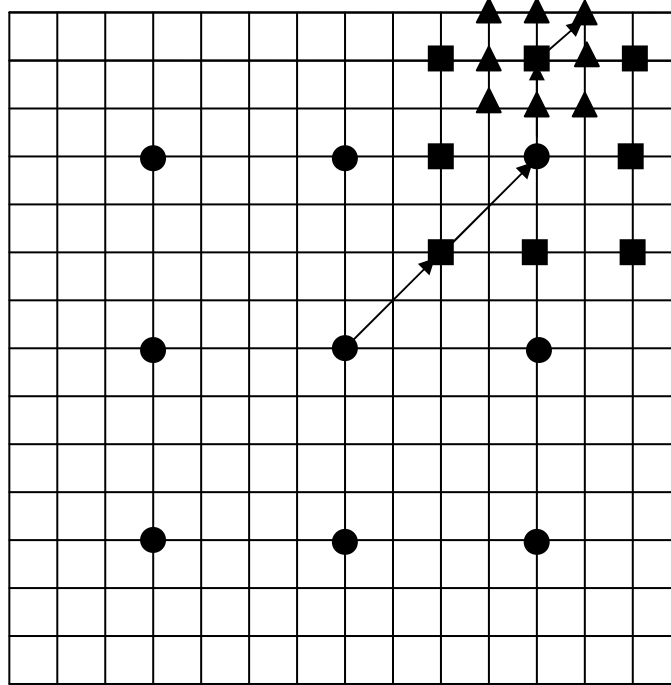
Figure 2.8: The three step search method. This algorithm is based on a coarse-to-fine approach with logarithmic decreasing in step size as shown. The initial step size is half of the maximum motion displacement d.

For each step, nine checking points are matched and the minimum BDM point of that step is chosen as the starting center of the next step. For d = 7, the number of checking points required is(9 + 8 + 8)=25. For larger search window (i.e. larger d), 3SS can be easily extended to n-steps using the same searching strategy with the number of checking points required equals to [1 + 8 log$_2$(d + 1)].

2.3.5 COMPARISON OF ALGORITHMS

TABLE 2.1: Comparison of algorithms for a search range of R=7.

SEARCH ALGORITHM	TOTAL NUMBER OF SEARCH POINTS	AVERAGE NO. OF SEARCH POINTS
EBMA	18271	184.56
SUB-PIXEL EBMA	67803	684.87
THREE STEP	2127	21.48

Table 2.1 compares the total number of search points required for the different type of algorithms. As compared to EBMA, the search points for sub-pixel EBMA havebeen quadrupled been quadrupled and for three step search method, it has been reduced by six times.

CHAPTER 3

ERROR CONTROL IN

VIDEO COMMUNICATIONS

AND

BASICS OF CDMA

Error control techniques

CDMA Channel BER

Matched filter concept

The video is first compressed by a video encoder to reduce the data rate, and the compressed bit stream is then segmented into fixed or variable length packets and multiplexed with other data types, such as audio. The packets might be sent directly over the network, if the network guarantees bit-error-free transmission. Otherwise, they usually undergo a channel encoding stage, typically using forward error correction (FEC) and interleaving, to protect them from transmission errors. At the received end, the receiver packets are FEC coded and unpacked, and the resulting bit stream is then input to the video decoder to reconstruct the original video.

Transmission errors can be roughly classified into two categories: random bit errors and erasure errors. Random bit errors are caused by imperfections of physical channels, which result in bit inversion, insertion and deletion. When fixed-length coding is used, a random bit error will affect only one codeword, and the damage caused is generally acceptable. But if VLC (e.g., Huffman coding) is used, random bit errors can desynchronize the coded information, such that many following bits are undecodable until the next synchronization codeword appears. Erasure errors, on the other hand, can be caused by packet loss in packet networks such as the Internet, burst errors in storage media. Random bit errors in VLC coded streams can also cause effective erasure errors, since a single bit error can lead to many following bits becoming undecodable, hence useless.

Error control in video communications is very challenging for several reasons. First, compressed video streams are extremely vulnerable to transmission errors because of the use of temporal predictive coding and VLC by the source coder. Due to the use of temporal prediction, a single erroneously recovered sample can lead to errors in the following samples in the same and following frames as illustrated in Figure 3.1.

Second , with VLC, the effect of a bit error is equivalent to that of an erasure error, causing damage over a large portion of a video frame.
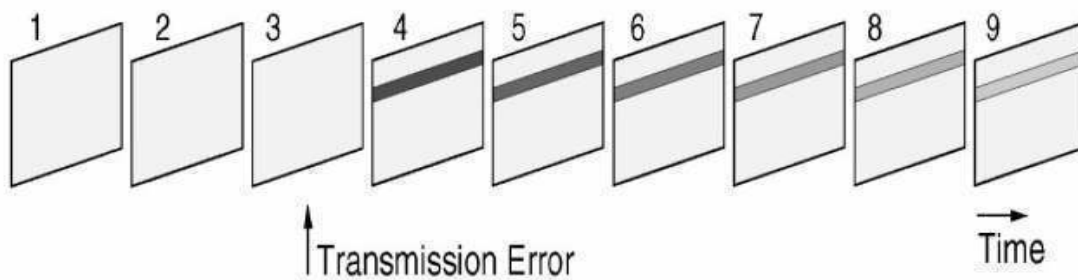

Figure 3.1: Illustration of spatiotemporal error propagation

To make the compressed bit stream resilient to transmission errors, one must add redundancy to the stream, so that it is possible to detect and correct errors, Typically, this is done at the channel by using FEC, which operates over coded bit streams generated by a source coder.

3.1 ERROR CONTROL TECHNIQUES

Many a techniques have been proposed for reducing error rates in the decoded video stream. They can be categorized as active (apply protection irrespective of whether error occurs or not) or reactive (used to recover from a prior error) [7] [11].

3.1.1 Layering

This active error control scheme tries to send data in the form of two or more layers, each with different priorities. The layering can be temporally achieved wherein the data of the I and P-frames are sent in together as high priority data and the B-frames which temporally depend on the others are sent as low priority data. But, the bottom line is that data of higher importance (the headers, the I-frame or the P-frame) is sent in a stream separate from the rest. This technique is also called data partitioning. The high priority stream should be kept small to optimize the bandwidth required. This also helps protect the stream easily because the size of redundancy introduced, when FEC is also brought in, will be small. The best approach is to rearrange the coded video information such that the important information is better protected and more likely to be received correctly. The data can also be partitioned by splitting the DCT coefficient in the frame into high priority ones and low priority ones.

3.1.2 Forward Error Correction (FEC)

FEC introduces some redundancy in the data bits so as to allow the receiver to predict what the source might have sent in despite corruption of the frame. FEC adds some redundancy so that errors can be detected and corrected. This comes at the cost of the overhead that eats away a portion of the network bandwidth. Care should be taken so as not to cause much loss in the compression efficiency. FEC does have an overhead but varying the strength of the error compression used can control it. Thus there is a tradeoff between overhead and video quality. BCH is a popular FEC algorithm. This paper deals with using varying strengths of BCH for error

correction. More details on BCH can be found in the later part of the chapter. There are different segments of the stream over which this control is exercised.

3.1.3 Retransmission

Retransmission based error recovery can provide good error resilience without incurring much bandwidth overhead because packets are retransmitted only when some indications exist that they are lost. However retransmission always involves additional transmission delay and thus has been widely known ineffective for interactive real-time video applications such as video conferencing. The playout time need to be adjusted in case we need to allow retransmission. This delay might be intolerable for interactive video applications. Thus there is a tradeoff between delay and quality. To allow enough time for retransmitted packets to arrive before their frames are referenced for the reconstruction of their dependent, the retransmission scheme must extend the temporal dependency distance of the frames. A scheme similar to ARQ (Automatic Repeat Request) is applied in certain implementations. Techniques like selective repeat can also be used.

3.1.4 Hybrid ARQ and other advanced options

ARQ has been rejected owing to its high latency. Variations of ARQ have been proposed to reduce the delay incurred. The prominent one is called the Hybrid ARQ. There are two classes of hybrid ARQ scheme.

• Type I: These schemes include parity bits for both error detection and error correction in every transmitted packet. If the received packet is uncorrectable, then a retransmission is requested. The transmitter resends the same packet from its inbuilt buffer. This strategy reduces the number of retransmission requests and thereby the delay incurred.

• Type II: This scheme sends the parity bits to the receiver only when needed. In the event of bit error, the packet can be recovered from the combination of information bit and parity bit. If uncorrectable, the receiver requests for the parity bit to be resent. The redundancy should be specially made out to recover a complete packet from just the parity bits. This scheme saves time and also bandwidth during retransmission.

Macroblock synchronization using additional information is used to reduce loss of information in the event of a Huffman code bit error. These bit errors corrupt the self synchronizing Huffman codes and lead to loss of the remaining portion of the frame. But,

indicating number of MB bits improves the situation. The decoder will be able to substitute the macroblock from the previous frame in the current frame thereby compensating for the loss to an extent.

Other schemes that rely on the dependence factor of frames have been suggested. When a frame is send from the source and it gets corrupted / lost *en route,* then the decoder can still function properly if the frame was not used at the later stage. There are encoding strategies that perform something like in Figure.2.2 to reduce the effect of frame drops. But, any strategy boils down to the idea of increasing the importance of some video information. The suggestions may or may not be standardized globally.
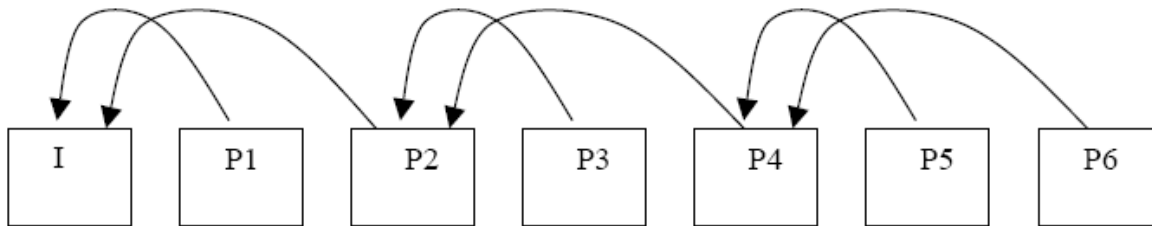
Figure 3.2: Varied coding which aims to decrease the effect of frame drops. In the above figure, if the Inter frame 1, 3, 5 … are lost, the video quality is not affected too much. But, it does increase the important of the super-parents.

3.1.5 BCH Error coding

A block code is something that discretely identifies each message symbol. If there are k-bit data symbols 'u', then the transformed packet representing an l-bit symbol is called as a code word. This set of code words for each of the symbol is called as a Block code. A particular subclass of these block codes is Linear codes wherein the symbols are from the binary field. Cyclic codes are another subclass of Linear codes, which employs shift registers for calculating the encoding and syndrome. Bose, Chaudhuri and Hocquenghem (BCH) codes form a large class of powerful random error-correcting cyclic codes.This class of codes is a generalization of the Hamming codes for multiple-error correction. The implementation uses Binary BCH for its error correction operation. Reed-Solomon code is an example of non-binary BCH. But, it is out of the scope of this thesis. The encoder and decoder for binary BCH codes have been written in matlab for the implementation. The Block length and the Error correcting capability (max. no. of errors

the code corrects) decides the strength. The encoder tries to fix the value of 'm' such that the codeword length can be bounded as: $2^{m-1} <$ length $<= 2^m-1$. [13].

Redundant data (parity bits)	Data bits

Figure 3.3: Codeword generated by the BCH encoder.

3.2 CDMA Channel BER

Code Division Multiple Access (CDMA) is a cellular technology also known as IS- 95. CDMA is a digital spread-spectrum modulation technique used mainly with personal communications devices such as mobile phones. In the spread spectrum system, the digital information signal modulates a digital carrier to produce a spread spectrum signal. CDMA employs different transmission techniques in the forward and the reverse directions. CDMA digitizes the conversation and tags it with a special frequency code. The data is then scattered across the frequency band in a pseudorandom pattern. The receiving device is instructed to decipher only the data corresponding to a particular code to reconstruct the signal. The CDMA systems have two stages of modulation to ensure spread spectrum communication. The bit error rate (BER) that would occur for a CDMA system that does not use forward error correction depends on the E_b/No ratio. The energy per bit to noise ratio (E_b/No), is the energy in the demodulated data bit, to the noise energy in the same bit. It is similar to the signal to noise ratio. The minimum allowable E_b/No that can be used for a particular system depends on the forward error correction scheme used, and the type of data being sent. Voice communications typically requires a BER better than 1% or 0.01. This is assuming some forward error correction is used. Thus when simulating a CDMA channel BER, an upper bound of around 1% should be good estimate[8].

3.2.1 Matched Filter Detector

The standard MF is a single-user detection (SUD) which utilizes the user's own signature sequence to make the best possible estimate of user's transmitted sequence from the raw chip data received at the user equipment (UE). The detection algorithm completely ignores the presence of convolutions due to multipaths in the receiver environment as well as the multiple

access interference (MAI) due to other users sharing the resources [9]. For the DS-CDMA system, MF detector for the ith user becomes

$$\hat{\overline{b}}_{i,MF}^{D} = S_{i}^{H} \overline{r}$$

where $S = diag\begin{bmatrix} \overline{S}_{i} & \overline{S}_{i} & \cdots & \overline{S}_{i} \end{bmatrix}$, and $\overline{S}_{i}^{H} = \begin{bmatrix} 0 & \cdots & s_{i} & \cdots & 0 \end{bmatrix}$

$$\overline{r} = S\overline{b} + \overline{n}$$

Where $\overline{r}$ is the received signal vector, $\overline{n}$ is the (NG+L-1)-d channel noise vector, S is a NG × NK block diagonal matrix with the matrix of spreading codes forming the diagonal elements and $\overline{b}$ is an NK –d vector containing the user symbols with the structures defined as follows:

$$S = diag\begin{bmatrix} \overline{S} & \overline{S} & \cdots & \overline{S} \end{bmatrix}, \overline{S} = \begin{bmatrix} s_{1} & s_{2} & \cdots & s_{k} \end{bmatrix}$$

3.3 Quality estimation using PSNR

Signal-to-noise (SNR) measures are estimates of the quality of a reconstructed image compared with an original image. The basic idea is to compute a single number that reflects the quality of the reconstructed image. Reconstructed images with higher metrics are judged better. In fact, traditional SNR measures do not equate with human subjective perception. Several research groups are working on perceptual measures, but for now we will use the signal-to-noise measures because they are easier to compute. It is important to note that higher measures do not always mean better quality. But, SNR is a fairly good estimate of the perceived quality. The actual metric computed is the peak signal-to-reconstructed image measure, which is called PSNR. Assume we are given a source image f(i,j) that contains N by N pixels and a reconstructed image F(i,j) where F is reconstructed by decoding the encoded version of f(i,j). Error metrics are computed on the luminance signal only so the pixel values f(i,j) range between black (0) and white (255). The mean squared error (MSE) of the reconstructed image can be computed as follows:

Mean square error MSE $= \left(\Sigma \left[f(i,j) - F(i,j) \right]^{2} \right) / N^{2}$

Root mean square error RMSE $= \sqrt{MSE}$

PSNR in decibels (dB) $= 20 \log_{10} \left(255 / RMSE \right)$

Typical PSNR values range between 20 and 40. The actual value is not meaningful,

but the comparison between two values for different reconstructed images gives one measure of quality.

3.4 RATE CONTROL

The FEC introduces redundancy in the stream which causes the overhead. To ensure conformity with the non-FEC video stream, the bitrate needs to be deliberately reduced. Adjusting quantization levels dynamically is an effective solution, as video information cannot be predicted easily.

3.4.1 Effect of Quantization levels

The coefficient obtained after applying DCT are quantized based on a previously established factor. The quantization parameter QUANT is obtained for each macroblock. Figure 2.4 illustrates the variation of quality with the quantization level. The quantization level is also a reflection of the bitrate. Higher the quantization, the lesser bits it takes to encode a particular picture. Thus, a model can be derived for bitrate versus quantization level and the dependence can be used in altering the bitrate for a video stream. In most cases a direct linear proportion can be assumed.

| (a) | (b) | (c) |

Figure 3.4 The effect of quantization level on a frame. Frame (a) is the original picture being encoded Frame (b) is the picture after encoding with a QUANT level of 1.PSNR= 35.5943. Frame (c) corresponds   to QUANT level of 2.5.

CHAPTER 4

SENSITIVITY OF STREAM

TO BIT ERRORS

Influence of each bit on reconstructed quality

Summary of bit sensitivity experiments

This chapter is devoted to present some hypothesis regarding the effect of bit errors on the overall video quality. The performance metric in all discussions is the Peak Signal to Noise Ratio (PSNR). The syntax for a macroblock is as given in Figure 3.1. The importance and effect of each bit is the point of discussion in this chapter. The meaning of each bit will be explained as and when needed. Those bits that lead to a decline in PSNR level, when altered, are considered as the *Weak bits*[10].

COD	CBPCM	MODB	CBPB	CBPY	DQUANT	MVD	MVD$_2$	MVD$_3$	MVD$_4$	Block data

Figure 4.1: Structure of the macroblock layer. The shaded fields are the subjects of discussion in this chapter

4.1 Complete frame drop

This section discusses the type of bit errors and violations that lead to a complete frame drop. Any bit leading to a frame drop is also considered as a weak bit. An error in a chronologically backward frame is of lesser significance than in one before. The weak bit in an Intra frame is weaker than the one in an Inter frame. Hence, the I-frame bits are more important than the P and B frame bits. Figure 4.2 emphasizes the importance of the first Intra frame. From that figure, it can be inferred that loss of other frames does not have a significant effect.
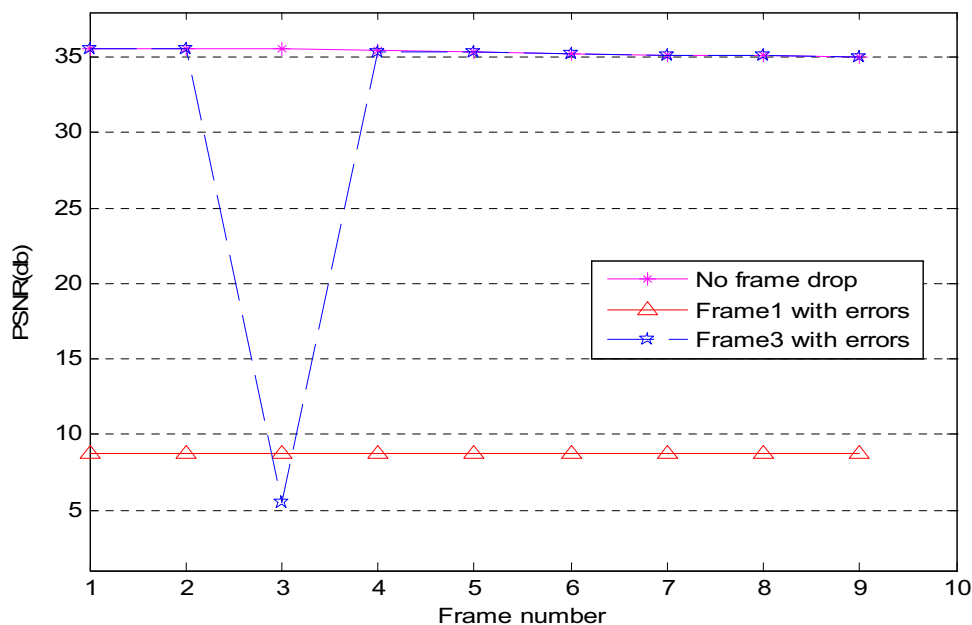

Figure 4.2. Sensitivity of each frame to a single frame drop.Experiment was performed for three different conditions

4.2 Frame resynchronization

The frame resynchronization and its associated data loss are caused when error occurs in the data portion of the picture. When a frame loses synchronization, only the data received thus far is salvaged. The remaining gets substituted by information from the previous frame.

The main reasons for losing synchronization are:

- Presence of an invalid Huffman code in the block

- DCT coefficient index out of bounds

- Quantizer out of range (leading to clipping)

- mvx (or) mvy out of range: searching for sync

- Invalid motion vector, MCBPC, MBTYPE, chrominance CBP (or) CBPY code

The variable length coding employed in H.263 is self-synchronizing. Hence, even a single bit error could cause an invalid code and lead to loss of synchronization. The decoder discards all information henceforth until the next picture start code. This type of error kills only the macroblock starting at that point. Thus, some useful information is indeed passed down history. They are not the most virulent errors. These data errors follow the same pattern as the header errors. The error in I frame is more devastating than those in the other P frames.

4.3 Hypothesis

Based on the picture level and macroblock level syntax, many hypotheses can be stated. Those relevant to the bit sensitivity will be analyzed and discussed in the subsequent section. The weak bits in the stream can be prioritized as follows:

1. The Intra PSC has the highest precedence order. When lost, the whole stream is rendered useless. The Intra header weak bits can cause frame drop and hence is also of high precedence.

2. VLC errors, Motion vector errors & other header errors in the Intra frame can cause resynchronization. These errors have extensive visual quality deterioration.

3. The Intra coefficient errors are long lived and hence are to be protected.

4. Inter PSC has the next highest precedence as it affects only that particular frame and has minimal effect on future frames. This is because each Inter frame has very less information. Only few blocks would have changed. Inter header has the next highest precedence as it affects only that particular frame and has minimal effect on future frames.

5. VLC errors, Motion vector errors & other data errors mandate resynchronization and lead to frame drop. This has minimal effect on overall quality.

6. Inter coefficient errors are also long lived. Though it rarely happens since the macroblock might not be coded owing to "no variation". This does have some influence because it misleads the decoder by providing wrong information. The wrong information passes down history and affects the future frames too.

7. COD errors lead to one or two macroblock drop. They can cause resynchronization and also deteriorate quality by misleading the decoder.

4.4 Influence of each bit on reconstructed quality

The sensitivity experiments performed can be categorized as chronologically oriented and non chronological. That is because most errors trickle down the stream owing to interdependence. The non-chronological experiments try to compare the effect of CBPY with the effect of MV. It basically deals with PSNR levels inside the same frame, for different types of error. This section explores into the relative weakness of each bit. The Intra frame bits are consider separate from the Inter frame bits, as the syntax is slightly different between both. (For example: I frame bits have no COD bit, while P frame bits use them).

4.4.1 Intra frame bits
The effect of each type of bit is discussed below. However, the temporal effects still impose on the I frame bits. The effects of the PSC & other weak header bits that cause frame drop have been discussed in the previous section. The effect of other important bits has been monitored by performing a "bit killing" operation, wherein the critical bits are flipped. In this thesis we have considered only the luminance components and not the chrominance components

- **Coded Block Pattern for luminance (CBPY)**
This is a variable length codeword giving a pattern number signifying those Y blocks in the macroblock for which at least one non-INTRADC transform coefficient is transmitted. When altered, the decoder misunderstands the information present in that particular macroblock and expects different number of non-INTRADC coefficients. It may also lead to invalid code and

eventually loss of synchronization. Figure 4.3 plots the variation of PSNR with corruption pattern. While decoding an MB with a corrupted CBPY, the decoder encroaches the other subsequent MBs. Thus these errors have a greater extend of damage.
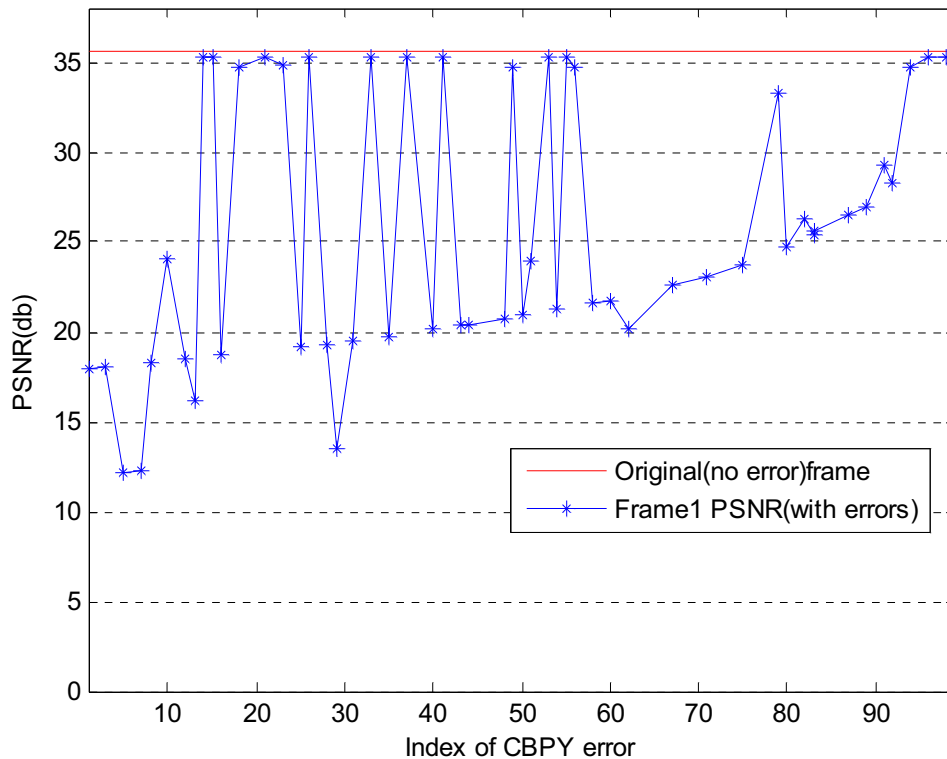

Fig 4.3: Sensitivity of overall PSNR to CBPY error in Intra frame. The larger the index, further down in spatial domain the error is. The slope is upwards owing to the variation in the resynchronization time for each case.

● **DC coefficient for INTRA blocks (INTRADC)**

This is the most important information derived after performing DCT on the macroblock. They are of a fixed length of 8 bits. When altered, these bits mislead the decoder, but do not lead to loss of synchronization. The whole macroblock is affect owing to the theory of DCT. The true video information is lost and the video quality is proportionally reduced. Figure 4 illustrates the behavior noted during experimentation.
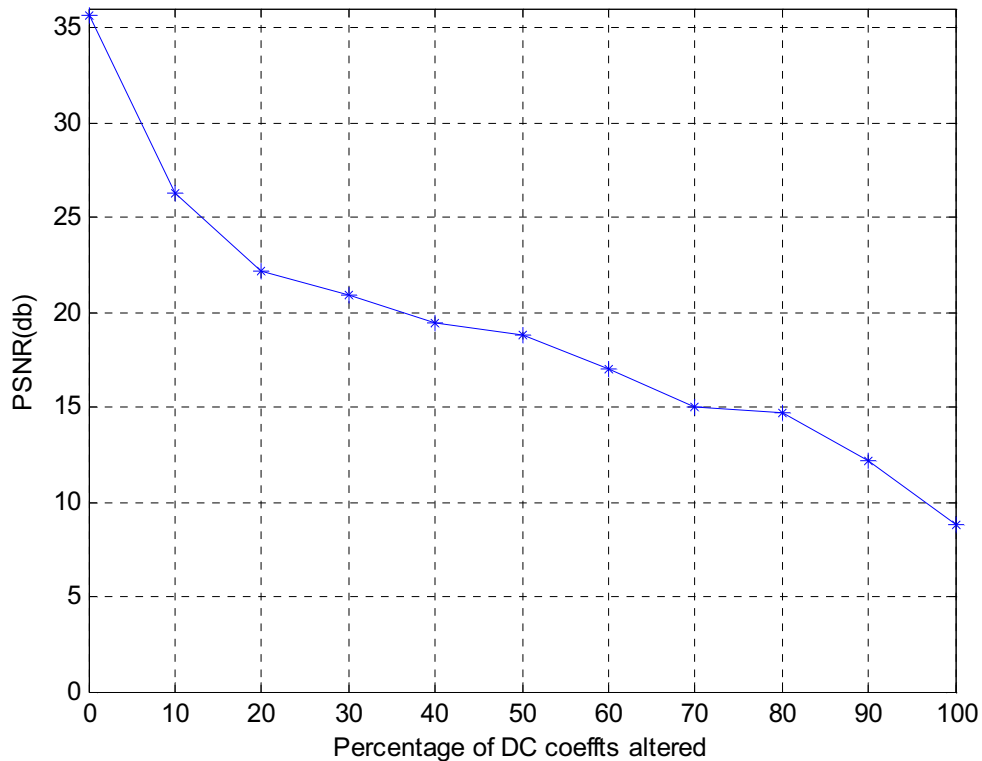
Fig 4.4 Sensitivity of video stream to DC coefficient corruption in Intra frame. More the DC coefficient corrupted, lower the video quality.

- **Transform Coefficient (TCOEF)**

These are the AC coefficient generated after doing DCT. The errors will flood the whole macroblock during spatial IDCT. This may also lead to loss of synchronization and mislead the decoder in most cases. It had direct effects on the reconstructed video quality. The behavior noted is quite similar to that of DC coefficients but for a different reason (resynchronization).

4.4.2 Inter frame bits

The effect of each type of bit is discussed below. However, the temporal effects are not very prominent for P frame bits. Hence the main focus in this section is to throw light on the non-temporal effects. The relative weakness of each P frame bit is studied. The effect of all bits has been monitored by performing a "bit killing" operation

- **Coded macroblock indication (COD) with value '1'**

COD is only present in pictures that are not of type "INTRA", for each MB in these pictures. If set to '1', no further information is transmitted for this MB; in that case the decoder shall treat the MB as an INTER macroblock with motion vector for the whole block equal to zero and with no coefficient data. When this bit if altered, the decoder assumes the MB is coded and starts encroaching on the successive MBs. This is catastrophic (relatively) as it has a greater damage area.

- **Coded macroblock indication (COD) with value '0'**

A COD set to '0' implies that the MB is indeed coded and has coefficients. When this bit is altered, the decoder assumes the MB is not coded. Thus the current MB is skipped. The quality deteriorates based on how much information the MB previously contained. It will make wrong interpretation for the actual data and fouls up the successive MBs as well.

- **CBPY**

CBPY is a variable length codeword signifying those Y blocks for which at least one non-INTRADC transform coefficient is transmitted. When altered the decoder makes wrong judgments about the contents of the macroblock. It may lead to loss of synchronization and eating away of successive MBs. Figure 5 summarizes the effect of CBPY bits (in P frame #3). It can be noted that they have minimal effect on the overall PSNR.
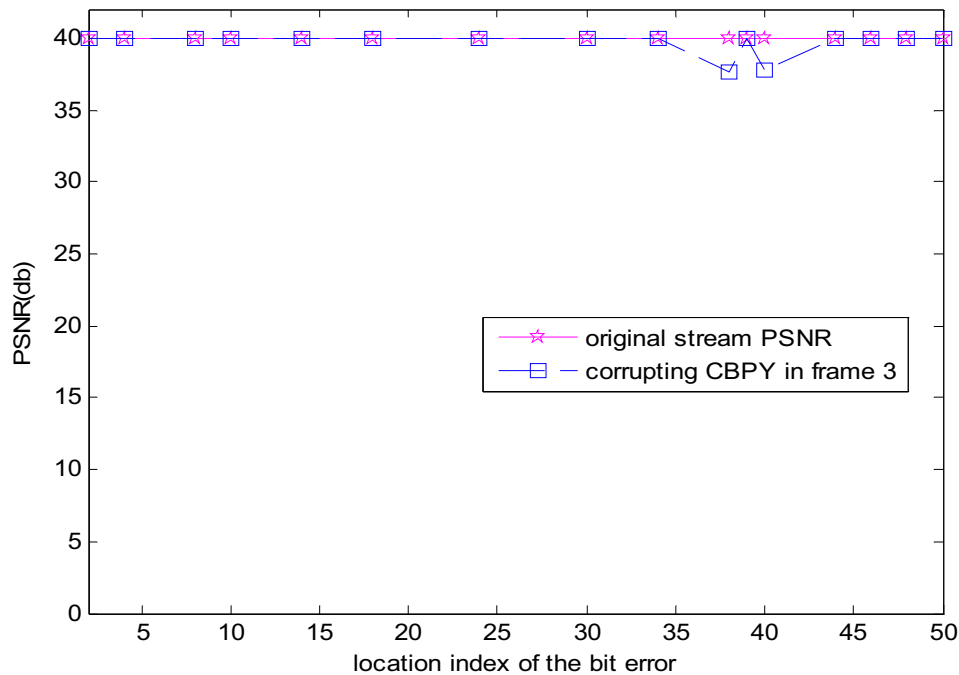
Fig.4.5 Sensitivity of overall PSNR to CBPY error in Inter frame. There
is minimal effect on overall PSNR.

- **Motion Vector Data (MVD)**

MVD is included for all INTER macroblocks and consists of a variable length codeword for the horizontal component followed by a variable length codeword for the vertical component. When corrupted they lead to a loss of synchronization and invalid motion compensations. The effect for the motion vectors is similar to that of CBPY bits.

- **DC Coefficient**

The P frames contain the DC coefficient only when they are coded. They usually do not contain the TCOEF. Owing to DCT the error in a DC coefficient will creep in during spatial IDCT. The corruption of these coefficients misleads the decoder by providing wrong intensity levels. Fig 6 summarizes the effect of altering this information. It can be observed that if the DC coefficients of Intra frame is affected PSNR degrades much, but if the DC coefficients of inter frame is altered there is not much effect in the PSNR value.
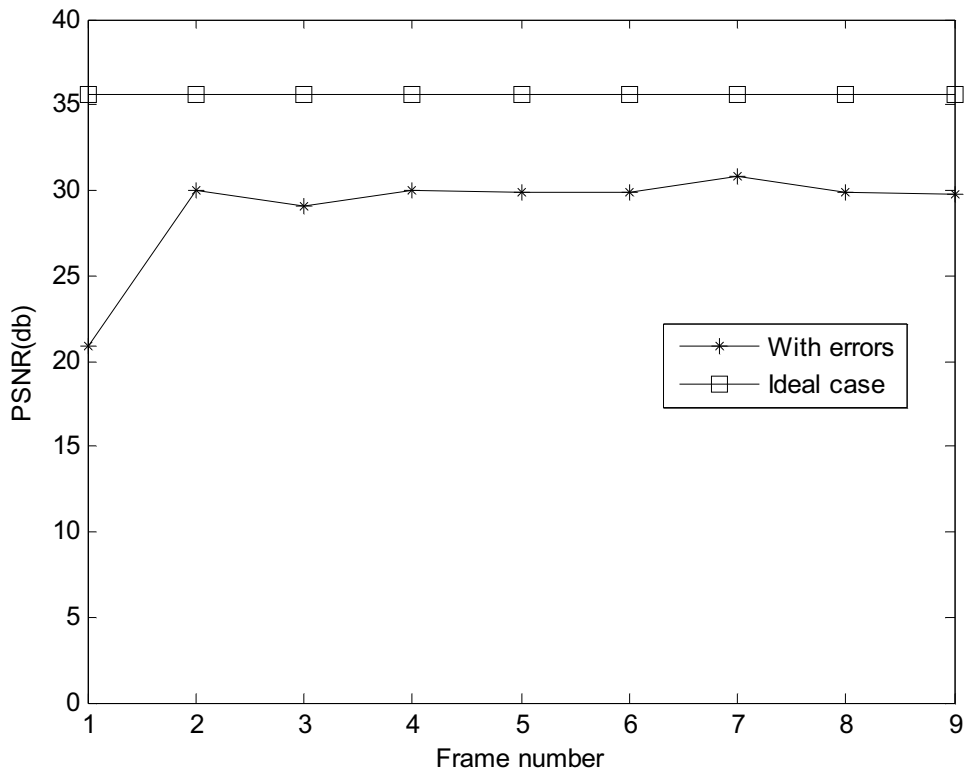
Fig.4.6 Effect of DC coefficients of video quality. The x-axis denotes the
frame being operated on. Each time all the DC coefficients present
in that frame are corrupted.

4.5 SUMMARY OF BIT SENSITIVITY EXPERIMENTS

The main cause for deterioration in quality at high BER is the frame drop (or) loss of frame synchronization. The important points to note from the experiments are:

• An error in a later part of a frame is less important than one occurring before (in spatial domain). This effect is attributed to the resynchronization procedure.

• The first Intra frame has been found to be important for the overall video quality.

• The most important segments in a frame are the macroblock header bits (CBPY). The next important values correspond to the coefficients.

CHAPTER 5

FORWARD ERROR CORRECTION SCHEMES
USING ERROR CORRECTING CODES

Simulation methodology

Result analysis and discussion

Drawbacks

Error control schemes are necessary to obtain good video quality in the presence of high bit error rate. Traditionally, FEC codes have been used for real-time services because they maintain a constant throughput and a bounded delay. FEC codes can be chosen to guarantee certain error rate requirements for the worst channel conditions. The goal of this thesis is to decrease the number of erroneous bits using error correcting codes (ECC). Two FEC scheme has been discussed in this thesis. In the first forward error correction scheme, the data is not prioritized. and various segments in the video stream are given equal importance. The FEC coder supplements each segment with the same number of parity bits. BCH (63, 51, 2) is used for this case. The second scheme is the outcome of the sensitivity analysis experiments. In this scheme, a stronger FEC is applied to the first frame (the first I frame) as the important data is found in it. Hence BCH (63, 24, 7) is used for the first frame and BCH (63, 57, 1) is used for the rest inter frames. The main aim of the algorithm is to strengthen the video stream against most possible error cases and also reduce the percentage of the overhead bits arising because of BCH coding. The strength of the BCH can be varied based on system average BER.

5.1 SIMULATION METHODOLOGY

The EBMA for H.263 was written in matlab. Half pixel precision is used for motion compensation. The module has two segments. One for encoding the raw QCIF video stream and the other to decode the H.263 stream provided. After H.263 encoding, the resultant bit stream was BCH encoded. BCH (63, 45, 3) was used. Then the resultant bit stream was scattered using the Walsh code. Additive white Gaussian noise was added to the resultant spreaded bit stream. AWGN noise was added for a particular SNR after measuring the strength of the resultant stream. Each point on the graph corresponds to only one error count value.

The decoder after receiving the bit stream first performs CDMA demodulation using the matched filter concept and then passes it through the FEC BCH program to correct the errors present based on the particular FEC strategy used. After error correction it decodes the data to generate the raw video again. Finally, the PSNR values are calculated by comparing the input raw video with the output. The value is plotted against the noise factor that was artificially fed in. The value is plotted against the BER, since only the relative performance is of significance.

The simulation was also carried out by varying the signal-to-noise ratio for different BCH strengths. BCH(127,106,3),BCH(127,99,4), BCH(127,85,6) was adopted to see the effect of signal-to-noise ratio on bit error rate. It was observed that if the number of redundant bits is increased or the BCH code can correct more errors, then the bit error rate is achieved at lower signal-to-noise ration. But the disadvantage is that the percentage of the overhead bits is increased.
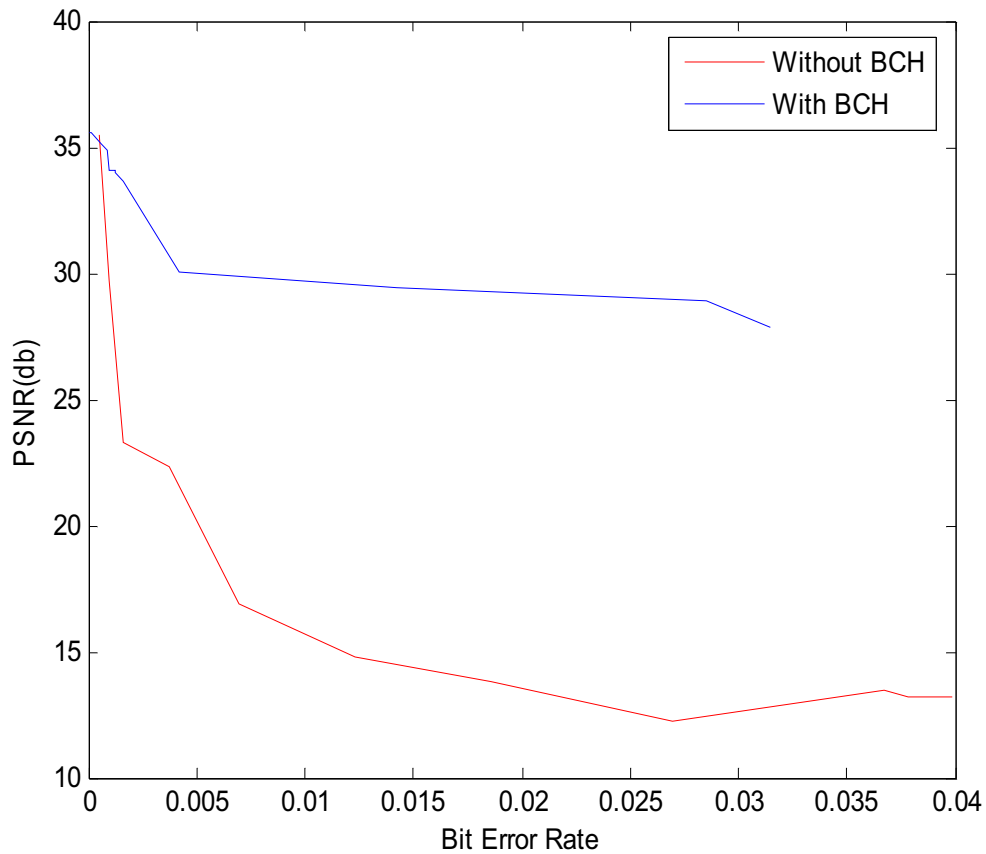

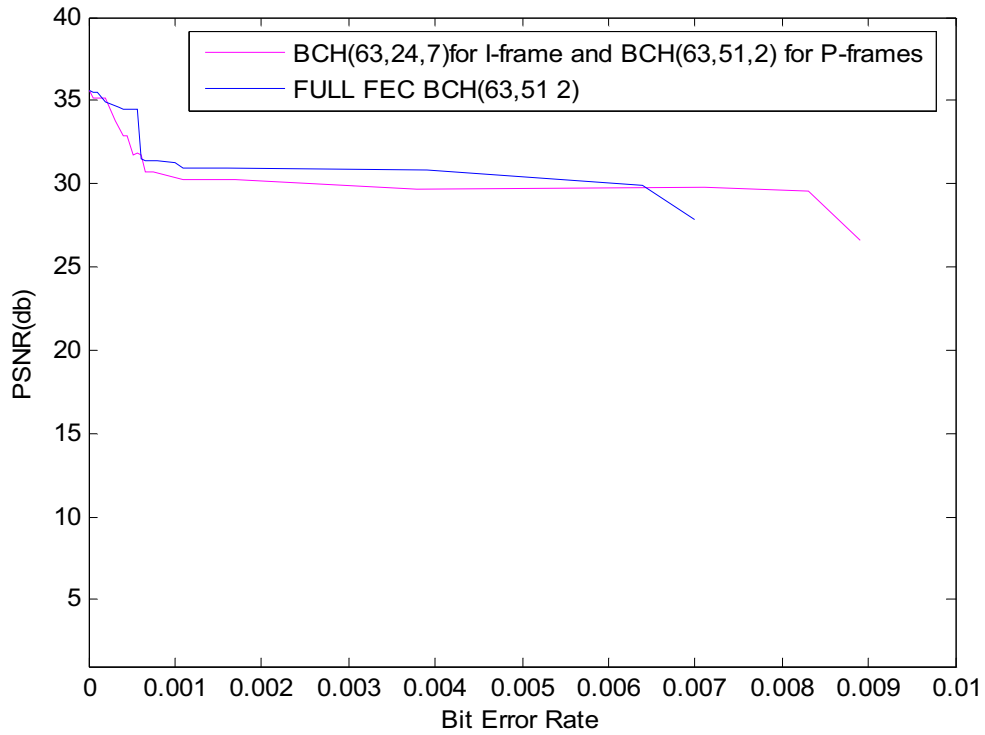Figure 5.1: FEC quality analysis using BCH (63, 51,2)

Figure 5.2: FEC quality analysis for two schemes:(i)BCH(63,24,7) for I-frame and BCH(63,51,2) for P-frame ,(ii) full FEC using BCH(63,51,2)
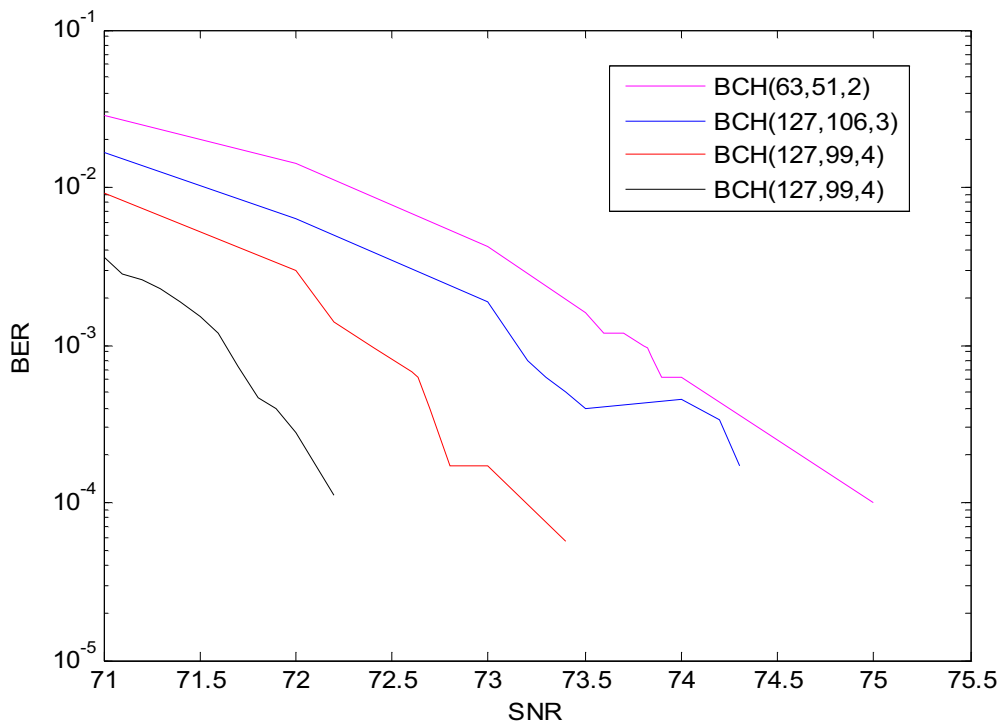

Figure 5.3: Variation of bit error rate (BER) with SNR for different BCH strength.

5.2 RESULT ANALYSIS AND DISCUSSION

The simulation was performed for 9 frames of the "akiyo" video stream and the figure 5.1 was obtained. The error rate was varied from 0-4%. The errors were randomly distributed using different SNR values. For this only BCH (63, 45, 3) was adopted. The figure shows the effect of bit error rate on PSNR values for two different cases: (i) without BCH and (ii) with BCH. When we are using error correcting code i.e.,BCH code ,error correction is made prior to the decoding of data, so the error count is less as compared to the case (i).So the PSNR value is better as compared to without BCH case. It was found that beyond BER of 0.5%, most FEC protected streams suffered loss of video.

Figure 5.2 shows the variation of bit error rate (BER) with SNR for different BCH strength. It can be seen that when the code can correct more errors, the same bit error rate is achieved at lower signal to noise ratio. So we will get better PSNR values at lower SNR. But this comes at the expense of the overhead bits. The percentage of the overhead bits increase as the number of redundant bits is increased.

From the bit sensitivity analysis we found that the first intraframe is most important for the overall video quality. So we can go for a stronger BCH code for the intraframe as other frames are generated using the first intra frame. Hence BCH (63, 24, 7) is used for it. For the interframes, BCH (63, 51, 2) is used. From the figure 5.3, we find the performance is almost same if we use BCH (63, 24, 7) for intraframe or full FEC using BCH (63, 51, 2). But it produces more redundancy as compared to the other scheme. It is also stable at higher bit error rate because weak bits are well protected. The new algorithm is stable till BER of 0.8% as compared to 0.5% of full FEC using BCH (63, 51, 2).

5.3 DRAWBACKS

In this thesis we have considered only the luminance components. This thesis does not does not deal in protecting the chrominance components. This is the major drawback of the thesis. Also CDMA transmission was considered without channel in this work. But in actual CDMA transmission, apart from noise, the bitstream is also affected by channel parameters and which has not been considered in this thesis.

CHAPTER 6

CONCLUSION

Conclusion and future work

6.1 CONCLUSION

In this thesis, we tried to provide better video quality for transmission over wireless CDMA channels. Owing to the applied compression, H.263 is very vulnerable to bit errors due to fading on radio channels. Hence extraneous effort is needed to control error in the video stream.

The bit sensitivity experiment was performed to find out the effect of each bit to the overall PSNR value of the video. The important points to note from the experiments were:

- An error in a later part of a frame is less important than one occurring before (in spatial domain). This effect is attributed to the resynchronization procedure.

- The first Intra frame has been found to be important for the overall video quality.

- The most important segments in a frame are the macroblock header bits (CBPY). The next important values correspond to the coefficients.

FEC scheme has been implemented to control error in the video stream as compressed video stream is fragile in the presence of error. BCH of different strength is also used to observe the effect of SNR on BER. The strength of the FEC code provides another degree of freedom in improving quality. Also from the bit sensitivity analysis we found that the first intraframe is most important for the overall video quality. So we go for a stronger BCH for the intraframe as other frames are generated using the first intra frame. Hence BCH (63, 24, 7) is used for it. For the interframes, BCH (63, 51, 2) is used. The following observations were made.

- Beyond BER of 0.5%, the loss of synchronization problem creeps in which results in loss of video quality.

- When BCH codes of different strength was used, it was found that the code which can correct more errors gives the same bit error rate at lower signal-to-noise ratio but the redundancy is increased.

- When we applied BCH (63, 24, 7) to the first frame, more redundancy was produced as compared to the other scheme.

- It is also stable at higher bit error rate because weak bits are well protected. The new algorithm is stable till BER of 0.8% as compared to 0.5% of full FEC using BCH (63, 51, 2).

6.1 FUTURE WORK

In this thesis we have considered the importance of luminance components only. We can extend this work for chrominance components also in the near future.

CDMA transmission was considered without channel in this work. But in actual CDMA transmission, apart from noise, the bitstream is also affected by channel parameters and which has not been considered in this thesis. So to make the transmission system more practical, channel parameters should be considered during the transmission of the bits and which is an interesting work of the future.

REFERENCES

[1] Rijkse, Karel,"H.263: Video Coding for Low-Bit-Rate Communication", <u>IEEE Communication Magazine</u>. Vol. 34, (December 1996): pp.42-45.

[2] Dufaux,F. and Moscheni, F.,"Motion Estimation Techniques for Digital TV: A Review and a New Contribution". <u>Proc.IEEE</u>.Vol.83, no.6, (June 1995): pp.858-876.

[3] Aroh Barjatya,"Block Matching Algorithms for Motion Estimation".<u>IEEE</u>. Student Member, (April 2004).

[4] Jain, Jaswant, and Jain, Anil,"Displacement Measurement and Its Application in Interframe Image Coding".<u>IEEE Transactions on Communications</u>.Vol.COM-29, No.12 (December 1981).

[5] Wu, Siu-Wai and Gersho, Allen,"Joint Estimation of Forward and Backward Motion Vectors for Interpolative Prediction of Video".<u>IEEE Transactions On Image Processing</u>.Vol.3, No.5 (September 1994).

[6] Chen,Lei and Lee,Chung-wei,"Multi-level secure video streaming Over SRTP".<u>43rd ACM Southeast Conference</u>.(March18-20,2005).

[7] Rhee, I., "Error Control Techniques for Interactive Low-Bit rate video transmission".<u>ACM SIGCOMM '98</u>, (1998).

[8] Toshun, A., and Feng, W.,"On improving quality of video for H.263 over wireless CDMA networks".<u>IEEE Wireless Communications and Networking Conference</u>, (2000).

[9] Waheed, K., and Salem, F.M.," Blind Information-Theoretic MultiUser Detection Algorithms for DS-CDMA and WCDMA Downlink System". <u>IEEE Transactions on Neural Networks</u>. Vol. 16, no. 4, (July 2005).

[10] Zhao, C., Nqan, K.N.,"Error Sensitivity Analysis for Macroblock Bitstream".<u>Proceedings of IEEE TENCON</u>, (2002).

[11] Wang, Y., Ostermann, J., and Zhang Ya-Qin. <u>Video Processing and Communications</u>. Upper Saddle River, New Jersey: Prentice Hall, 2002.

[12] Ohm, J.-R.<u>Multimedia Communication Technology</u>. New York: Springer International Edition, 2004.

[13] Bose, Ranjan.<u>Information Theory Coding and Cryptography.</u>New Delhi: Tata McGraw-Hill Publishing Company Limited, 2002.