SPATIALLY SCALABLE VIDEO CODING (SSVC) USING MOTION COMPENSATED RECURSIVE TEMPORAL FILTERING (MCRTF)

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF TECHNOLOGY

IN

ELECTRONIC SYSTEMS AND COMMUNICATIONS

BY

BEDADATTA BEDANTA

ROLL NO. - 213EE1280

Under the Guidance of

PROF. DIPTI PATRA

Department of Electrical Engineering

National Institute of Technology, Rourkela-769008

2015

National Institute Of Technology, Rourkela

CERTIFICATE

This is to certify that the thesis entitled "**SPATIALLY SCALABLE VIDEO CODING (SSVC) USING MOTION COMPENSATED RECURSIVE TEMPORAL FILTERING (MCRTF)**"submitted by**BEDADATTA BEDANTA** in partial fulfillment of the requirements for theaward of Master of Technology Degree in **Electrical Engineering** with specialization in **"ELECTRONIC SYSTEMS AND COMMUNICATION"** at National Institute of Technology, Rourkela is an authentic work carried out by him 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:

Prof. DiptiPatra

Department of Electrical Engineering

National Institute of Technology

Rourkela

ACKNOWLEDGEMENT

Firstly, I would like to thank Prof. DiptiPatra, Department of Electrical Engineering, National Institute of Technology, Rourkela as my guide, who encouraged and challenged me throughout my project work and guided me on writing conference papers, thesis and to put effort on our workings. She is a perfect motivator and I am glad to have guide like her.

After that, I would like to thank, Prof. Anup Kumar Panda, head of department, electrical engineering, National Institute of Technology, Rourkela and professor of my specialization Prof. Prasanna Kumar Sahu, Prof. Susmita Das, Dr. Supratim Gupta and Prof. K. RatnaSubhashini.

I would also like thanks my fellow lab mates in Image Processing & Computer Vision Lab: Saka harsabardhan, RajkumarMaharaju, G.S.pranav for their help and support.

BedadattaBedanta

Abstract

Through the following years, streaming makers will be progressively tasked supplying enhanced streams of video to gadgets as mobile phones and set top boxes, alongside diverse quality variants for clients to get content on general Internet. While there have been various ways to deal with this issue, including different bit rate feature, one exceptionally solid competitor will be a H.264 expansion called Scalable Video Coding (SVC). It encodes video into "layers," beginning with the "base" layer, which contains the most minimal information of the bit-stream, and then moving towards "enhanced layers" which includes the information to scale up the output. Also SVC gives support for different resolutions inside a single compressed bit stream which is known as spatial scalabilility. In this thesis a problem on SSVC has been addressed. The video sequences had been made scalable in spatial domain. In order to make it more efficient for real time applications, motion compensated recursive temporal filtering (MCRTF) has been implemented. This scheme enhances the efficiency of the components of a visual signal. The temporal filter used here helps in reducing noise arising from the plurality of the frames and the improvised output with reduced noise is used in the process of predictive encoding. Also it eliminates the inherent drift, which arises due to difference between encoder and decoder. As visual signals are always subjected to temporal correlation, motion compensation from the adjacent frames and using it as the reference during the process of predictive coding is of prior importance. The conventional and the proposed method have been used during the encoding process of various video sequences in the spatial domain andan analytical study on that has been carried out.

Contents

List of figuresand tables

CHAPTER 1

1. Introduction

1.1 Scalable Video Coding

The advancement in communication infrastructure, growing demand of internet, technological development have paved the way for interactive media applications such as video conferencing, streaming videos over internet etc. As an intuitive outcome of this, a large group of consumers have been formed which request for the best feature quality wherever they are and whatever their network condition is. Hence, before transmitting the bit stream, the recipient's attributes must be observed and best quality subject is supposed to be provided. As traditionally the regular same link is transmitted starting from the gadgets such as mobile phones to high-performance devices such as HDTV, the emergence of these heterogeneous and non-deterministic systems came up with a big problem for traditional encoder , because they don't consider the 'adaptability' in the streaming applications. To overcome this downside, the feature of scalability introduced in video coding with the ability to support a network infrastructure having number of variants from the view point of video processing capability. The scalable video coding (SVC) is based on the principle of partitioning a single bit stream into multiple number of bit streams containing different but complementary constituents, known as 'layers'. The following Figure demonstrates the concept behind scalable video coding by showing how the input video sequence gets segmented into four different layers, which enables the decoder to choose and decode the layers according to the device attributes and network constraints.This layered structure in SVC can be referred as a concoction of layers namely 'base layer' and 'enhancement layers'. The base layer is meant for the lowest resolution video application where as enhancement layers are the refined than base layers and meant for higher resolution applications [3].Due to multidirectional functionalities and wide range of applications, scalable profiles have been preferred by the encoders rather than the traditional methods of video encoding and proved as a better one for better coding efficiency as compared to non-scalable profiles.

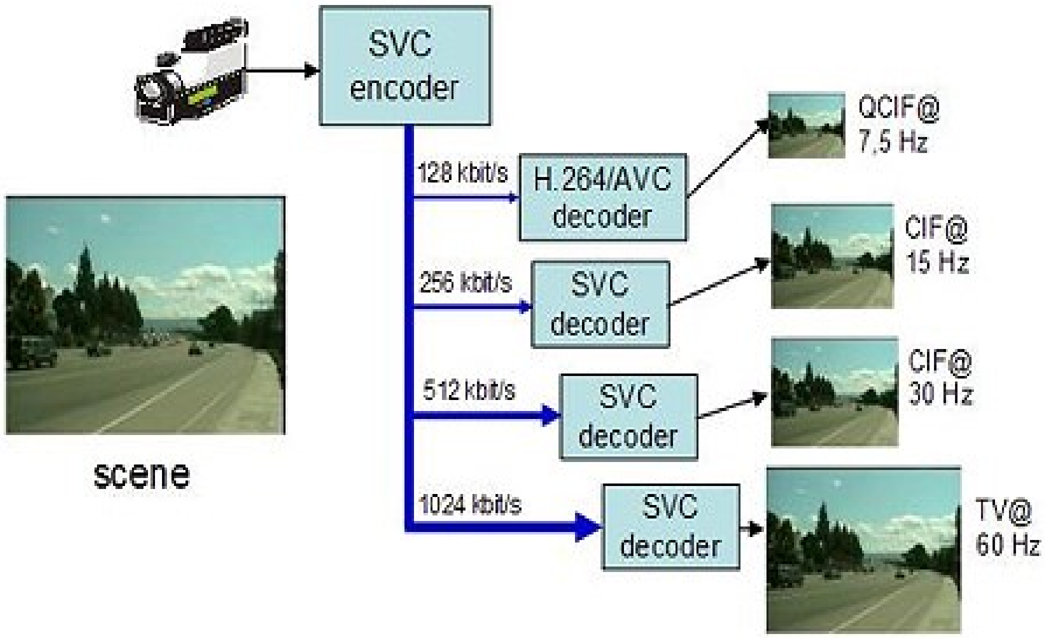

Fig. 1.1.1 Structure of SVC

As a result of this scalable mode, video coding has achieved a widespread range of application and getting utilized in the real time multimedia applications. Hence for circumventing the disadvantages of traditional encoding and answering the needs of various recipients according to the capability of their devices and network attributes, inclusion of scalability in video coding profile seems to be the most preferred alternative.

1.2 Spatial scalability in scalable video coding

Apart from the functionalities depicted above, Scalability in the process of video encoding is also able to work in various modes. The types of scalability inherited from the SVC with the well-designed core coding tools can be broadly classified as: Spatial, SNR and temporal scalability. Among these three modes, Spatial scalable video coding (SSVC) uses the information from all the layers for bringing down the overall size, so that the altogether size of streams can be reduced. This process of utilizing information between the layers is known as interlayer prediction and this constitutes the base of SVC. Figure 2 shown below explains the role of interlayer dependencies in spatial scalable profile.

Figure 1.2.1: Interlayer dependencies in spatial scalable video coding

1.3 Motion Compensated Temporal Filtering (MCTF)

If temporal decomposition will be carried out on a video sequence without implementing motion compensation, the region having motion subjected to blurriness. For eliminating this blurriness present in the frames with low frequency, temporal filtering is applied along the trajectory of motion which can be referred as MCTF. Ohm is the pioneer behind the concept of motion compensated temporal filtering. Researcher Choi and Woods developed this idea further more. It also helps in bringing down the energy present in the high frequency frame, hence consequently supports the enhancement in the efficiency in compression. In the MCTF frame work the motion estimation method used is ‘bi-directional motion estimation’, in which method the process of motion estimation is used for all the blocks present in the frame irrespective of the presence of motion. The videos as Akiyo, News, etc., where the motion is limited and background is fixed, it is quite obvious thatin a frame, most of the blocks will stay static. As the bidirectional motion estimation utilization makes MCTF slow, here removal of the non moving blocks can make MCTF more reliable i.e. the stationary blocks in a video will be identified initially and these blocks won't go through motion estimation. Moreover, MCTF can be a substantial tool for realizing scalable video coding and it has also been proved as a way for improving coding efficiency.

1.4 Motivation

The rapid progress of technology resulted in an exigent group of clients demanding for a better quality of service in multimedia applications. The high bit rates that emerges out of the different sort of Raw videos results in making their transmission through the preconceived channels an arduous task. Even if the issue of storing and transporting of the videos will behandled some way, the power needed for processing and to manage such huge amount of data would make the fabrication of the hardware for the receiver exorbitant. So, dispatching videos of consumer quality will be impossible with raw videos. Hence the compression of raw videos is inevitable for video processing. Again swift rise in number of internet users, numerous devices with a wide range of variation in capabilities resulted in a substantial need of an adaptable, multi-flow bit stream. As an answer to the requirement scalable video coding came into picture. It has also been further modified with new efficient tools and made to support various modes in scalability. As SVC is a direct extension to H.264 with veryless Complexity with the approach layering and numerous functionalities supporting enhancements to transmission and storage applications, For the last few years it has drawn the attention of many researchers and researches have been carried out to build it as the most efficient video codec which will be reliable and subjected to real time applications.

1.5 Literature survey

In the field of multimedia applications video compression plays a key role. The idea of compressing a video began at 1960 with an analog video phone system. Then in 1980 with the availed technology a video codec was designed on the basis of differential pulse code modulation (DPCM). Then,with the help of vivid investigation and research lots of improvement has been brought in video coding such as H.261 in 1988, MPEG-1&MPEG-2 from 1989-95, MPEG-4 in 1995, H.263 in 1996 and many more. Later H.264 and H.265 came into picture, which are considered as a benchmark in the field of video coding. Addition of scalability feature in video coding gave a new direction to the encoders. Many researchers have been researching on various directions of scalable profiles of video encoding. Most of the studies in this literature focused on inclusion of scalability along with additional features and tools in order to make it more reliable in various applications. In this literature survey, I have observed the gradual growth of the video encoding techniques starting from simple codec to a superior encoder as SVC. Along with it, how new features are also getting added to the encoding technique has also been studied.

M. Schobinger, B. Zehner, F. Matthiesen, U.Totzek, J. Hartl, U.Reimann [1] In this paper a video codec has been designed basing on the principle of DPCM. Here a two dimensional intra-field prediction structure has been used. But, the difference between the loops used for prediction results in drift, which can lead towards an acute issue of artifacts.

S.J. Choi, J.W. Woods [2] It has overcome the pit fall of drift. But need of a huge of amount memory can't be avoided while encoding the large blocks. Also the presence of temporal splitting has made it lack of temporal quality.

T. Wiegand, G. J. Sullivan, G. Bjøntegaard, A. Luthra [3]It gives a precise sketch of scalable video coding. Here the major goals and vision of SVC such as boosting up the compression performance, portrayal of both network friendly conversational and non-conversational video has been prioritized.

G. J. Sullivan and T. Wiegand [4] The paper depicts all the details of video codec design beginning from the fundamental concept to how several attributes got incorporated with it and made it standardized internationally.

H. Schwarz, D. Marpe, T. Wiegand [5] This paper outlines the various tools employed in SVC to drop the coding loss and consequently elevating the coding efficiency. It also concentrates on the possibility of incorporating hierarchical prediction structure in order to impart SVC various modes like temporal scalability with a layered approach, consequently rising the coding efficiency.

D. Grois and O. Hadar [6]This paper explains about the integration of Region of interest (ROI) in Scalable video coding. Traditionally the enhancement layers are of fixed size, where as this method employs ROI in it. Adaptive spatial preprocessing also has been put into practice here.

A. Munteanu [7]It gives a brief idea about motion compensated temporal filtering. In this work, a method has been proposed on the basis of motion compensated temporal filtering which can be used in the video encoding in order to Control the distortion fluctuations in video coding systems**.**

1.5Thesis objectives

The thesis strives for:

- To encode the video sequences employing the concept of inter layer prediction in order to make it scalable in spatial domain i.e. making it able to have more than oneresolutions in the single sequence of video and within the same stream.
- To implement a fast method for estimating motion in the spatial scalable video coding, in order to make it more reliable and efficient for real time multimedia applications.
- To incorporate a simple new method of motion compensated recursive temporal filtering (MCRTF) in spatial scalable video coding and comparison of the proposed method with conventional method

1.6Thesis contributions

- Encoding a video sequences and making it spatially scalable.
- Development of a faster technique of motion estimation and compensation in SSVC.
- Implementation of motion compensated recursive temporal filter over spatially scalable video coding to avoid the effect of artifacts.

CHAPTER 2

2. Background theory

2.1 Process of encoding and decoding of Video

A Video unifies a continual sequence of images from a video storage media. The still images are lack of temporal information, but due to the a time variant property videos carry the temporal information. Basing on the applications videos can be systematized into two different categories as: conversational and streaming video. While conferencing via video, video calling etc. falls under conversational video category, multimedia applications over internet, video-on-demand comes under the steaming video group. The Video signals carries both the spatial and temporal aspect of the signal i.e. it alters with respect to both space and time. The signals which fluctuates along with time gets a huge similitude between themselves. Again in a video sequence the constituent frames got a very little period of time like some millisecond. But practically the physical movements in milliseconds is very less. Hence consequently it comes up with the aspect of repetitions in temporal domain referred as temporal redundancy. So, moreover for the compression of a video we need to deal with spatial and temporal redundancy. For spatial redundancy Transform domain techniques are apparently the best way and as far as temporal redundancies are concerned, method of 'predictive coding' is superlative. As video signals are the amalgamated form of spatially and temporally fluctuating signals, both transform domain technique and predictive coding can be hybridized to achieve a efficient compression. The following figure shows the basic structure of a hybrid video encoder. As the video signals are 3-dimensional by nature, For the process of encoding, these dimensions are required to be arranged according to the aspect of spatial and temporal domain. Just as shown in the figure, Transform domain technique and predictive coding technique both has been incorporated to exploit the redundancy up to a considerable extent. The operations executed here can be classified into two groups such as :interframe and intraframe mode. In interframe mode the frame is splitted up into into several blocks called ‘macroblocks’.

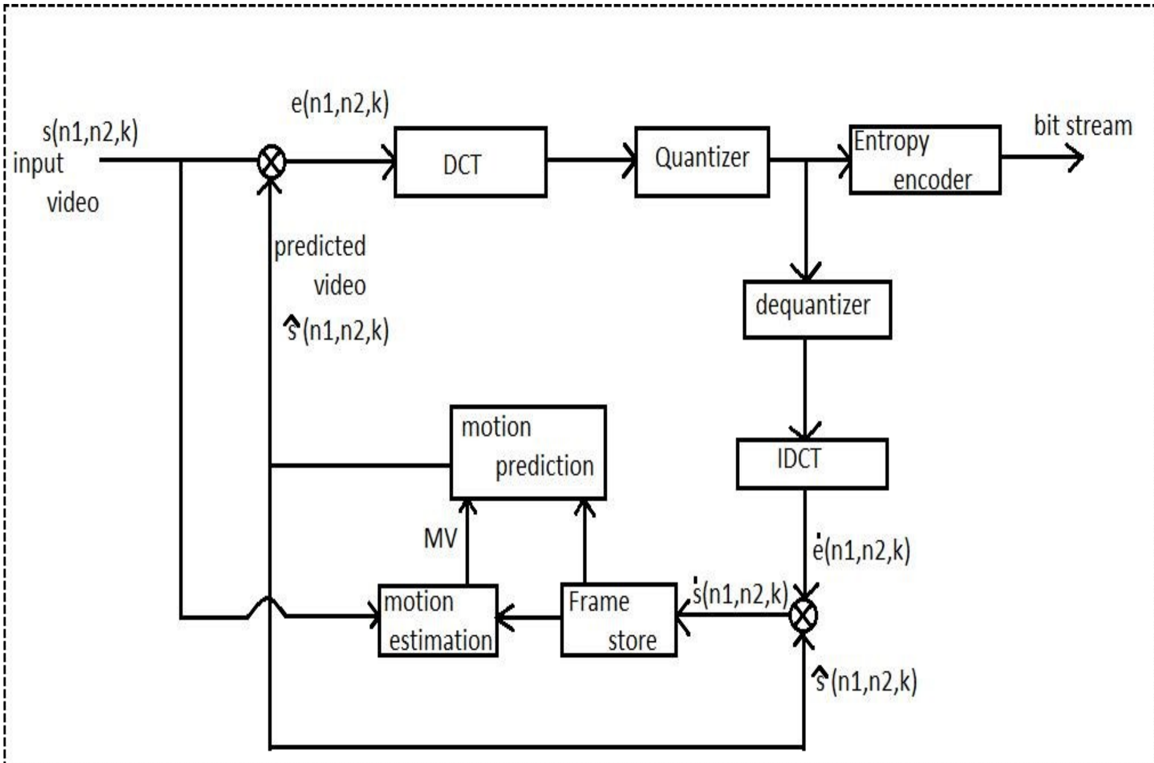

Figure2.1.1Block diagram of video encoder

Rather than encoding the raw pixel values, the encoder strives for obtaining the block same as reference frame. The whole process is executed using a block matching algorithm. Then the encoder seeks for the motion vector, which can be outlined as the information about the degree of movement taken place from frame to frame. The complete governing process for ascertaining the motion vector is known as motion estimation. Motion estimation can be executed in various ways such as forward motion estimation, backward motion estimation and bi-directional motion estimation. When the future frame anticipates the previous frame, it is referred as backward motion estimation and when the process gets reversed i.e. past frame tries to predict about the future frame, it is referred as Forward motion estimation. Also the averaging of both forward and backward motion estimation can result in a better estimate. The process of intra-frame coding can be elucidated as several techniques for compression, executed in accordance with the information which is present in the current frame, and not relatedto remaining frames of the video sequence. To be specific, beyond the current frame temporal processing is restricted in this process. The above said depicts the encoding process of a video in a précised manner. After encoding, the encoded data must gone through the process of decoding. The process of decoding is a reversed process of encoding. The following figure shows the simplified structure of the video decoder:

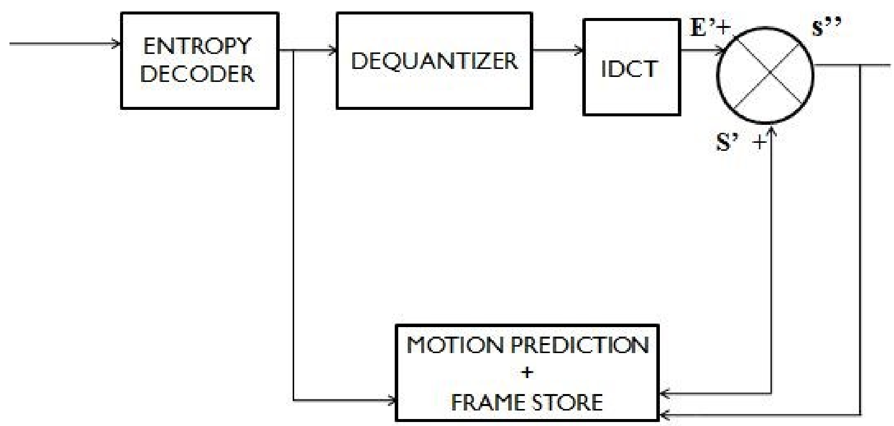

Figure2.1.2 Block diagram of video decoder

2.2 Spatial scalable video coding

The scalability in video coding can be executed in various modes, which can be broadly classified as: temporal scalability, spatial scalability and quality scalability. The quality scalability is a special modified version of spatial scalability. By making a video sequence Temporally scalable multiple frame rates can be obtained from a single bit-stream. The implementation of spatial scalability makes the encoder able to acquire number of resolutions for a single video sequence. Moreover, the spatial scalability is a way of presenting the video sequence,in different resolutions, such that the improvement of the low-resolution based layers will be reliant on each enhancement layer, which is achieved by using an approach of a layering. Summarizing the facts altogether, It can be specified that Scalable video coding, extension of H.264/AVC imparts a technique forreutilizing the encoded sequence having lesser resolution, for encoding theanalogous sequence with comparatively better resolution.The figure shown below, delineatea hypothetical Spatial Scalable video codec (SSVC). Also how various tools got incorporated in the SSVC has been depicted . However,to comprehend SVC design with more lucidity ,detailed concepts have also been emphasized. The scalable design of video codec mainly directed towards two objectives i.e. bringing down the loss in coding efficiency and cutting down the complexity at decoder sidein contrast to traditional single-layer coding. single-layer coding, can be defined as the encoding of a given sequence of video without the inclusion of the profile of scalability. While implementing scalability in video sequences, the substantial challenges could be spatial, temporal,and quality scalability [8] [11]. Out of these three, the mostfundamental type of bit stream scalability, i.e. the spatial scalabilityappears to be the most strenuous one, which hasattained a noteworthy improvement in contrast to simulcast solution.The availed video data with lowest resolution in a spatially scalable bit-stream is popularly called as the base layer which can be decoded with the help of a simple single-layer-decoder and the remaining video data with better resolution is familiar as the enhancementlayer. The procedure for anticipating the enhancement layer by making use of the information availed from the previous contemporary layer of lower resolution is called inter-layer prediction which got inherited from the source known as reference layer. Along with it different varieties of prediction methods can be executed such as inter-imageprediction which follows a methodology of temporal prediction among the number of images available in a single layer of resolution and intra-imageprediction which is referred as the procedure of spatial prediction within a single image of a certain resolution layer. During the design of SVC the spatial dependencies existing between the inter layers plays a vital role. Adaptable and pliant features such as Scaling, parameters for windowing, Flexibility while cropping etc. can be supported by its implementation. Apart from inter layer prediction, Image pyramids play a pivot role in Spatial scalable video coding (SSVC). With the help of image pyramid the correlation between the variants of an image on the basis of resolution i.e. lower to higher resolution can be obtained. This affinity between the images during video processing is being noticed very commonly. Hence involvement image pyramid in video processing especially in spatial mode is of great importance. For increasing efficiency of spatial scalability in video encoding a hierarchy of images ranging from coarse to fine is being utilized [15]. By the use of filtration and decimation, the given original sequence of video of high-resolution can bebrought down to lower resolution. From conceptual point of view SSVC can be designed as an image pyramid which utilizes the process of single loop.

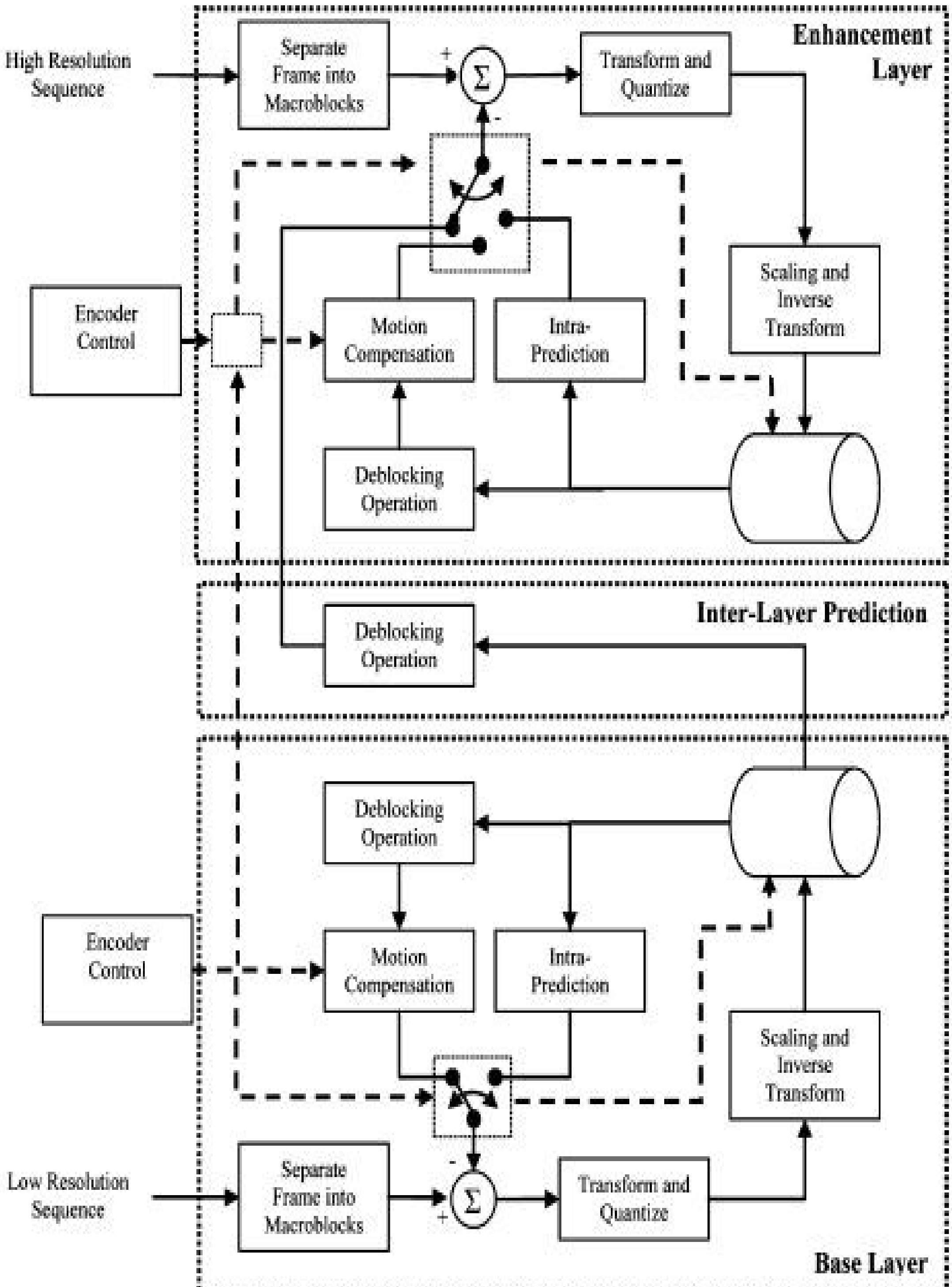

Fig. 2.2.1Detailed structure of spatial scalability in the SVC design

It got an additional major benefit of bringing down the complexity in the decoder side as because the decoder of SVC gets executed on a single loop of motion compensation independent of the enhancement layers going through the process of decoding. For further development in SSVC, rate control algorithm in multi-layer approach, improvements in the estimation techniques of motion, re-sampling operators can also be added.

2.2.1 Requirements from the spatial scalable video coding standard

Being a pivot mode of scalable video coding Spatial scalability in the video encoding standard is subjected towards achieving various goals. Some of the major attributes intended to be obtained by SSVC are listed as follows:

- Obtaining a remarkable enhancement in coding gain i.e. better coding efficiency in contrast to simulcasting for the resolutions present in individual bit-streams.
- Similitude in the coding gain when compared to the approach of single layer encoding for every subset of the bit-stream.
- To reduce any increment in the complexity at decoding side as much as possible
- To support the base layer which is backward compatible
- To achieve the flexibility for supporting the adaptability of bit-stream after going through the process of encoding.

2.2.2 Functionalities and applications

Spatial Scalable video coding is a featured extension of H.264 with a very negligible addition of complexity, which got manifold implications and performances. Out of the innumerable beneficial functionalities some are outlined as follows:

- To put into effect the approach of layering which includes a single base layer, along with one or more enhancement layers. As an additional asset, the generated base layer is an acquiescent to H.264/AVC.
- It got the ability to recreate the signals with lower resolution from the partial bit-streams.

- The most substantial property of SSVC is the ability to decode a bit stream partially, as because this partial decoding attribute consequently allows the following outlines which play a vital role in encoding and transmission.
    - ❖ In some circumstances where the bit stream has been lost partially, it Performs degradation in a very graceful manner.
    - ❖ Adaptability in case of bit rate
    - ❖ Supporting the adaptation for Format
    - ❖ Property of reconciliation for power variation
- It is advantageous during the transmission services while dealing with precariousness which can be either the requirement of resolution in the terminal end, unpredictability in channel condition or devices with various processing and displaying capability.

2.3 Motion compensation:

Video is an outcome of a sequence of consecutive frames ordered in accordance with time. If judged from realistic prospective, the movement within a short span of time is very much limited. Hence repeatability arises due to the similarity and the plurality of the frames which can be referred as temporal redundancy. This redundancy or which can be simply referred as the echo of the nearly similar information between the frames. Therefore the fundamental perception is not to encode each and every frame of the sequence individually and to avoid the encoding of the near-pixel values and hence, consequently reaching towards compression. In the methodology of compression the task of prediction based on motion compensation is eminent.Motion compensation is a method based on algorithm which is utilized in the process of predicting a frame in a video, in accordance to past or future frames by considering the movement of the objects present in the video sequence.

2.3.1 Various stages of motion compensation

The process of motion compensation goes through several stages which have been portrayed in the following figure precisely. The selections made in the previous stages influence strongly the stages executed later. For perceiving the concerns completely which arises with this sort of compression technique, scrutinizing each and every the stage vividly is of major importance.

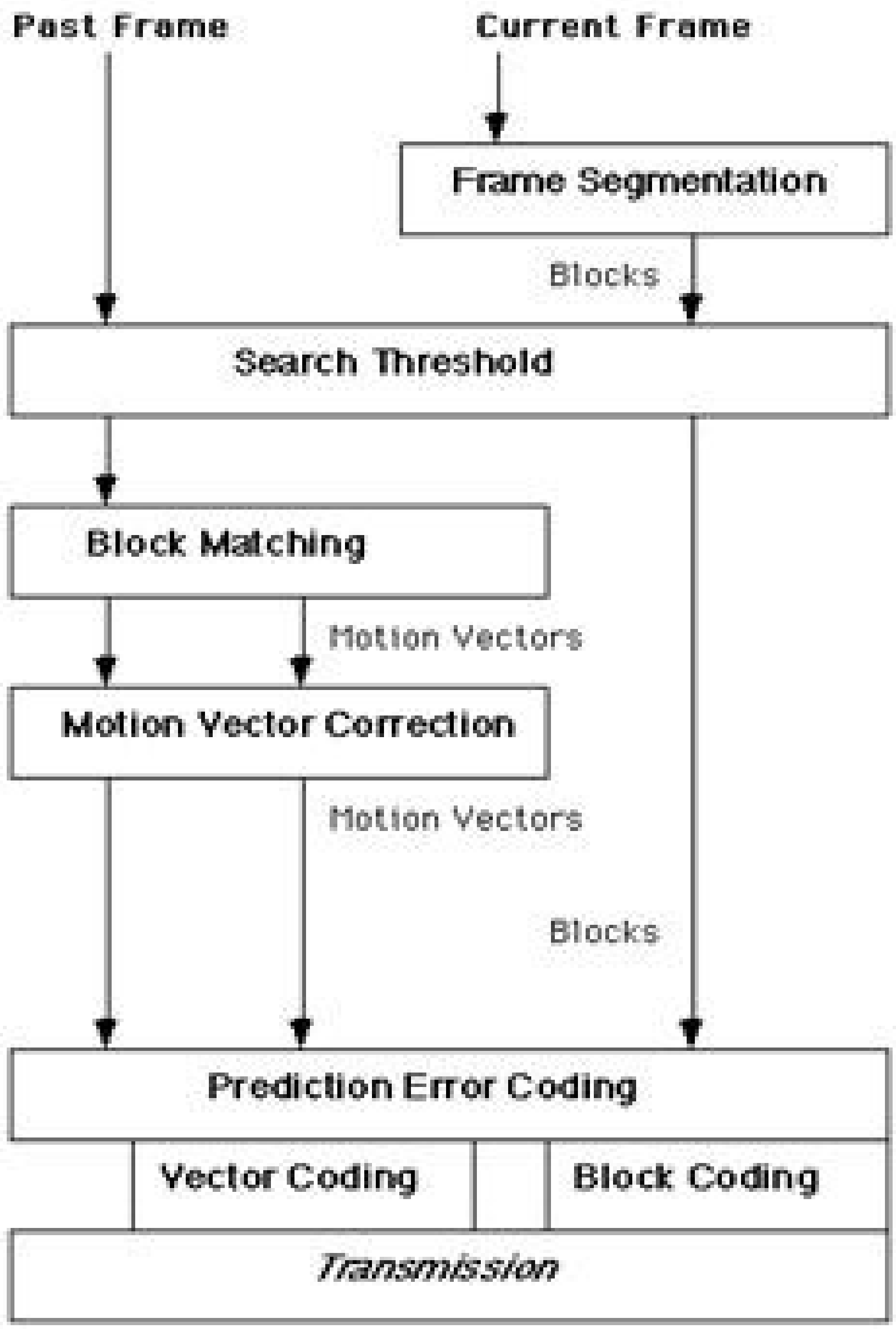

Fig 2.3.1 Overview of the process of motion compensation

Frame Segmentation

In this stage the frame of the video on which compression is going to be held is parted into rectangular blocks, conditionally of equal size and not overlapped. Certainly the rectangular square frames gives a better performance and hence widely used. The Size of the block chosen has a dominant effect on the execution and efficiency of the coding technique. Just because, if the size of the blocks is chosen large, then obviously the number of the total blocks will be turned down. As a result of it evaluation for the motion vectors will be reduced and therefore the number of motion vectors to be transmitted will also be less. But the information sent over for correction will be more. If the case will be reversed i.e. the block size will be chosen small then consequently the increment will occur in the total number of blocks and for obtaining the motion vector number of evaluations will be more. But the requirement of correction factor will be comparatively less. It also have been observed that if the size of the block chosen will be much small, that can lead to a noise sensitive system of compression. Hence the size of the block must be a covenant among the motion vectors reduction and escalation of the refinement of matching blocks. Moreover, the dependencies existing in between the chosen block size, resolution of the frame, coding efficiency must be analyzed minutely.

Search Threshold

For estimating that, whether the objects in the frame have undergone through a significant degree of movement or not,searching for a threshold is required. Initially, subtraction of the candidate and the target block has carried out. Then the value obtained then gets compared with the threshold value fixed before, which is the minimal criteria for determining the movement of a object. If it has been found that the difference between the above said blocks is smaller than the pre-decided threshold, then certainly it will be assumed that the concern object in the frame has not undergone any motion.

Block Matching

During encoding,Block matching is the most prolonged process. While carrying out this process a frame is picked up as current frame and the corresponding blocks are treated as target blocks. Comparison is being done between every target block and the contemporary block present in the immediate previous frame in order to acquire the block of matching. In fact what is being done is, when the current frame gets delivered and reconstructed at the receiver side the matching block found earlier can act as a stand in for the concerned block of the current frame [10]. In a frame only the luminance components can be subjected to block matching.

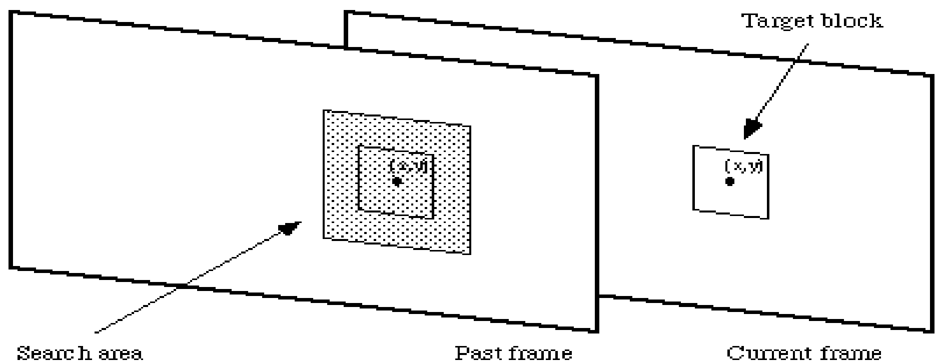

Fig2.3.2 Blocks from a current and previous frame along with the search area

In order to improve the matching criteria, pixels can be incorporated in the predefined searcharea along with the effective enhancement of the resolution property by the implication hypothetical candidate blocks.

Matching Criteria

However we aim at building a compressed block which resembles the block its replacing. Therefore the criteria for matching or the distortion function should be chosen accurately. The matching criteria for the block matching can be defined as a parameter used for quantifying the equivalence among the block of target and the candidate block. The number of evaluations carried out for obtaining the matching criteria is conditional on the size of search area chosen. Increment in the size of search area is directly proportional to the number of computations taking place for the search area. Hence its clearly comprehendible that there exists a very strong influence of matching criteria for the efficient outcome in compression. In case, the matching criteria or distortion function has not been fixed properly, then it will turn out as bad matches. Thus many matching criteria has been developed which can be used effectively in the compression techniques.

Motion vector correction

After executing the block matching methodology over the blocks, somewhere the matching block will be found which can be enacted as the ultimate alternative to replace the current one. Afterward, motion vector is determined which informs the encoder about where the matching block obtained from the past frame is positioned in accordance with the target block present in the frame referred as current frame. In some cases the motion vectors (MV) are not authentic i.e. they won't indicate the motion genuinely. The reason behind this may be presence of noise, inadequate efficiency of the algorithm used for matching of the blocks etc. This inaccuracy can be circumvented by the use of the algorithms which are dependent spatially. Application of the technique of smoothing can also be a way to find out the erroneous vectors and also recommend a substitute. But the use of the methodology for smoothing the motion vector also bears a dark side. It may adjoin remarkable complexity in the algorithm being used for video compression. Therefore the implementation of these techniques is suggested only when subjected to immediate need and much beneficial for any application. Again Smoothing can be resulted as a bad coding of the object which are small in size just because in some cases motion vectors are considered erratic while they are accurate in reality. if smoothing will be executed for such kind of motion vectors, it will badly affect the coding efficiency of the concerned frame. Only a sophisticated and ideal methodology for correction will be able to eradicate such kind of issue. Instead of the cons depicted above, still smoothing is an appreciable technique in compression as because it can bring down the requirement for a huge consignment of data needed for transmitting the motion vector and the reason lies behind this that the information is compressed appreciably.Consequently the process compression of the motion vector can be carried out with much effectiveness and substantially improve the efficiency of the compression.

Vector coding

The methodology of deciding a motion vector during the process of video compression has been discussed so far. After motion vectors getting obtained, it is provided to the bit stream. Commonly, more than half of the data compressed, includes the motion vector. Hence for this reason the coding efficiency of the motion vector influences strongly the coding gain achieved in the whole process of compression. The compression also can be executed further because of the similarities existing among the motion vectors obtained. But it should be taken care of that the compression carried out in the above said process must bear no loss. Fortunately, many researchers have worked on this context and number of algorithms for compression came into picture.

Prediction Error coding

As discussed previously, ideally, we need to construct a perfect replication of the given original frame undergone through compression. But hardly the above said can be accomplished successfully. Hence the distinctness between the frame anticipated and the raw frame should go through the process of encoding. Commonly, this task can be performed basing upon block to block and provided that the fragment of the frame encoded must be considerably dissimilar from each other. For this purpose Transform coding can be utilized.

2.4 Motion Estimation:

The video signals are abundant in temporal redundancy. Exploitation of which is inevitable for a better coding efficiency. To achieve the purpose stated above, motion compensation is a significant tool. In fact, a great deal of resemblance between the consecutive frames forms a basis for motion estimation. The fundamental goal of applying motion estimation is acquiring the motion vector in order to anticipate the best alternative block during the process of prediction. For this purpose two frames are needed to be selected, out of them one is current frame and another is referred as frame for reference. The motion estimation process inspects the blocks that have not been gone through any motion and the blocks which have gone through. For a transparent comprehension a figure has been shown below. In between two frames shown in the upper portion the first one refers to the current frame and the second one is the next frame anticipated. Here we will observe how motion estimation will address this issue and what procedure it will follow to present the changes taken place in between the two frames. Let us observe first, the changes taken place from frame to frame. Then a search has been carried out basing on some matching criteria and as shown in the figure various results will be obtained. To sum up , it can be said that subtracting desired and predicted image the erroneous vector can be found which is subjected to encoding first and transmission afterwards.

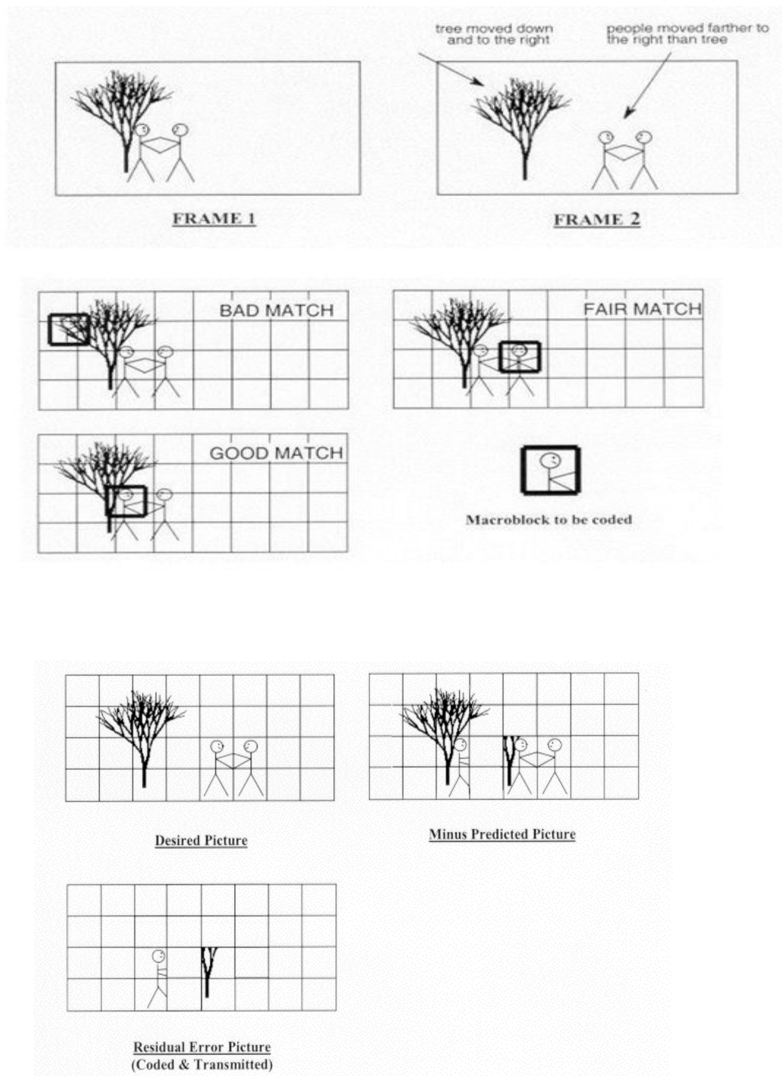

Fig.2.4.1 Showing the processes altogether for estimation of motion

2.5 Role of temporal filtering in Scalable Video coding Standard

In spite of enormous development in scalable video coding, distortion is still a tough issue. The issue of 'Distortion' can be referred as the artifacts that affect the quality of the video adversely when get decoded at the receiver side. The grounds behind this issue can be broadly classified into two categories such as: quantization of the transform co-efficient and the prediction using motion compensation. Analyzing the above said, it can be found that first source of artifact is while performing quantization, relying upon the coarseness, Some edges will arise which are annoying from the visual prospective. Moving towards the second reason which is the prediction process using motion compensation. In the process of motion compensated prediction, the blocks to be predicted from the adjacent frames which have been decoded locally and consequently carries disruptions at the boundary of the block. Commonly, they replicate to the interior section of the signal used for prediction. Hence researches have been carried out on deblocking filtering. Also the standardized codec H.264/AVC elucidates a filter for deblocking that follows the algorithm which strives to discriminate among various disruptions by performing the methodology of boundary analysis. In this context it can be assumed, relying on the fact that what sort of neighboring blocks are being used (intra coded/ inter coded), boundary conditions are identified i.e. either critical or not. The next step followed is filtration of the edges in vertical and horizontal fashion in the spatial domain. Henceforth refinement can be done substantially in the overall resolution and the prediction signal, Still in the decoder side during decoding of the low resolution video signal artifacts will be perceived. So as an answer to this problem a novel approach of coding can be utilized which is based on filtering the artifacts in the spatially lined up images temporally. Moreover filtering of a video signal temporally is of great importance. Therefore, implementation of Motion compensated temporal filtering (MCTF) in the SVC environment will certainly enhance the efficiency of coding. MCTF is a kind of wavelet transform. Because of that it can be enacted either by traditional methodology or by the use of lifting technique. Traditionally, when the low pass and high pass filtering is applied, it follows the process of parallel computation. But when both are implemented in the scheme of lifting it may be looked as a serial computation. While carrying out the motion compensated temporal filtering, the reconstruction of the filter is of great importance as the quality of the result rely on this. Hence to fulfill this requirement the filter should be designed using the lifting scheme. Besides, for MCTF which is integer-pel accurate, application in the traditional mode as well as lifting scheme is identical. But while MCTF is based on sub-pixel accuracy, the outcome will be furnished with some fine attributes proficient in SVC environment. Furthermore, employing the lifting approach the number of computation can be declined as it supports the computation of the transform in-place, where the data provided as input is substituted by the transform data despite entailing a storage place for the output distinctly.

CHAPTER 3

3.Proposed technique for spatial scalable video coding using motion compensated recursive temporal filtering

3.1 Introduction

Infrastructure development of networks, invention of number of devices, elevation in the number of internet users etc. together become a cause for growth in the number of video applications. Also transmission of video and accessing it over internet from different terminals, personal digital assistance has come into picture in the modern technological environment. Therefore, the scenario portrayed above demands for a content carrying video with fidelities distinct from each other, for instance, content of high quality which is meant for storage applications and editing in future. Traditionally when it was needed to send a video over some networks which are diversified from each other widely, in that case the processing was being done offline. The above methodology can be accomplished only by two ways, one is to transcode the content which has gone through the compression already and the other is to construct and accumulate various versions of the same encoded video to be transmitted. But these solutions are not satisfactory as because surrounded by many difficulties like memory requirement, bandwidth allotment etc. In addition to it, conveyance of videos over diversified channels which are vulnerable to errors came up with difficulties like erroneous bit, loss of packets, error propagation and many more. These issues cannot be abandoned as it is, because it adversely affects the transmitted video quality at decoder side. In some instances there is also a possibility of receiving ineffectual purposeless output at decoder. Moreover all the complications and difficulties delineated above are required to address an alternative, which must be robust, impenetrable to error, adaptable with the end terminals. The solution is none other than to enact scalability in the profile of video encoding. Scalable video coding is flexible enough to compliant with spatial and temporal resolution, displaying capability of the devices, requirements for transmission. SVC can be further improved and made more reliable for multimedia applications by revamping the techniques of motion estimation. Because if the entire process will be analyzed, it will be found that the motion estimation is the process which consume more than 60% of time. Also it is the most influential one while deciding coding efficiency. Carrying out motion estimation in SVC mainly aims at extracting the true motion and discarding redundancy accordingly [13]. So, very clearly upgrading the techniques for motion estimation is of vital importance to increase the compression efficiency. Along with this another important issue also exists as an obstacle for efficient video compression i.e. Noise reduction. Because presence of noise not only affects the concerned frame but also reduces the effectiveness of the subsequent tasks going to be carried out further. So, filtering methods are needed to be employed to circumvent the above said issues. Commonly, these kinds of issues can be dealt with two courses of action. Out of them one is to construct the filter with the attribute of adaptability. Another one is to exploit the nonstationarity by the process of motion compensation. Both of the approaches can eliminate the motion artifacts arising during video processing effectively.

3.2 Improved Motion-Estimation by better motion-vector coding

Motion estimation can be defined as a methodology to obtain the motion vectors by the help of which we can get the information about the movement of the objects from frame to frame in a video sequence. Motion estimation follows various methods like forward motion estimation, backward motion estimation and bi-directional motion estimation. To decipher it properly we need to take two frames from a video sequence, out of which lets take one current frame entitled as K and another frame K-1, referred as past frame. Now let us consider that the current frame will be encoded in accordance with the past frame i.e. the frame K will be encoded referring the frame K-1. Elaborating the above said, it can be stated that we are aiming at anticipating how the current frame is going to be on the basis of the objects position in the immediate adjacent frame. First motion vector is going to be found out and then it will be implemented on the frame we have considered as reference i.e. the K-1. Then motion vector (MV) and reference will together construct a frame which can be referred as predicted frame. Afterwards the difference between the original current frame i.e. K and predicted frame is taken into account and only that difference goes through the process of encoding. If we will analyze the whole process, it is observable that for the prediction we are referring the past frame which means we are going back in the prospective of time and trying to estimate the future, hence this can be designated as Forward Prediction. For further analysis we need to reverse the case i.e. we will consider the future frame as the reference frame. Taking the future frame as a reference frames sounds confusing because as it has been told that the frame is 'future' frame that indicates that it has not been received yet. So it is questionable that how to take an obscure content as reference. To clarify it, it can be stated that to carry out the above said process properly we need to accumulate the frames previously i.e. a prior storage of the sequence of the frames, which will enable us to have the frame referred as future frame in the storage memory location. Now the same procedure as we have followed previously for the first case will be followed. This time, referring the future frame, the current frame is getting predicted. So clearly we are going backward in time. Hence it can be referred as 'backward prediction'. Sometimes the past frame doesn't have enough information to anticipate the future frame. In that case forward prediction can't be put into action and only backward prediction works. The situation also reverses in some cases i.e. the future frame doesn't have much information about the past frame and hence unable to predict it. In this scenario one should take help of the forward prediction. As previously we are storing the frames, a better idea can also be suggested which is a halfway i.e. collaborating both of the techniques. Blending of both the techniques inherits the new efficient one, known as 'bi-directional prediction'. This prediction technique is a better alternative than the previously discussed ones. Because it is based on averaging of both forward and backward prediction technique and it's a well known fact taking average provides a finer result than individual estimation as it even outs the error occurred at individual calculation i.e. if one computation is erroneous that could be rectified by one accurate estimate. In addition to it, in some instances occlusion persists in the past frame and in some cases in the future frame. In such cases, bi-directional prediction is very much advantageous. During encoding, motion estimation is lengthiest process as it consumes a major portion of the time taken for the whole process. It can be accomplished in two ways pixel matching and block matching. Pixel to pixel matching a very accurate method as it explicitly finds suitable matches. But unfortunately, it is not a preferable solution as it takes a huge amount of time. So the alternative left is block-matching. In block matching a target block is picked from the current frame and goes through a process of comparison with the contemporary immediate past frame in order to obtain a suitable matching block which can stand as a substitute to the current one. For achieving the 'perfect match' some constraints get decided which can be stated as 'matching criteria'. This can also be referred as 'distortion function' being used as a parameter to extract the similitude between the blocks chosen. The method followed in this work aims at reducing the time required for motion estimation. This idea was coined by lallauret and barba. The motion vectors are found for the whole sequence and then grouped into blocks. it has been found that whether the vectors are identical or not. If not so then all will be transmitted but it they are found to be identical then only one will be sent which can save a huge amount of time. This above procedure can be summed up as:

Step1: From a bit sequence consecutive frames have been chosen

Step 2: The frames obtained are getting segmented into number of non overlapping bocks;

Step 3: Then motion vectors are computed as they describe the degree of displacement from frame to frame.

Step 4: Once the motion vectors assigned to the concerned sequence for their representation, they are grouped into several blocks.

Step 5: Then for each block, comparison between the motion vectors have been carried out.

Step 6: if it has been found that in a block the motion vectors present are same, then from the concerned group only one will be transmitted.

Step 7: if a block doesn't contain a similar group of motion vectors then it goes through the conventional method of vector encoding.

3.3 Motion compensated Recursive Temporal Filtering

When we are dealing with the bit rates varying from each other in a wide range, along with diversified devices and contrasting network condition, implementation of scalability in the level of bit stream is ineluctable. The compression methodology prevailed earlier, used to follow a scheme of compression designed on the basis of motion compensated prediction which employs a recursive loop in the process of prediction. As discussed earlier, a hybridized coding technique is required to be enacted during encoding of a video in order to exploit spatial and temporal redundancy. From the practical prospective if we will analyze, the exact duplication of motion compensation at the encoder side is not possible for the decoder. Because in encoder the lossy component like quantizer has been used. So some error is expected in this case and the error originated gets assembled at the decoded video sequence. As every frame undergoes the process of motion compensation, after a certain interval of time the assembled error will be substantially huge and consequently results in remarkable distortion. This severe disparity can be outlined as 'drift'. The inheritance of this drift property, which originates due to the dissimilarity between the encoder and decoder, builds an unfavorable environment for scalable encoding and needless to say makes the compression process amateurish by resulting in a bad coding gain. Various approaches has been emanated to limit the drift property during hybrid encoding[8][9]. Some of the approaches has been enlisted below such as drift compensation where it has been attempted to optimize the drift value during encoding so that an equilibrium can be maintained. The main pro of this method is, it just focuses on limiting the drift by optimizing it at encoder without making any alternation at the decoder. Some other methods are also there which performs better than the above said such that drift clipping, drift leaking, progressive FGS etc. In the drift clipping technique a certain value gets decided previously as a maximum drift value. The decision is being taken on the basis of the step size of the quantizer of the base layer. If the chosen value is reached, then the dynamically the drift gets restricted. In drift leaking, on the basis of operational target and characteristic of the sequence a drift coefficient is being fixed known as leak co-efficient. By multiplying drift and leak coefficient drift can also be restricted.

Recent researches have unveiled that motion compensated temporal filtering (MCTF) can be solution to this. This can act like a tool for further increasing the efficiency of SVC. Ohm has given the first thought of utilizing MCTF [12][13]. Then the concept was thrived properly by choi and woods. In MCTF a set of trajectories of motion has been given and then temporal filtering is being done over the trajectory of motion on the frames of the considered video sequence. Here a straight forward approach has been used i.e. recursive filtering to Spatial scalable video coding. To precisely elucidate the design of a filter, it can be stated that a mechanism used to eliminate some frequency from certain signal and to support some other frequency to traverse through the output. The term 'Recursive' only defines the trait of the filter. Recursive means repetitions. This filter utilizes the output yielded previously to evaluate the current output. Here the motion compensation will be held in the spatial domain. Suppose two consecutive frames are there, then the odd numbered frame will be considered as candidate frame and the even numbered frame is considered as reference frame. Then by following a proper block matching algorithm predicted frame can be found. The method followed for prediction strongly influences the compression efficiency. Because if a refined prediction strategy will be followed then it will result in the declination of the significant bits in the availed error frame. The error frame becomes a major source for the inheritance of the updated frame. For the process of updating the information is needed to be accumulated from the motion vector. Afterwards the updated frame can be obtained which is constructed by making use of the availed error frame and the motion vector.

CHAPTER 4

Results and Discussion

4.1 Scalable encoding of a video sequence in spatial domain

The video sequence which has been taken into consideration is ‘Rhinos.avi’, which is at complete resolution. First it has gone through low pass filtering and next to it down sampling for the base or lower resolution. Then Encoding has been taken place. After that the difference present among the video sequence taken as input and the video decoded at base layer locally is evaluated. The difference obtained from the above said procedure has been coded by utilizing discrete cosine transform. The two bit streams are stored. In order to get the video of lowest resolution or base resolution, the base layer is required to be decoded. For acquiring the full resolution, decoding has been carried out on both base and enhancement layer. Then up-sampling and filtering have been executed on the decoded frames which are of base resolution and then the output obtained from the two has been summed up. The scale used for quantization is of equal size for both the layers. Consequently the video sequences which are decompressed are of equal quality but differ at resolution [14]. This is referred as spatial scalability. For the experiment the set up has been used as follows: block is of size 8*8 pixel, search window of size 16*16, bicubic interpolation, variable length coding has been used. 50 frames have been taken for experimental purpose.Variable Length coding has been used. Decoding the base layer yields a lower spatial resolution video and adding the enhanced layer to the up-sampled base layer results in a higher resolution video.The reconstructed block plays the role of the reference block.

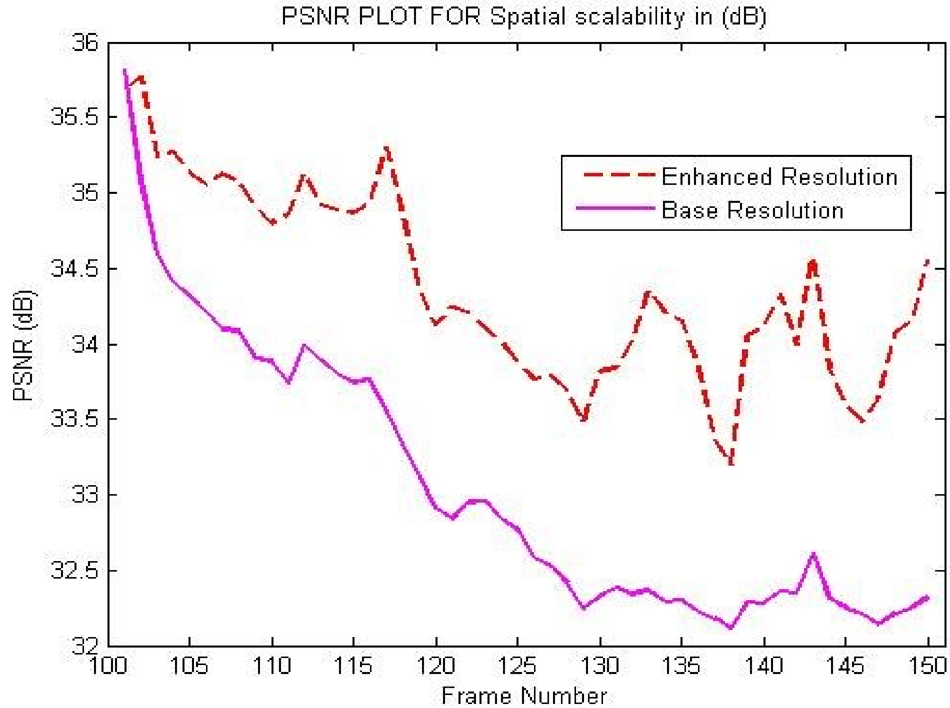

Fig. 4.1.1 PSNR plot for spatial scalable video 'Rhinos.avi'

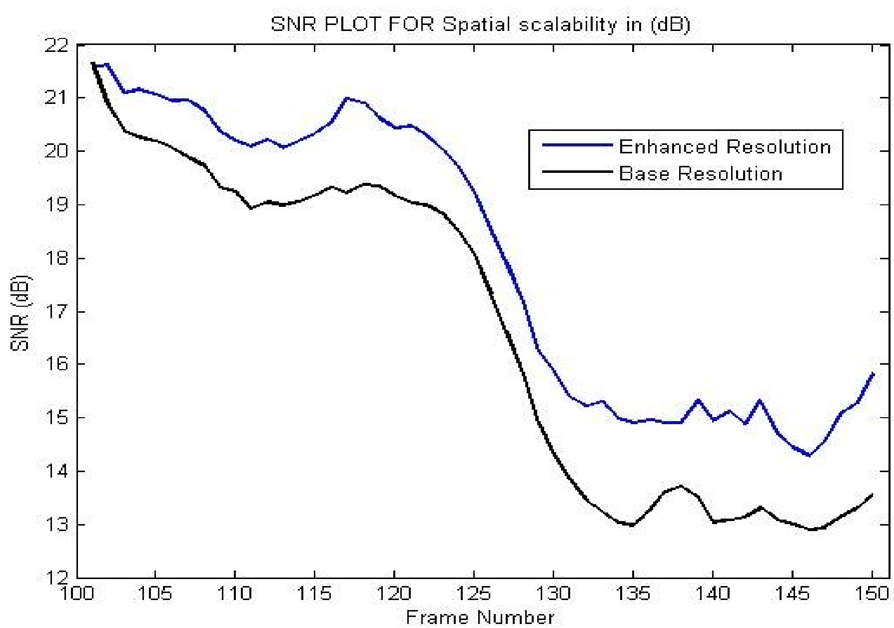

Fig. 4.1.2 SNR plot for spatial scalable video 'Rhinos.avi'

4.2 Improvising the method of motion estimation

For improving motion estimation following algorithm has been implemented:

Step1: From a bit sequence consecutive frames have been chosen

Step 2: The frames obtained are getting segmented into number of non-overlapping blocks;

Step 3: Then motion vectors are computed as they describe the degree of displacement from frame to frame.

Step 4: Once the motion vectors assigned to the concerned sequence for their representation, they are grouped into several blocks.

Step 5: Then for each block, comparison between the motion vectors have been carried out.

Step 6: if it has been found that in a block the motion vectors present are same, then from the concerned group only one will be transmitted.

Step 7: if a block doesn't contain a similar group of motion vectors then it goes through the conventional method of vector encoding.

Following above said methodology the motion estimation can executed in an efficient way. The video sequence used here is 'Rhinos.avi'. The SNR and PSNR values have been plotted for the considered video sequence. 20 frames from the video sequence have been taken for experimental purpose.

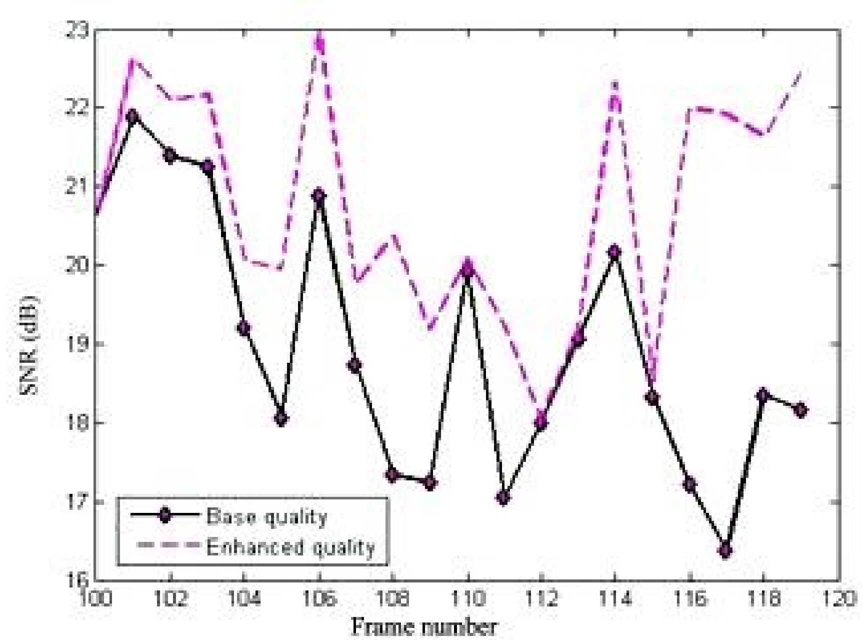

Fig. 4.2.1 SNR plot for 'Rhinos.avi' with improvised motion estimation

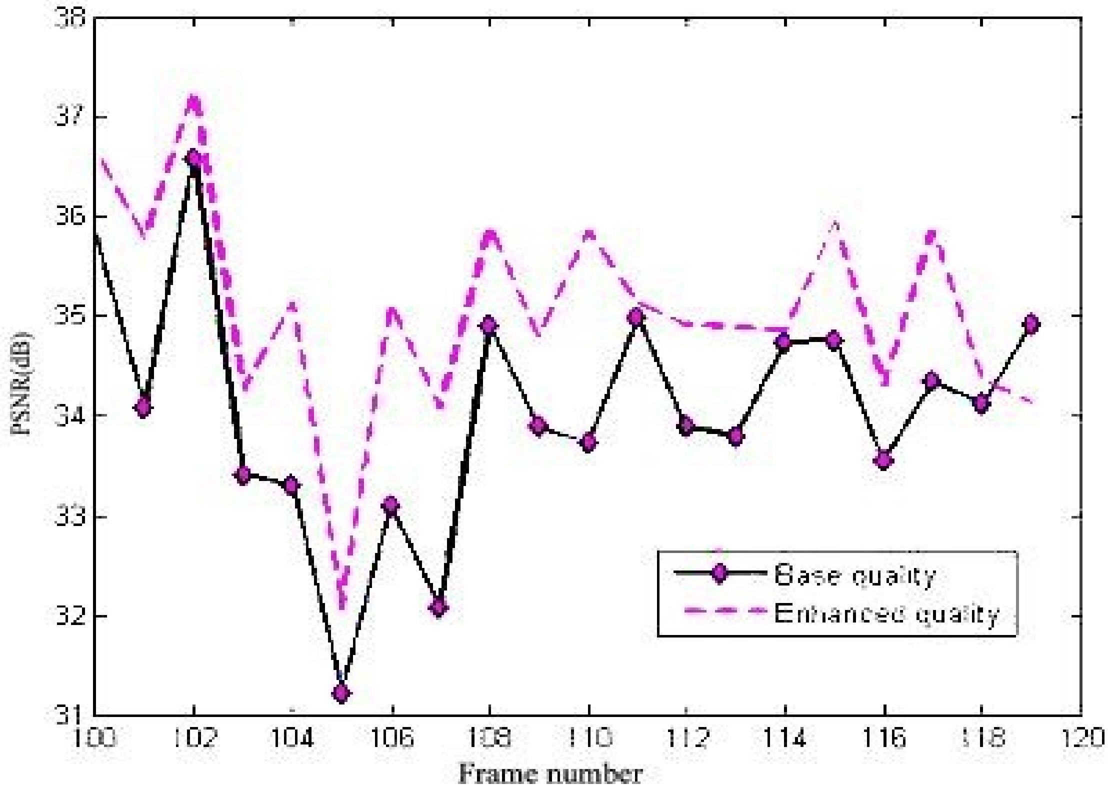

Fig. 4.2.2 PSNR plot for ‘Rhinos.avi’ with improvised motion estimation

4.3 Implementing temporal filtering in various search method

For cutting down the noise existing in a video sequence, temporal filtering is a most commonly used tool. As in a video we are dealing with both spatial and temporal domain this can be referred as spatiotemporal filtering broadly. Temporal filtering can also be considered as a special case of spatiotemporal filtering [16]. In a video sequence plurality of frames in abundance in a small interval make them similar up to a great extent. This correlation benefits the spatiotemporal filtering as it enables the later one for circumventing noise up to a considerable extent. In some of the filtering scheme, the filter assumes that the video sequence is a stationary one. But unfortunately it results in a filter that can be a reason behind the distortion in a certain region of the sequence. Because this assumption won’t work at spatial and temporal edges. So motion compensation can’t be avoided while doing temporal filtering

.in this work a temporal filtering scheme has been along with motion compensation. The video sequence taken here is 'rhinos.avi'. Temporal filtering has been implemented on the video with motion compensation. Also three of the popular search methods i.e. three step, four step and diamond search has been implemented on the sequence taken to analyze the performance. The following figures fig. 4.3.1 and 4.3.2 show the SNR and PSNR value for base and enhancement layer of the test video sequence after the execution of the temporal filtering scheme. Number of frames taken into consideration is 50.

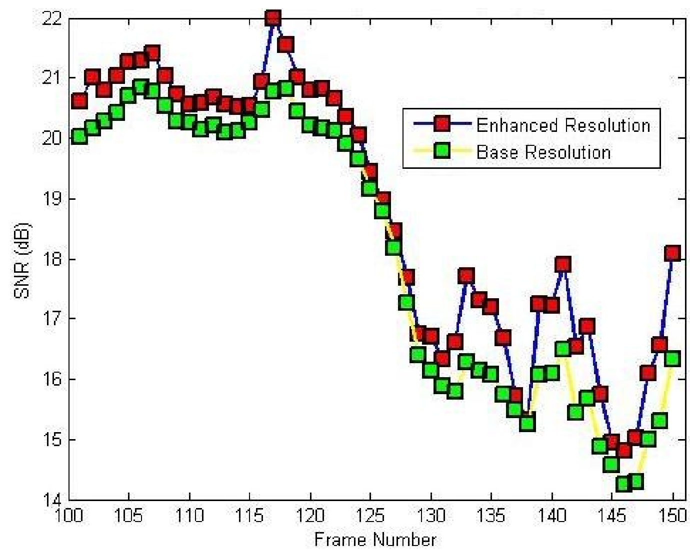

Fig. 4.3.1 SNR plot for 'rhinos.avi' using temporal filtering

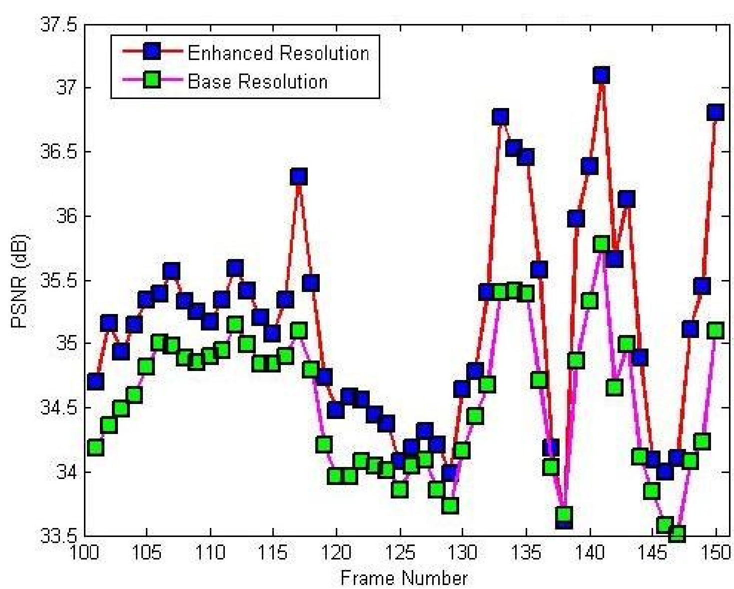

Fig. 4.3.2 PSNR plot for 'rhinos.avi' using temporal filtering

4.4 Results and Analysis

The insights of the previous sections have been summed up here and analysis have been carried out .The comparison has been studied between two videos i.e. ‘rhinos.avi’ and. The SNR,PSNR, SSIM have been evaluated for the video sequence along with computational time for various search method i.e. full search, three step, four step and diamond search. The table 4.4.1 shows the above said for ‘rhinos’ video and 4.4.3 shows about ‘news.avi’. similarly 4.4.2 and 4.4.4 table show the result for the video encoded using the proposed method. The improvised method has been executed on the video sequence. Full search is the method which gives the most accurate output but it is computationally intensive. However next to it three step, four step, diamond search gives gradual reduction in evaluation time respectively. The improved motion estimation technique works with least computation time but with distortion in output. Hence to enhance the output, filtering has been implemented as filtering is a well known method for removal of noise and also holds perfectly good along with motion compensation when we want to remove severe annoying artifacts arose due to drift. All these procedures which is used for revamping the efficiency of compression is intended to cope up with scalable video coding to make it more reliable for the real time multimedia application. Again the mode of scalability chosen is spatial. The encoding has been done spatial domain for certain number of consecutive frames taken from the video sequence. Moreover the collaboration of temporal filtering with I proved motion estimation method has been found as the efficient one for compression in spatial domain of SVC.

TABLE 4.4.1

Comparison of SNR, PSNR and SSIM Along and time of computation for video sequence 'Rhinos.avi'

Search method	**SNR**		**PSNR**		**SSIM**	**Computation Time**
	BL	***EL***	***BL***	***EL***		
Full search	19.66	20.70	34.01	35.01	0.9154	80.2535
Three step Search	18.24	19.56	33.64	33.45	0.7568	75.6528
Four step Search	18.62	18.62	31.65	34.62	0.6845	69.2451
Diamond Search	16.95	17.62	33.45	31.62	0.6328	57.3568

TABLE 4.4.2

implementing proposed methods on 'Rhinos.avi'						
Method Used	**SNR**		**PSNR**		**SSIM**	**Computation Time**
	BL	***EL***	***BL***	***EL***		
Conventional method with filtering	20.66	20.70	35.65	36.01	0.9354	89.2568
Improved ME	19.05	19.21	34.01	35.01	0.8232	62.2536
Improved ME with filtering	20.56	19.95	34.65	35.68	0.8956	99.2564

TABLE 4.4.3

Comparison of SNR, PSNR and SSIM Along and time of computation for video sequence 'news.avi'

Search method	**SNR**		**PSNR**		**SSIM**	**Computation Time**
	BL	***EL***	***BL***	***EL***		
Full search	20.33	20.71	39.94	40.32	0.8925	56.3592
Three step Search	19.70	19.21	38.64	39.45	0.7568	45.2536
Four step Search	19.62	18.95	37.65	38.35	0.7265	40.3568
Diamond Search	18.56	18.24	37.45	38.52	0.6501	35.6529

TABLE 4.4.4

Encoding 'news.avi' with proposed method						
Method Used	**SNR**		**PSNR**		**SSIM**	**Computation Time**
	BL	***EL***	***BL***	***EL***		
Conventional method with filtering	21.36	21.76	40.65	41.36	0.8825	75.2568
Improved ME	19.96	19.95	38.35	38.69	0.7235	50.6254
Improved ME with filtering	20.62	21.65	39.65	40.35	0.8023	65.2536

CHAPTER 5

CONCLUSION AND FUTURE SCOPE

5.1 Conclusion:

As compared to traditional single layer video encoding techniques, scalable encoding of the bit streams results in remarkable loss reduction from the prospective of coding efficiency. In this thesis the test video sequences have undergone spatial scalable encoding. Then extending further revamping the motion estimation is carried out. Motion estimation is a vital but most time consuming part in the process of video encoding. Hence, by reducing the number of search points. In this work we have followed a scheme to improvise the estimation of motion and consequently reducing computational complexity. To improve the output further-more also filtering has been carried in temporal domain. Different search methods also have been implemented for the video sequences taken and computational time has been calculated for each along with the developed one. The quality of the outputs is measured on the basis of parameters SNR, PSNR, SSIM. The comparison has been carried out between the developed method and conventional method. The insights of the whole work have been collected in the fourth section which tends to a reliable implication of scalable video coding in the field of multimedia applications such as TV broadcasting, mobile videos, streaming videos over internet etc. with beneficial traits like higher coding gain, graceful degradation when dealing with lossy environment, hierarchical prediction structure etc.

5.1 Future scope:

The work done in this thesis can be extended and enhanced further as follows:

- Concept of region-of-interest (ROI) can be implemented in proposed method.
- Super resolution can be incorporated in the frames to improve the quality
- The proposed method can better work in the combined scalability environment.

- Better technique for motion estimation should be developed to reduce the time needed for evaluation i.e. reduction in computational complexity.
- Implementing this with fine granular scalability will give a more refined output.

References:

[1] M.Schobinger, B.Zehner, F.Matthiesen, U.totzek, J. hartl, U.Reimann,"*A Single-Chip Adaptive DPCM Video Codec*" IEEE journal of solid-state circuits, Vol. 24,pp. 481–484 ,august 1989.

[2] S.J. Choi, J.W. Woods,"*Motion compensated 3-D sub-band coding of video*"IEEE Trans. Image Proc., vol. 8, pp. 155 - 167, Feb 1999.

[3] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra,"*Overview of the H.264/AVC video coding standard*," IEEE Transactions on Circuitsand Systems for Video Technology,vol. 13, pp. 560-576, July 2003.

[4] G. J. Sullivan and T. Wiegand, "*Video compression – from concepts to the H.264/AVC standard,*" Proceedings of IEEE, vol. 93, pp. 18-31, Jan 2005.

[5] H.Schwarz, D.Marpe, T.Wiegand, "*Overview of the Scalable Video Coding Extension of the H.264/AVC Standard*" IEEE Transactions on Circuits and Systems for video Technology,vol.17, September 2007.

[6] G.Dan, and O.Hadar. "*Efficient region-of-interest scalable video coding with adaptive bit-rate control.*" Advances in Multimedia 2013,vol. 2013,Aug 2013.

[7] A. Munteanu, et al., "*Control of the distortion variation in video coding systems based on motion compensated temporal filtering,"*Proc. IEEE Int. Conf. Image Process. (ICIP 2003), Barcelona, vol.2,pp. 11-61,Sept. 2003.

[8] B. Ben Fradj, A.O. Zaid.“*Scalable video coding using motion-compensated temporal filtering*” In Visual Information Processing (EUVIP), 2011 3rd European Workshop, pp.50 –55, july 2011.

[9] B.Girod, S. Han, “*Optimum update for motion-compensated lifting*,” IEEE Sig. Process. Lett., vol. 12, pp. 150-153, Feb. 2005

[10] A. Kundu, “*Modified Block Matching Algorithm for Fast Block Motion Estimation*,” International Conference on Signal and Image Processing(ICSIP), 2010.

[11] J.R. Ohm, "*Advances in Scalable Video Coding*", Proceedings of the IEEE, vol. 93, pp.42-56, Jan. 2005.

[12] Y. Chen, YKWang, M Gabbouj “*Regionally adaptive filtering for asymmetric stereoscopic video coding*” IEEE international symposium on circuits and systems, vol.2, pp.2585-2588,May 2009

[13] B. Lee ,M. Kim. “*An efficient inter-prediction mode decision method for temporal scalability coding with hierarchical b-picture structure broadcasting*” IEEE Transcations on broadcasting, vol.58,pp.285-290, feb.2012

[14] K. S. Thyagarajan. Still image and video compression with MATLAB. A John Wiley and Sons, Inc., 2011

[15] Y.Liu, Z.G.Li, Y.C. Soh. *“Rate control of h.264/avc scalable extension.”* IEEE Transactions on Circuits and Systems for Video Technology, vol.18,pp. 116 –121, jan. 2008.

[16] M. Domanski, A. Luczak, and S. Mackowiak. “*Spatio-temporal scalabilityfor mpeg video coding”* IEEE Transactions on Circuits and Systems for Video Technology, vol.10, pp.1088-1093, july 2000