August 2012

Master's Degree Thesis

Advanced H.264/AVC Error Resilient Video Coding over Erroneous Wireless Network

Graduate School of Chosun University

Department of Information and Communication Engineering

Shaikhul Islam Chowdhury

Advanced H.264/AVC Error Resilient Video Coding over Erroneous Wireless Network

무선 에러 네트워크에서 향상된 **H.264/AVC** 에러
리질리언스 비디오 코딩에 관한 연구

August 24, 2012

Graduate School of Chosun University

Department of Information and Communication Engineering

Shaikhul Islam Chowdhury

Advanced H.264/AVC Error Resilient Video Coding over Erroneous Wireless Network

Advisor: Prof. Jae-Young Pyun

This thesis is submitted to Chosun University in partial fulfillment of the requirements for a Master's degree in Engineering

April 2012

Graduate School of Chosun University

Department of Information and Communication Engineering

쵸우두리 샤이쿨 이슬람 석사학위 논문을 인준함

위원장 조선대학교 교수 박종안

위 원 조선대학교 교수 한승조

위 원 조선대학교 교수 변재영

2012년 6월

조선대학교 대학원

Table of Contents

List of Tables

List of Figures

Acronyms

AVC	:	Advanced Video Coding
CABAC	:	Context Adaptive Binary Adaptive Coding
DP	:	Data Partitioning
ER	:	Error Resilience
FMO	:	Flexible Macroblock Ordering
FPS	:	Frame per Second
GOP	:	Group of Pictures
IDR	:	Instantaneous Data Refresh
ISO	:	International Organization for Standardization
ITU	:	International Telecommunication Union
ITU-T	:	Telecommunication Standardization Sector
JVT	:	Joint Video Team
MB	:	Macroblock
MBAmap	:	Macroblock Slice Allocation Map
MDC	:	Multiple Description Coding
NALU	:	NAL Unit
PSNR	:	Peak Signal-to-Noise Ratio
QCIF	:	Quarter Common Intermediate Format
PS	:	Parameter Set
PPS	:	Picture Parameter Set
RPS	:	Reference Picture Selection
RS	:	Redundant Slices
RTP	:	Real Time Transmission Protocol
SAD	:	Sum of Absolute Differences
SPS	:	Sequence Parameter Set

VCEG : Video Coding Experts Group
VLC : Variable Length Coding

요약

무선 에러 네트워크에서 향상된 **H.264/AVC** 에러 리질리언스 비디오 코딩에 관한 연구

Shaikhul Islam Chowdhury
지도 교수: 변재영 교수, **Ph. D**
정보통신공학과
조선대학교 대학원

화상 전화, 화상 회의, 영상 스트리밍과 같은 실시간 멀티미디어 통신의 수요 증가에 따라 영상 압축과 전송 강인성은 모든 영상 코덱 표준에서 중요한 부분을 차지하고 있다. 과거의 인터넷은 정적인 화면 뿐이었지만 오늘날의 인터넷은 영상 스트리밍이 가미된 동적인 화면들이 대부분이다. 이동 통신도 마찬가지의 현상을 보인다. 가령, 과거의 휴대 전화는 음성 통화 용으로만 사용되었지만, 오늘날은 데이터 서비스가 보다 중요한 부분을 차지한다. 멀티미디어 통신의 수요가 오늘날의 통신에 활력을 불어넣고 있다. 모두가 월드컵, 총선과 같은 그들이 선호하는 방송 혹은 스트리밍 프로그램 상에서 선명한 고화질 영상을 제공받기를 원한다. 그러나 실시간 비디오 서비스는 유독 이동통신 시스템 상에서는 무선 채널의 품질 차이로 인해 어려운 과제로 남아있다.

비디오 전송 서비스를 제공하기 위해서는 영상 압축 코덱이 필요하다. 이러한 압축 기법들의 조합이 높은 압축률을 제공하지만, 동시에 인코딩된 비트열 데이터는 에러에 더욱 취약해진다. 이는 에러에 취약한 무선 네트워크 환경에서 프레임 에러가 발생될 경우 왜곡된 프레임의 복원이 어렵기 때문이다. 영상 데이터는 에러에 취약한 채널에서 패킷 손실, 무선 채널에서 발생하는 페이딩 현상 등으로 인해 화질이 현저히 저하된다. 그러므로, 화질 저하를 완화시키기 위해 인코더에는 에러 리질리언스 기법이 필요하다. 사전 방지를 위한 에러 리질리언스 기법은 인코더에서 디코더의 에러 처리를 보다 수월하게 하는 지능적 설계에 기반을 두고 있다.

본 이론에서, 에러에 취약한 네트워크에서의 영상 전송 성능 평가를 위해 여타 표준들과 대체 가능한 에러 리질리언스 기법들이 기술되고 비교되었다. 특히 랜덤 인트라 리프레쉬 기법이 강력한 에러 복원 기법으로 사용될 때 H.264/AVC 에 적용된 다수 참조 프레임 기법이 분석되었다. 그리고 나서 제안된 참조 프레임 선택법에 기반을 둔 인트라 리프레쉬 기법은 랜덤 인트라 리프레쉬와 결합된 다수 참조 프레임 기법의 결점을 극복하기 위해 존재한다.

I. Introduction

A. General Overview

The 3rd generation of mobile systems designed to enable multimedia communication. H.264, the latest video coding standards, provides video compression, error resilience and error concealment tools. In wireless communication, there is always some data loss due to transmission errors. If there is no error, there is no need to add any redundancy to make encoded bit-stream robust. But due to transmission error, encoder has to use some error resiliency techniques so that if any error happens in transmission, decoder could mitigate the distortion. Video compression and error resilience are oxymoron, so encoder needs to use tradeoff between these two. In general, video coding, transmission and decoding process are given in Fig. 1.1.

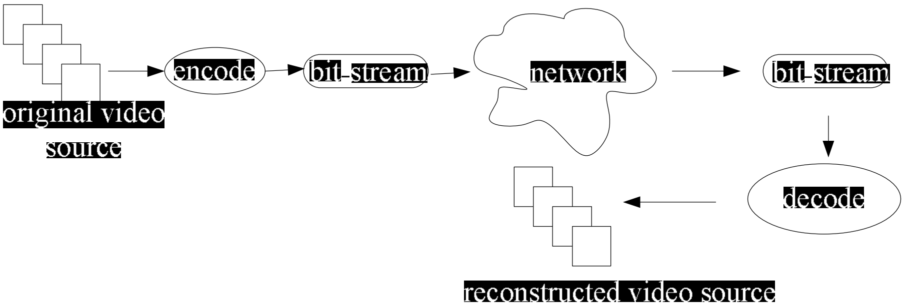


Fig 1.1 General video codec procedure

Many researchers are working on how to improve the performance of video codec. To follow some defined pattern especially in encoded bit stream and decoding, International Telecommunication Union (ITU) and International Organization for Standardization (ISO) provide standardization. Recently, they joined under the group Joint Video Team (JVT) and created H.264 video codec which is the most popular in recent years.

B. Motivation and Thesis Objective

In H.264, to mitigate error, several methods are used, classified mainly in three groups, at encoder, at decoder and feedback based both at encoder and decoder. In addition to these methods, there are also error resilience methods, which mainly used to stop transmission errors. Commonly used error resilient methods are intra placement, picture segmentation, reference picture selection (RPS), data partitioning (DP), redundant slices (RS), parameter sets (PS), and flexible macroblock ordering (FMO). All these types of methods can be used independently or together. However, performance of different error resilience methods varies in erroneous network. Also, the performance is affected by the frame sequence structure, rapid scene change, or fast object moving.

In this thesis, different standard and alternative error resilient methods are investigated. A problem using multiple reference frames with intra refresh in erroneous network is explained and solution is proposed to solve the problem.

C. Thesis Contribution

In this thesis, a new method is proposed to use intra refresh with multiple reference frames properly. By using proposed method, positive gain in video quality can be achieved in erroneous network. This thesis also surveys different error resilience methods, how to use error models in video transmission. The main contribution are given as follows

New method: A new method is designed to select intra-coded MBs in reference selection in motion estimation.

Simulation: The new method is implemented in H.264/AVC reference software, JM 17.2 encoder, and to test the performance of the proposed method and compare with standard method.

D. Thesis Organization

The remaining of this thesis is organized as follows. In chapter II, principles of H.264 video codec is described. Here main focus was given to H.264 coding structure, prediction methods, rate distortion optimization, profiles and levels, and error propagation and a brief of H.264 reference software version 17.2. In chapter III, standard and alternative error resilience methods are described. Among the standard error resilience methods, intra placement, picture segmentation, reference picture selection, redundant pictures, multiple description coding are described and them some combined error resilience methods are described. In chapter IV, a problem is explained in using multiple reference frames with random intra refresh and then to solve the problem a new method is proposed. In chapter V, Implementation of the proposed method is described. Simulation environment, simulation parameters are described and then performance comparison with standard error resilience methods and discussion are given and finally in chapter VI, conclusion is given.

II. Background

A. Overview of H.264

In this chapter an overview of H.264 is given covering core topics like what H.264 stands for, its basic principles, coding structure, prediction techniques, profiles and levels. Then how error propagates in erroneous network is described.

1. What is H.264?

H.264/ AVC is a recently completed video compression standard jointly developed by the ITU-T VCEG and the ISO/IEC MPEG standards committees. The standard promises much higher compression than that possible with earlier standards. It allows coding of non-interlaced and interlaced video very efficiently, and even at high bit rates provides more acceptable visual quality than earlier standards. Further, the standard supports flexibilities in coding as well as organization of coded data that can increase resilience to errors or losses. As might be expected, the increase in coding efficiency and coding flexibility comes at the expense of an increase in complexity with respect to earlier standards [1].

H.264 Advanced Video Coding is an industry standard for video coding, but it is also a popular format for coded video, a set of tools for video compression and a stage in a continuously evolving digital video communication landscape. H.264 is a method and format for video compression, the process of converting digital video into a format that takes up less capacity when it is stored or transmitted.

An encoder converts video into a compressed format and a decoder reconstructs compressed video back into an uncompressed format as shown in Fig 2.1. In a typical application of H.264 such as remote surveillance, video from a camera is encoded or compressed using H.264 to produce an H.264 bit stream. This is sent across a network to a decoder which reconstructs a version of the source video.

H.264/AVC standard defines a format or syntax for compressed video and a method for decoding this syntax to produce a displayable video sequence. The standard document does not actually specify how to encode digital video – this is left to the manufacturer of a video encoder – but in practice the encoder is likely to mirror the steps of the decoding process. H.264/AVC describes a set of tools or methods for video compression. The standard specifies how video coded with these tools should be represented and decoded [2].

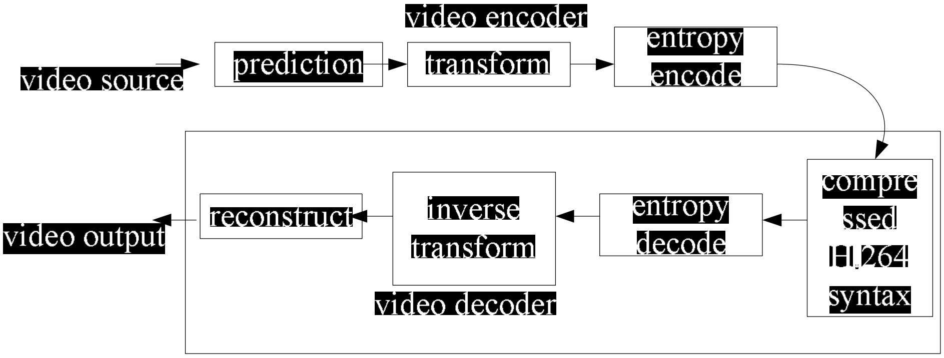


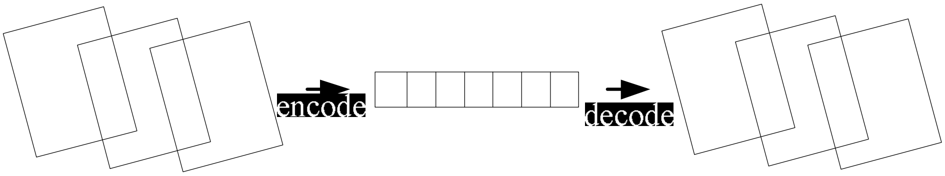
Fig 2.1 Scope of H.264 standard

2. Principles of H.264

An H.264 video encoder carries out prediction, transforming and encoding processes (Fig 2.1) to produce a compressed H.264 bit stream. An H.264 video decoder carries out the complementary processes of decoding, inverse transform and reconstruction to produce a decoded video sequence. This compressed bit stream is stored or transmitted and can be decoded to reconstruct the video sequence. The decoded version is, in general, not identical to the original sequence because H.264 is a lossy compression format, i.e. some picture quality is lost during compression.

A frame or field to be coded is processed by an H.264-compatible video encoder. As well as coding and sending the frame as part of the coded bit stream or coded file, the encoder reconstructs the frame, i.e. creates a copy of the decoded frame A, which will eventually be produced by the decoder. This reconstructed copy may be stored in a coded picture buffer, CPB, and used during the encoding of further frames. The decoder receives the coded bit stream and decodes frame A, for display or further processing. At the same time, the decoder may store a copy of frame A, in a decoded picture buffer, DPB, to be used during the decoding of further frames.

Data is processed in units of a MB corresponding to 16 × 16 displayed pixels. In the encoder, a prediction MB is generated and subtracted from the current MB to form a residual MB; this is transformed, quantized and encoded. In parallel, the quantized data are re-scaled and inverse transformed and added to the prediction MB to reconstruct a coded version of the frame which is stored for later predictions. In the decoder, a MB is decoded, re-scaled and inverse transformed to form a decoded residual MB. The decoder generates the same prediction that was created at the encoder and adds this to the residual to produce a decoded MB.



(a) Source frames (b) Encoded bitstream (c) Reconstructed frames

Fig 2.2 Video encoding and decoding process

3. Coding Structure of H.264

The basic coding structure of this standard is similar to that of earlier standards and is commonly referred to as motion-compensated—transforms coding structure. Coding of video is performed picture by picture. Each picture to be coded is first partitioned into a number of slices (it is possible to have one slice per picture also). Slices are individual coding units in this standard as compared to earlier standards as each slice is coded independently.

As in earlier standards, a slice consists of a sequence of MBs with each MB consisting of *16x16* luminance (Y) and associated two chrominance (C_b and C_r) components. Each MB's *16x16* luminance is partitioned into *16x16, 16x8, 8x16*, and *8x8*, and further, each *8x8* luminance can be sub-partitioned into *8x8, 8x4, 4x8 and 4x4*. The *4x4* sub-MB partition is called a block as shown in Fig. 2.3.

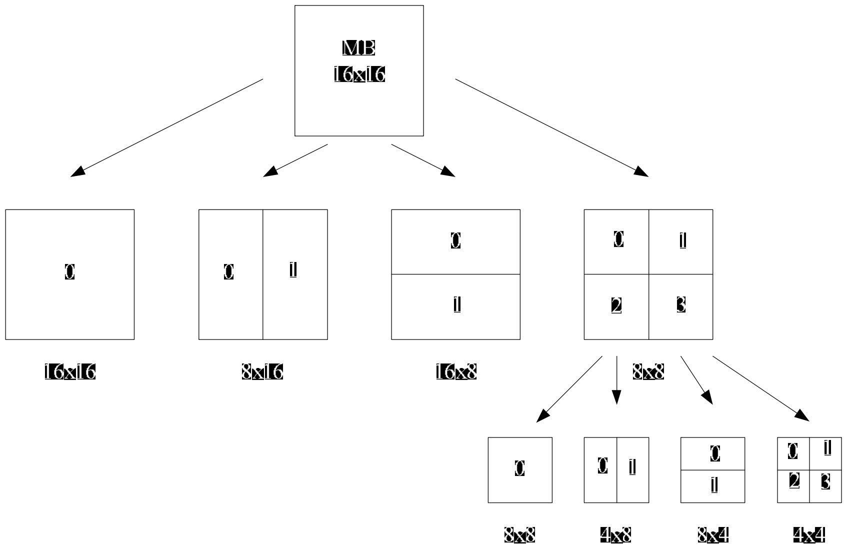


Fig 2.3 MB partitions and sub-MB partitions

Slices are individually coded and are the coding units, while pictures plus associated data can be considered as being the access units. There are three basic slices types: *I—(Intra)*, *P—(Predictive)*, and *B—(Bi-predictive)* slices. This is basically a nomenclature as well as functionality extension of the *I*-, *P*-, and *B*-picture concept of earlier standards. In H.264/MPEG-4 AVC standard, *I*-slice MBs are compressed without using any motion prediction (also true of all earlier standards as well) from the slices in other pictures. A special type of picture containing *I*-slices only called instantaneous decoder refresh (IDR) picture is defined such that any picture following an IDR picture does not use pictures prior to IDR picture as references for motion prediction. Thus all following coded pictures in decoding order can be decoded after decoding an IDR without the need of reference to any decoded picture prior to the IDR picture. IDR pictures can be used for random access or as entry points in a coded sequence. P-slices consist of MBs that can be compressed by using motion prediction, but P-slices can also have intra MBs.

MBs of a P-slice when using motion prediction must use one prediction only (uni-prediction). Unlike previous standards, the pixels used as reference for motion compensation can either be in past or in future in the display order. Also, both I- and P-slices may or may not be marked as used for reference. B-slices also consist of MBs that can be compressed by using motion prediction and like P-slices can also have intra MBs. MBs of a B-slice when using motion prediction can use two predictions (bi-prediction).

4. Prediction technique

The encoder forms a prediction of the current MB based on previously-coded data, either from the current frame using intra prediction or from other frames that have already been coded and transmitted using inter prediction as shown in Fig 2.4, 2.5 and 2.6. The encoder subtracts the prediction from the current MB to form a residual.

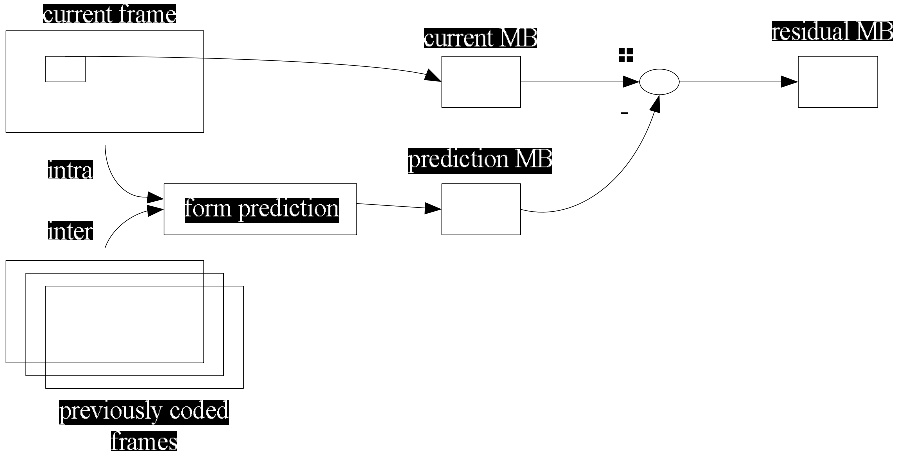


Fig 2.4 Inter/Intra prediction flow diagram

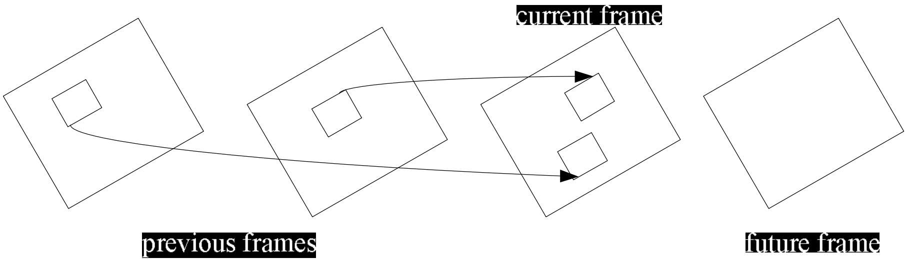


Fig 2.5 Inter prediction

The prediction methods supported by H.264 are more flexible than those in previous standards, enabling accurate predictions and hence efficient video compression. Intra prediction uses 16 × 16 and 4 × 4 block sizes to predict the MB from surrounding, previously coded pixels within the same frame. The values of the previously-coded neighboring pixels are extrapolated to form a prediction of the current MB.

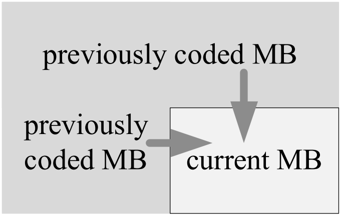


Fig 2.6 Intra prediction

5. Rate Distortion Optimization in H.264

The H.264/AVC video coding standard owes its major performance gains, relatively to previous standards, to the many different intra and inter MB coding modes supported by the video coding syntax. Although not all modes are allowed in every H.264/AVC profile, even for the simplest profiles, such as the baseline profile, the encoder has a plethora of possibilities to encode each MB, which makes it difficult to accomplish optimal MB coding mode decisions with low complexity. Besides the MB coding mode decision, for motion compensated inter coded MBs, finding the optimal motion vectors and MBs partitions is also not a straightforward task. In this context, rate distortion optimization (RDO) becomes a powerful tool, allowing the encoder to optimally select the best MB coding modes and motion vectors if applicable.

$$J_{mode} = D\left(MODE, QP\right) + \lambda_{MODE} \times R\left(MODE, QP\right), \qquad (2.1)$$

(2.2)

$$J_{MOTION} = D\left(mv\left(REF\right)\right) + \lambda_{MOTION} \times R\left(mv\left(REF\right)\right), \qquad (2.3)$$

In H.264/AVC JM reference software, the best MB mode decision is accomplished through RDO technique where the best MB mode is selected by minimizing Lagrangian cost function [4] shown in Eq. (2.1). Where MODE is one the allowable MB coding modes, QP is the quantization parameter, and D and R are respectively, the distortion (between the original and the reconstructed MB) and the number of bits that will be achieved by applying the corresponding MODE and QP. The value of λ_{MODE} as shown in Eq. (2.2), for inter and intra prediction, is defined in [5].

Motion estimation can also be accomplished through the same framework. In this case, the best motion vector and reference frame can be selected by minimizing the Lagrangian cost function shown in Eq. (2.3).

When deciding the best MB coding mode, notably between inter and intra coding modes, the RDO framework simply selects the mode which has lower RD cost. This RDO framework does not take into account other issues, besides rate and distortion optimization, such as the robustness of the bitstream in error-prone environments. Therefore some MBs are simple inter coded because of their best RD cost.

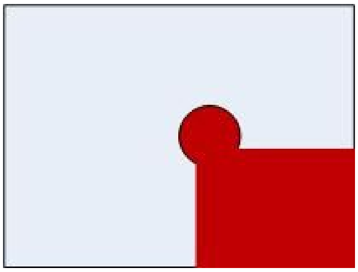
6. H.264 Profiles and Levels

H.264 supports a portfolio of ‘tools’. These are algorithms or processes for coding and decoding video and related operations. These include tools essential to any implementation of H.264 such as the basic 4 × 4 transform, optional or interchangeable tools such as CABAC or CAVLC entropy coding, tools for specific applications or scenarios such as the SI/SP tools which can support streaming video and tools not directly related to video coding such as VUI and SEI messages.

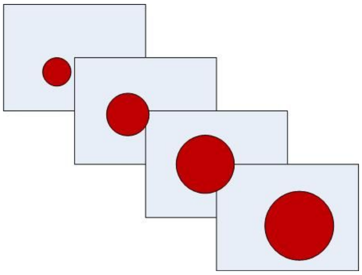
B. Error Propagation in Erroneous Network

The compression mechanism introduced in the previous section is capable of reducing the data rate more than 100 times while keeping the user-perceived quality almost unchanged. It has, however, some drawbacks to the robustness of such an encoded video stream against transmission errors. An error in a single bit of stream may cause considerable distortion due to the error propagation.

There are three possible sources of error propagation in an H.264/AVC encoded video stream: entropy code, spatial prediction and temporal prediction. Entropy codes assign code words to symbols to match codeword lengths with probabilities of the symbols – entropy codes are Variable Length Codes (VLCs). Hence, an error in a codeword may impact following code words as well, if the codeword boundaries are determined incorrectly. This is also a reason for discarding the entire packet if an error is detected.



(a) Error propagation in intra prediction (I frame)



(b) Error propagation in inter prediction (P frame)

Fig 2.7 Error propagation example caused in erroneous network

Spatially (intra) encoded MBs require the spatially neighboring reference blocks for decoding shown in Fig 2.5 (a). If these are erroneous, the MB whose reconstruction is based on them will be distorted, too. In practice, however, this problem does not occur as the entire packet is discarded if an error is detected. However, spatial error propagation becomes significant for the techniques utilizing information from damaged packets. Temporal (inter) prediction causes error propagation between the frames of the same video sequence shown in Fig 2.5 (b). If an error occurs in a frame, all following frames that use such a frame as a reference for the motion compensation will be reconstructed erroneously. The information in the video sequence is refreshed in each inter frame by the new information contained in the transmitted residuals; thus the sequences containing a higher amount of movement recover faster. Another source of refreshment is I frames and the intra-coded MBs inserted (randomly) in inter-predicted slices.

C. Performance Matrix

The performance of error resilience methods can be evaluated by measuring the end-user distortion in the presence of transmission errors with given characteristics and by analyzing the cost determined by increasing the rate and/or complexity. In this section, metrics are presented that are used throughout this book for performance evaluation. For image distortion, the MSE is most commonly quoted in terms of the equivalent reciprocal measured Peak Signal to Noise Ratio (PSNR).

D. Reference Software

The JVT, the group responsible for developing and maintaining H.264/AVC, publishes a reference implementation of the standard as a C software program, known as the Joint Model (JM). Software flow diagram of JM model is shown in Fig 2.8. The JM software consists of an encoder (*lencod*) that encodes a source video file into a coded H.264 file and a decoder (*ldecod*) that decodes an H.264 file into a decoded video file. The encoder and decoder are each controlled by a parameter file with default names *encoder.cfg* and *decoder.cfg*. The encoder creates a reconstructed video file, a copy of the decoded video file, and can optionally generate a trace file that records every syntax element of the coded sequence. To create packet loss scenario within the bit-stream, JM provides RTP_*loss* module that randomly drops some packets from the bit-stream. These three modules are command line programs; an example is given in Fig. 2.9.

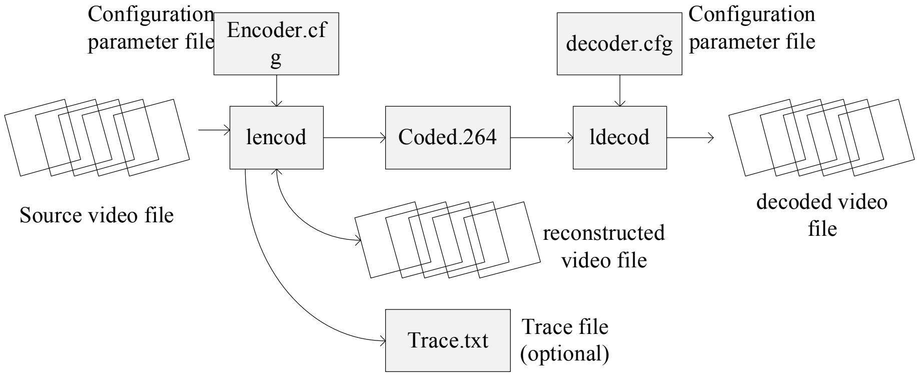


Fig 2.8: JM software flow diagram

Table 2.1 Available modules in JM software

Module	Command	Description
Encoder	lencod.exe encoder.cfg	To encode a video sequence
Decoder	decode.exe decoder.cfg	To decode a video sequence
RTP loss	rtp_loss original.264 erroneous.264 5	To create erroneous bitstream

III. Error Resilience Techniques

To reduce the distortion resulting from the errors and their propagation, several means were added to the H.264 standard. They are summarized nicely in Ostermann *et al.* (2004) and Wiegand *et al.* (2003). In Stockhammer and Hannuksela (2005), the benefits of the H.264/AVC error resilience features in the context of wireless transmission are emphasized. Besides the traditional error resilience schemes that are inherited from previous standards like H.263 or MPEG-4 to H.264, in this chapter some prominent error resilience methods are discussed which are more likely applicable for erroneous network.

Table 3.1 Example of random intra refresh order for QCIF sequence, 11x9 matrix with MB index number

95	94	24	60	97	22	90	21	61	77	43
9	65	28	41	50	11	23	18	82	26	73
87	36	71	15	29	52	70	10	64	44	45
63	8	66	88	35	32	33	39	0	80	72
78	25	93	67	68	85	48	19	89	1	7
53	75	49	46	54	42	13	4	81	30	40
57	3	79	84	86	55	98	56	58	16	47
91	69	92	59	62	5	51	83	38	76	27
34	96	17	37	31	6	20	14	12	2	74

A. Intra Placement

H.264 allows intra MB prediction in inter and intra frame. The best MB mode decision and motion estimation (ME) are accomplished through the RDO technique by minimizing Lagrangian cost function [5]. Allowable coding modes are SKIP, INTER 16X16, INTER 16X8, INTER 8X16, INTER 8X8, INTER 4X4, INTRA 16X16, INTRA 4x4. According to RDO, it always selects lowest RD cost. it does not take into account any other issues such as the robustness of the bit stream in erroneous network, therefore some MBs are simply inter coded because their best Inter mode RD cost is slightly lower than that of the best intra mode RD cost.

To select the intra cases, Random Intra Refresh (RIR) is introduced. It defines how many MBs should be intra-coded forcefully. For an example, if it is defined that for each frame 3 MBs should be intra-coded regardless of the RD cost. In that case, 3 MBs will be intra-coded.

To determine the MBs for intra coding, a randomly generated order is used, shown in Table 1. Although it is called RIR, it is not really a pure random refresh technique. This technique is basically a cyclic refresh technique for which the refresh order is not simply the raster scan order. The refresh order is randomly defined once before encoding and then it follows the order cyclically. The main advantage of using RIR is to stop error propagation. At the same time, the more intra refresh will be forced, the overall encoding performance decreases in terms of PSNR values. This is because RDO technique always tries to avoid intra-coded MBs. An example of using RIR to encode 25 frames of akiyo QCIF sequence and its corresponding PSNR values is given in Table 3.2.

Table 3.2 PSNR values of akiyo QCIF sequence

RIR	PSNR (dB)
×	33.707
o	33.353

In feedback based system, adaptive intra refresh [6] can be used according to the channel state. If a channel is stable enough it does not need to go for intra-coded MBs, but if the channel state is erroneous in that case intra-coded MBs can be used and thus it could stop the error propagation effectively.

In [5] RDO driven intra refresh is introduced. In this paper the authors use a user defined threshold value to determine if inter or intra mode decision should be used.

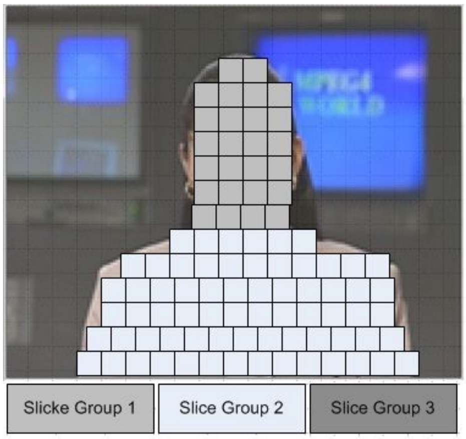
B. Picture Segmentation

A frame can be divided into one or more slice groups. For simplicity the standard limits of slice groups is eight. Each slice consists of number of MBs but one MB cannot be in multiple slices. MBs are assigned to slices in raster scan order. In FMO. MBs are assigned in different scan order.

Fig 3.1 2 slice groups in a QCIF sequence

To do so, each MB is statically assigned to a slice group using MB allocation map (MBAMap). Within a slice group, MBs are coded using normal scan order. FMO is available only in Baseline and Extended profiles but not in the main profile. There are 5 predefined FMO like type-0 row interleaving, type-1 dispersed or checker board like pattern, type-2 foreground with left over, type-3 box out, type-4 raster scan, type-5 wipe and 1 custom FMO which is left to video coder to decide the ordering. Among the predefined FMOs, checkerboard pattern performs best in erroneous network [7].

(a) Original akiyo QCIF frame



(b) Slice group 1 for face, slice group 2 for body, and slice group 3 for background

Fig. 3.3 Adaptive FMO

In type-6 explicit FMO new ordering can be introduced or one or more predefined ordering can be combined. In [8] the authors combined checkerboard like pattern and the spiral FMO. An example for customized FMO is shown in Fig 3.2. Here row interleaving is combined with a zig-zag scan order together. Thus, MBs in the first and third rows belong to the slice group 0 and MBs in second and fourth rows belong to the slice group 1. After zig-zag scan order its MBAMap would be as follows.

$MBAMap = \{ 0\ 0\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ 1\ 1\ 0\ 1\ 0\ 1\ 1 \}$

In [9] adaptive MB classification is used. The authors classify MBs into slice groups dynamically according to the frame content. For an example, if the background is constant in some consecutive frames, three slice groups can be used as shown in Fig. 3.3, i.e., one slice group for background part and other slice groups for foreground parts.

C. Reference Picture Selection

Reference picture selection of per picture, per slice or per MB can be used as an error resilience method. Encoding efficiency increases by using more reference frames as in that case encoder can select best RD cost from more candidates.

Quality of video by encoding a qcif sequence with single and multiple reference frames is shown in Table 3.3.

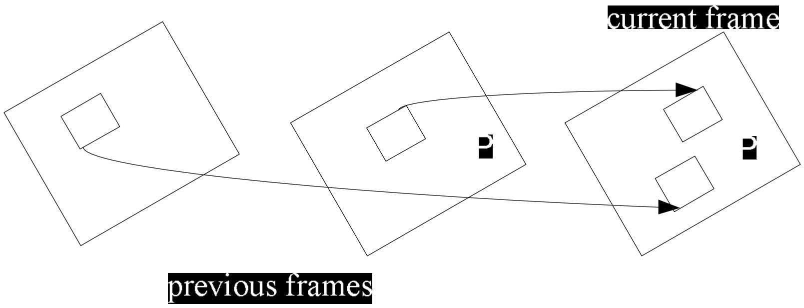


Fig. 3.4 Reference picture selection for an inter MB coding

In Fig.3.4, typical reference block selection for a MB in P frame is shown. Here, best matched block is found in previous reference I frame. The more number of reference frames encoder can use, the more choices it can use to select best MB.

D. Redundant Pictures/Slices

H.264 supports Redundant Pictures (RP). According to the standard each picture may be associated with one or more RPs, which a decoder can reconstruct in case a primary picture or parts of it is missing. H.264 does not

Table 3.3 PSNR values of akiyo QCIF sequence

No of Reference Frames	PSNR (dB)
1	33.707
7	33.809

(a) Face position marked with a green bordered box

(b) Redundant slice for face block

Fig. 3.5 Dynamic allocation of ROI

define how to generate redundant pictures as long as a decoded redundant picture is visually similar in appearance to the corresponding decoded primary picture.

Redundant Slices (RS) allow the encoder to include an additional representation of a slice with the original, primary coded slice in the same bit stream. Whenever any of the primary coded slices cannot be correctly decoded, the decoder can replace the primary slice with redundant representations. In [8] dynamic RSs are used. In the paper Region of Interest (ROI) part is identified from a frame and for that part RS is used. Fig. 3.5 shows a frame when foreground with left over FMO type 2 is implemented together with redundant slice feature.

E. Multiple Descriptions Coding

Multiple descriptions coding (MDC) represents a signal in more than one description in a way that high quality is achieved when all descriptions are reconstructed in combined. The most common MDC model refers to two descriptions with rates R1 and R2 that are sent with two erroneous channels. By receiving one description creates side distortion and by receiving both descriptions creates central distortion. By using redundant pictures lost descriptions can be recovered based on correctly received ones. In [10], MDC is used with redundant pictures. They use odd and even frames in different descriptors. When encoding, each primary picture in the odd/even description is predicted only from other pictures of the same description. RPs is placed in such a way that they can substitute a possibly lost primary picture. RPs is predicted from the previous frame in the input sequence. RPs is coded as P frames. It gives better performance in erroneous network. Figure 6 shows two descriptions of IPP... scenario.

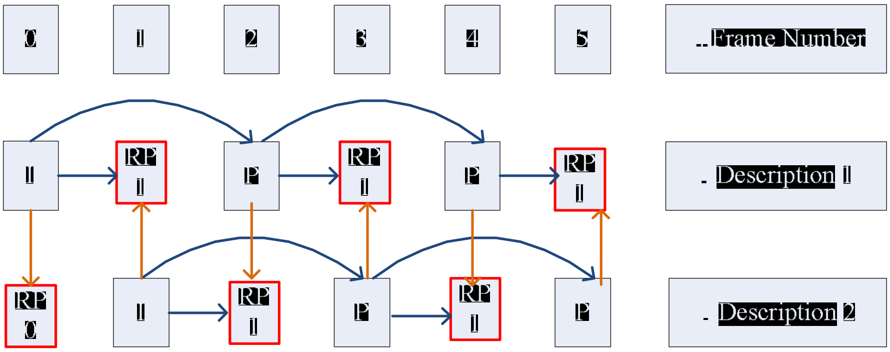


Fig. 3.6 Example of MDC with RSs

F. Joint Error Resilience Methods

To mitigate further errors over an erroneous network, different error resilience tools can be combined. Ex. the combination of RIR and FMO [6] and FMO and Redundant slices [8].

In [6] by using the feedback channel state, they define number of forcefully intra-coded MBs. In general, how many MBs will be used as intra-coded is defined statically before going to encode the video sequence. So by adjusting number of Intra MBs according to channel state gives better performance in erroneous network.

Table 3.4 PSNR values for akiyo QCIF sequence using FMO and RIR in error free channel

Method	PSNR (dB)
NoFMO + NoRIR	33.844
FMO + NoRIR	33.800
FMO + RIR	33.778

In Table 3.4, encoding performance by using RIR and FMO is shown for akiyo QCIF sequence in error free channel.

In error free channel overall performance decreases but in erroneous channel by using FMO along with RIR its performance will increase, because in case error propagation RIR fights better and for using FMO if a MB get distorted it could be concealed by using neighbor MBs.

IV. Proposed Error Resilience Method

As an error resilient method, typical video codecs including H.264/AVC use intra refresh coding to avoid error propagation on the distorted video sequence on the erroneous network [11]. The refresh order of intra refresh coding is not simply the raster scan order, but is randomly defined once before the start of encoding and afterwards intra refresh proceeds cyclically. In [12], it is shown that multiple reference frame (MRF) coding method combined with the cyclic intra refresh (CIR) reduces overall video quality in erroneous network. This letter proposes a solution to overcome this quality degradation problem.

A. Problem Definition

Moiron et al. [3] described the limitations of using MRF with CIR in an erroneous network, even though increasing the number of reference frames is better than using single reference frames in terms of the pure coding efficiency. In the present study, an experiment was performed on the JM 17.2 H.264/AVC to encode standard QCIF sequences (tennis and stefan) based on a MRF with and without CIR.

In this thesis, an experiment is done with JM 17.2 H.264/AVC encoding standard QCIF sequences (tennis and stefan) based on MRF with and without CIR, and corresponding PSNR values are given in Table 4.1 [13]. The GOP coding structure consisted of an initial intra-coded IDR frame and following inter coded P-frames, i.e., IPPP.., which is suitable for mobile applications.

It is shown that the quality of a H.264 video sequence is improved with MRF under an error-free network, but the video quality deteriorates in the presence of CIR over the erroneous network. This occurs because of a typical RD based reference selection, called the RDRS, on MRF. The RDRS calculates the RD cost of each coding block mode and chooses the optimal coding block as well as its mode with the minimum RD cost for each MB without consideration of the possible block damage on the erroneous network. When more multiple frames are allowed for a reference prediction, RDRS can select non-refreshed inter-coded MBs, as shown in Fig. 4.1 (a). For example, RDRS selects the shaded block in F_{n-1} for a reference of block in frame F_{n+1} based on RD cost strategy, which can cause error propagation into a block in F_{n+1} because the shaded portion in frame F_{n-2} represents the damaged area of the frame due to transmission errors.

If there are errors in the inter-coded MBs, the damage is propagated through the frames before the errors meet the refreshed blocks or slices. Non-refreshed MBs are more likely to be selected in multiple frames when more reference frames are used, causing error propagation.

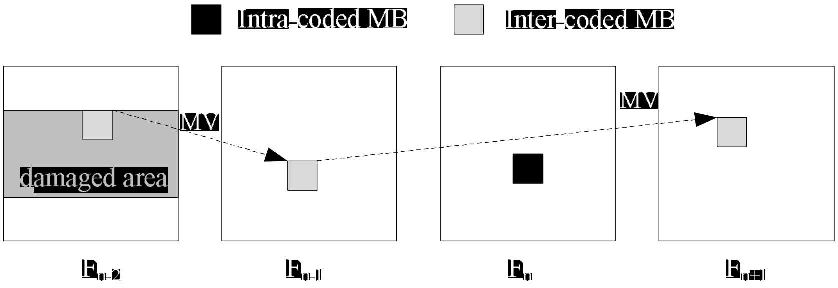
Table 4.1 Average PSNR values of H.264 video sequences encoded with and without CIR (packet loss occurred from 50th frame and PSNR was observed through whole frames)

No. of RF	CIR	PSNR (dB) for Tennis		PSNR (dB) for Stefan	
		0% PLR	5% PLR	0% PLR	5% PLR
1	X	34.35	32.09	34.72	23.28
7	X	34.39	32.05	34.74	23.30

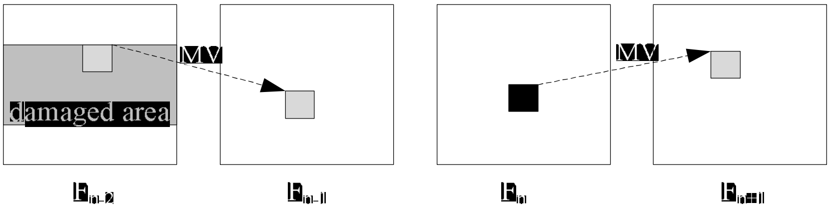
1	○	34.37	32.12	34.73	25.01
7	○	34.43	29.45	34.71	23.97

B. Proposed Method

An intra-refresh based reference selection (IRRS) scheme is proposed for the higher error robustness of MRF-based videos transferred over the erroneous network. The IRRS method attempts to choose a reference block that was intra-coded on MRF during motion estimation (ME), as shown in Fig. 4.1 (b). The IRRS searches the intra-coded blocks among the candidate reference blocks in the previous multiple reference frames (F_n, F_{n-1}, and F_{n-2}), and selects a block with the largest region refreshed by intra-coding. Fig 4.1b shows that the selected block in F_n is fully or partially refreshed with intra-coded MBs. Therefore, there could be limited error propagation on the erroneous network. The minimum requirement of IRRS is that it should recognize the coding modes of all MBs in a MRF to search for the best matched region among the reference frames.



(a) Limitations of MRF



(b) Proposed error resilient reference selection method

Fig. 4.1 Effects of MRF with intra refresh

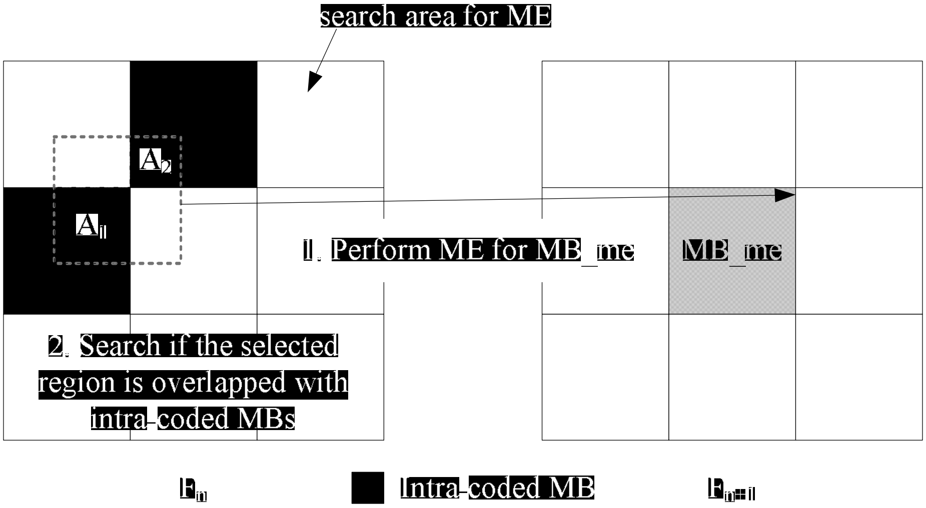


Fig. 4.2 The MV search procedure of the proposed error resilient reference selection method.

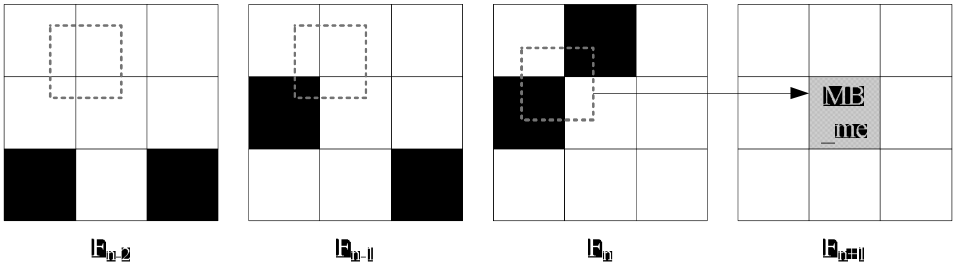


Fig. 4.3 Refresh ratio calculation procedure for current MB in reference frames

Fig. 4.2 shows the visualized motion vector (MV) search procedure of the proposed reference selection method. The dotted box in frame F_n is the block matched region that was chosen for the current MB depicted as a grey box in frame F_{n+1}. The matched region is fully or partially overlapped with the intra-coded MBs, which are shown as black boxes in frame F_n. A refresh ratio RR_i in IRRS is defined to select refreshed MB for references as follows:

$$RR_i = \frac{\sum_w A_w^i}{no.\ of\ pixels\ in\ MB}\ , \qquad (4.1)$$

Where A_w^i is w^{th} intra-coded area size (pixels) in i^{th} reference frame. For example, in Fig. 4.2, A_1 and A_2 are intra-coded areas, whereas dotted box implies the MB. This is calculated for each candidate MB selected from the RD cost optimization process in multiple reference frames. We used one candidate reference MB in each reference frame to be applied to RR_i in the simulation. After selection of candidate MBs, the IRRS chooses a MB with the maximum RR_i among them as a reference block for the current MB of F_{n+1}.

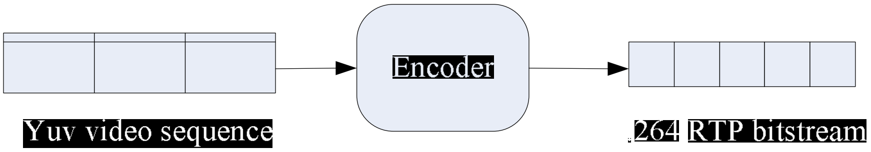
Fig. 4.3 shows the MV search procedure using RR_i for current MB of frame F_{n+1} in previous references frames. In frame F_{n-2}, the best matched MB does not overlapped with any intra-coded MBs, whereas in frame F_{n-1}, it overlapped partially with a intra-coded MB and in frame F_n, it overlapped with two intra-coded MBs, hence the block in frame F_n will be selected as best matched block for the current MB in frame F_{n+1} because of its highest RR_i value.

V. Simulation & Performance Comparison

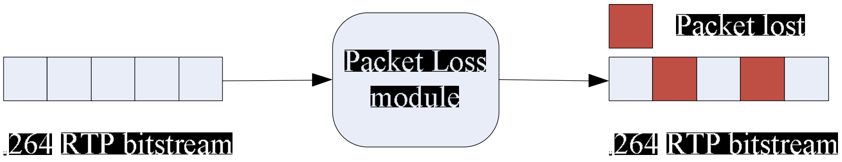
In this chapter, packet loss model, implementation of proposed method, simulation environment, simulation results and performance comparison are described.

A. Packet Loss Model

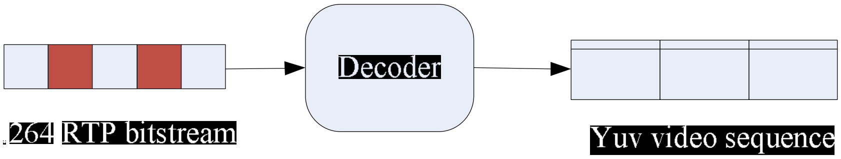
In wireless networks, packet loss may occur due to numerous reasons such as including link/node failures, route changes, and bit errors. These factors can cause both random and burst losses over the network. To investigate the effects of the video communications over such lossy networks,



(a) Encoder module



(b) RTP loss module



(c) Decoder module

Fig. 5.1 Available modules in H.264 reference software, JM 17.2

we first introduce a simple packet loss model that captures packet loss features in the network. As shown in Fig. 5.1, this model considers random packet loss available in JM reference software. It accepts real time protocol (RTP) coded bitstream and generates a erroneous RTP bitstream by randomly dropping fixed number of packets with different packet loss rate. This erroneous bitstream is then used in the decoder to reconstruct the YUV video sequence.

B. Implementation

IRRS scheme has been implemented in JM encoder version 17.2. It follows the flow chart given in Fig. 5.2 to calculate the refresh ratio for current MB in reference frames.

First, it finds the best matched block in reference frame, if it overlapped with intra-coded MBs, then calculate refresh ratio and then store it for further calculation. This procedure continues for all reference frames. Then in list prediction cost section, it selects the reference block with maximum refresh ratio as best matched block.

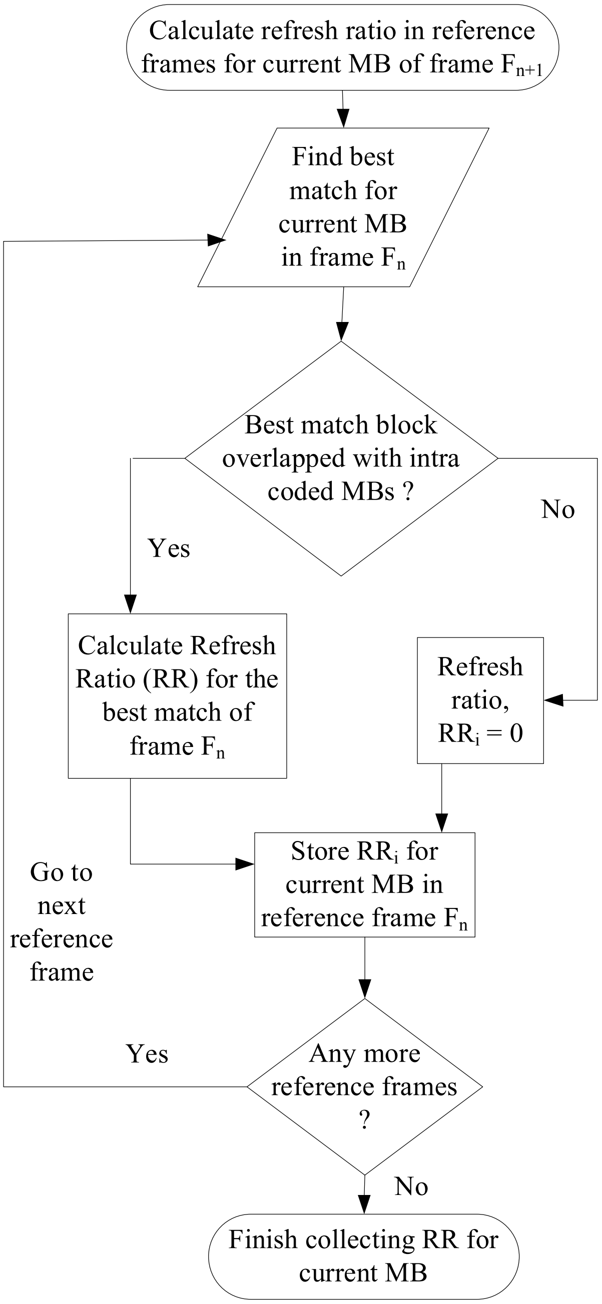


Fig. 5.2 Flow chart for refresh ratio calculation procedure for current MB in reference frames

C. Simulation Environment

Standard test video sequences (tennis, soccer, football, and stefan) were encoded at 30 frames per second based on the structure of seven reference frames. The H.264/AVC bit-stream was packetized in RTP mode. For the simulation of erroneous transmission, 0% to 10% of packets were randomly dropped from the RTP bit-stream. Error concealment at the decoder was set to frame copy. Number of candidate MBs for random intra refresh per frame was set to 1, 5, and 10. Basic characteristics of the video sequences are given in Table 5.1.

Table 5.1 Main characteristics of video test sequences

Test Sequences	Tennis	Soccer	Football	Stefan
Total Frames	148	148	258	298
Spatial Resolution	QCIF	QCIF	QCIF	QCIF
GOP	IPPP	IPPP	IPPP	IPPP
Motion Activity	Moderate	Fast	Fast	Fast
Scene change	Yes	No	No	No

Table 5.2 Encoder configuration setup file

Video sequence format	QCIF (176 x 144)
Profile level	High (100)
GOP	IPPP…
Number of reference frame	7
Random Intra Refresh	1, 5,10
Entropy coding	CABAC

Table 5.2 shows the encoder configuration parameter to encode the standard video sequences. After encoding the video sequences using RDRS and IRRS, RTP_loss module is used to randomly drop packets from encoded bit-streams with different packet loss rate from 0 to 10%. Then the decoder module is used to reconstruct the video sequence from erroneous bitstreams. PSNR value of luminance (Y) in dB is used as performance metric to compare the RDRS and IRRS method.

D. Simulation Results

The performance comparison between RDRS and IRRS has been done following three ways. In all cases the value of random intra refresh means the number of MBs forcefully intra coded per frame.

1) Individual frame analysis: Quality of each decoded frame is compared after transmitting the bit-stream over erroneous network of 5% packet loss rate. Corresponding results for RDRS and IRRS are shown in Fig. 5.3, 5.5, 5.7, and 5.9. In these figures, RDRS1 and IRRS1 imply RDRS and IRRS with single random intra refresh respectively.

2) Overall video quality over different packet loss rate: Quality of video sequences over erroneous network of different packet loss rate (0-10%) is compared between RDRS and IRRS. Here also single random intra refresh were used. Average PSNR in dB is used as performance metric. Corresponding results are shown in Fig.5.4, 5.6, 5.8, and 5.10.

3) Overall video quality for different random intra refresh: To check how RDRS and IRRS perform under different number of random intra refresh, 1, 5, and 10 candidate MBs are used and their corresponding results are given in Fig. 5.11 to 5.14. In these figures, RDRS1, RDRS5, RDRS10

mean RDRS with random intra refresh of 1, 5 and 10 respectively and IRRS1, IRRS5 and IRRS10 mean IRRS with random intra refresh of 1, 5 and 10 respectively. Overall performance gain of IRRS over RDRS is given in Table 5.3 to 5.6. Here $\Delta PSNR_1$ means the difference between average PSNR of IRRS1 and RDRS1. $\Delta PSNR_5$ corresponds to differences between IRRS5 and RDRS5 and $\Delta PSNR_{10}$ refers to IRRS10 and RDRS10.

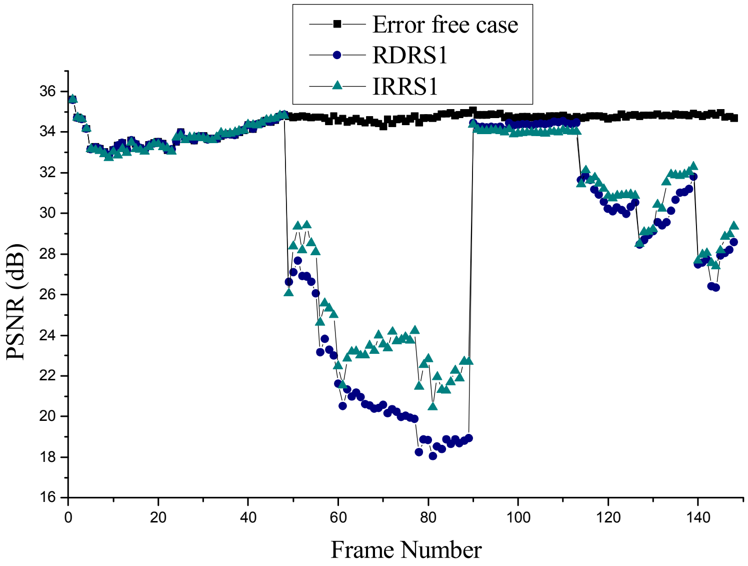


Fig. 5.3 Quality of tennis sequence with either RDRS or IRRS in erroneous network (PLR 5%).

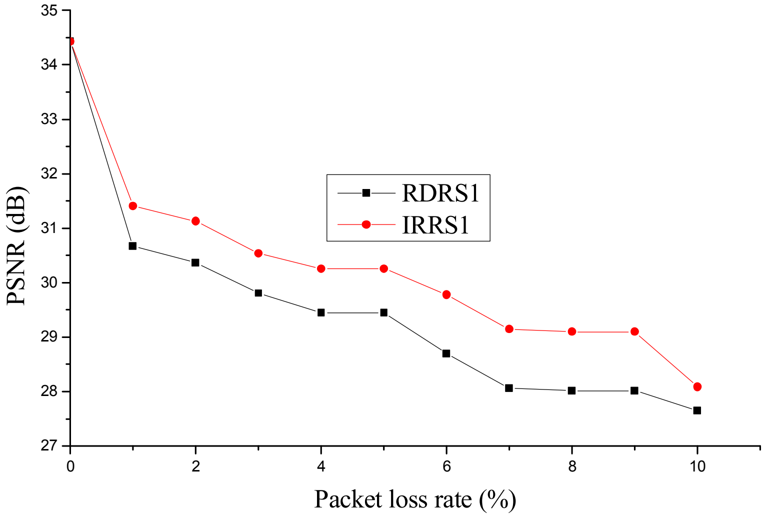


Fig. 5.4 Quality of tennis sequence with either RDRS or IRRS under different PLRs.

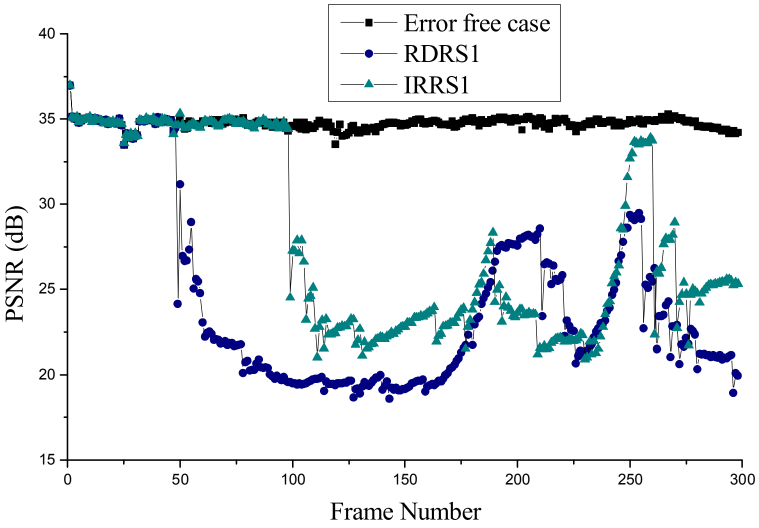


Fig. 5.5 Quality of stefan sequence with either RDRS or IRRS in erroneous network (PLR 5%).

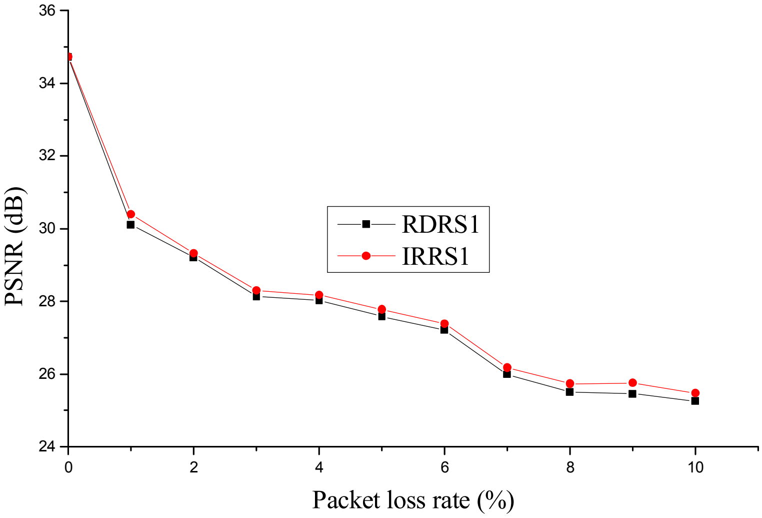


Fig. 5.6 Quality of stefan sequence with either RDRS or IRRS under different PLRs.

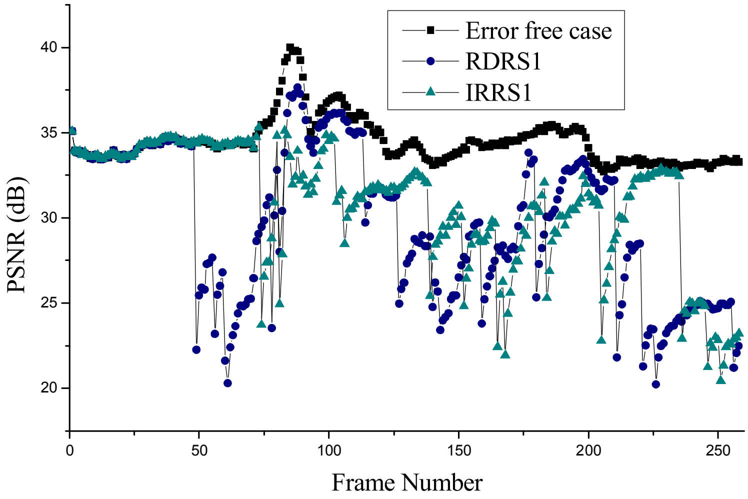


Fig. 5.7 Quality of football sequence with either RDRS or IRRS in erroneous network (PLR 5%).

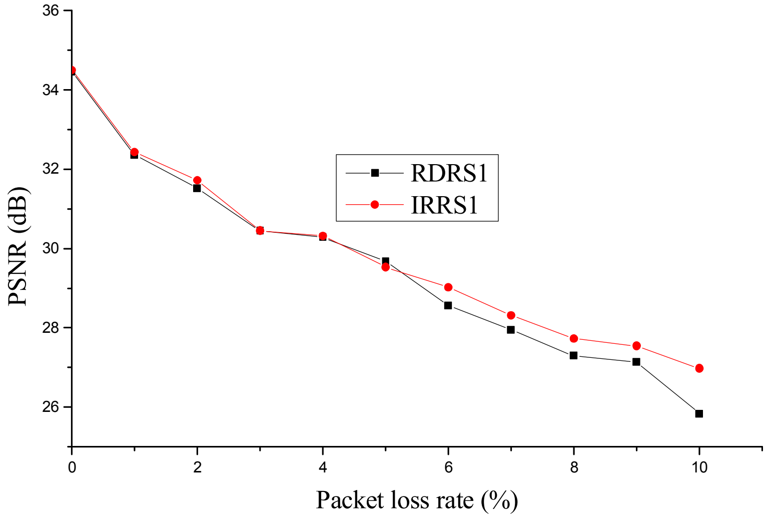


Fig. 5.8 Quality of football sequence with either RDRS or IRRS under different PLRs.

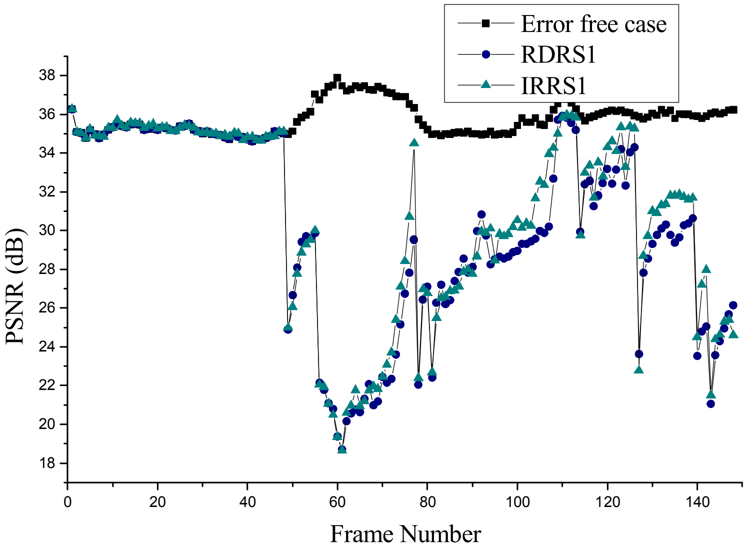


Fig. 5.9 Quality of soccer sequence with either RDRS or IRRS in erroneous network (PLR 5%).

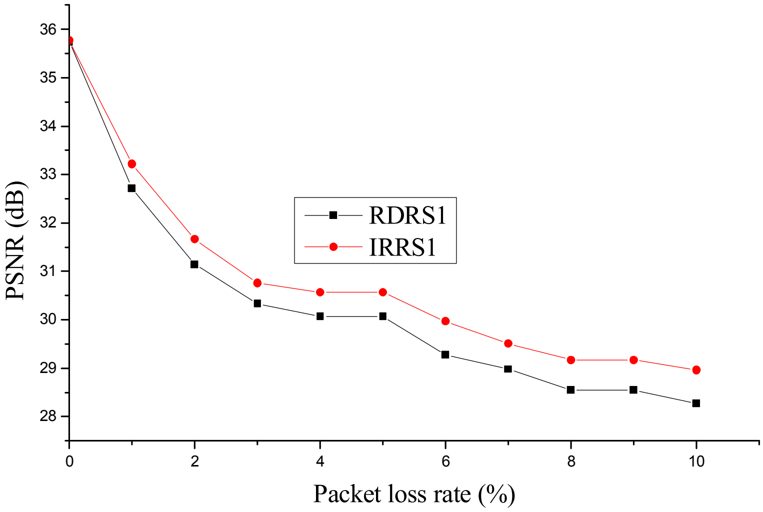


Fig. 5.10 Quality of soccer sequence with either RDRS or IRRS under different PLRs.

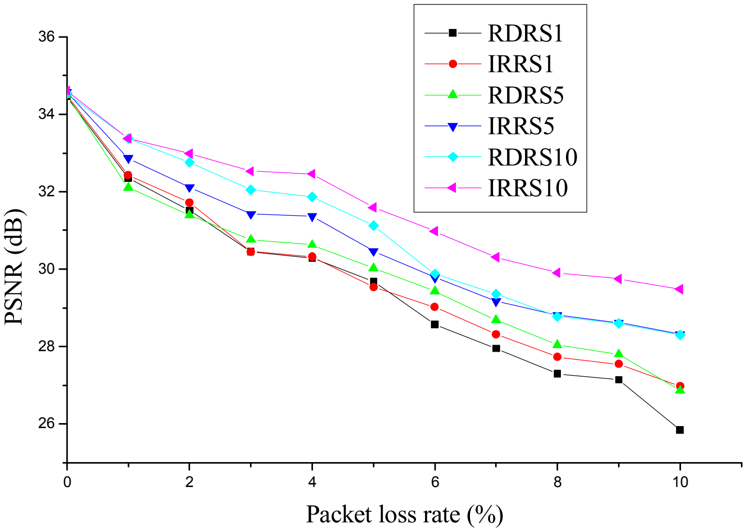


Fig. 5.11 Quality of football sequence with either RDRS or IRRS under different Intra Refresh.

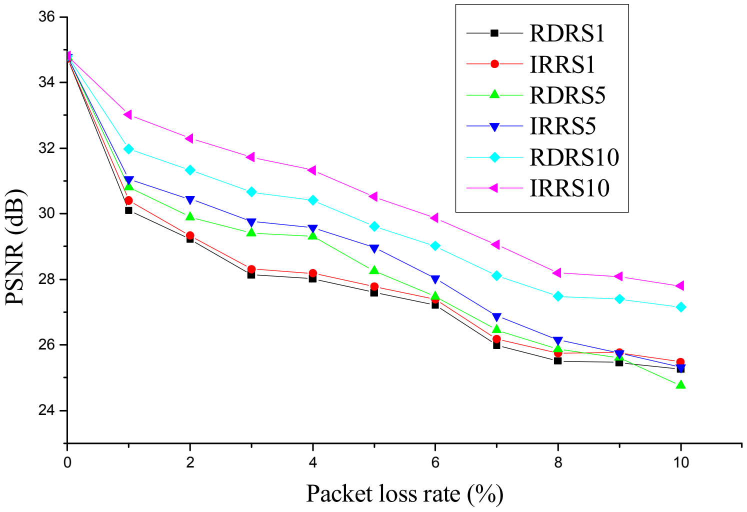


Fig. 5.12 Quality of stefan sequence with either RDRS or IRRS under different Intra Refresh.

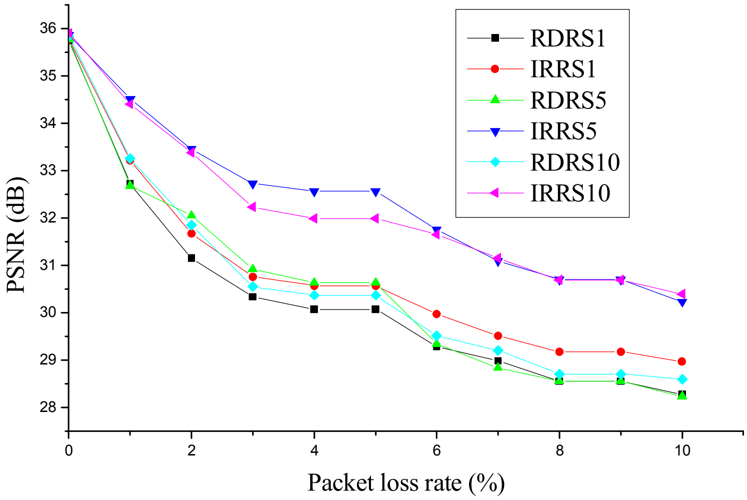


Fig. 5.13 Quality of soccer sequence with either RDRS or IRRS under different Intra Refresh.

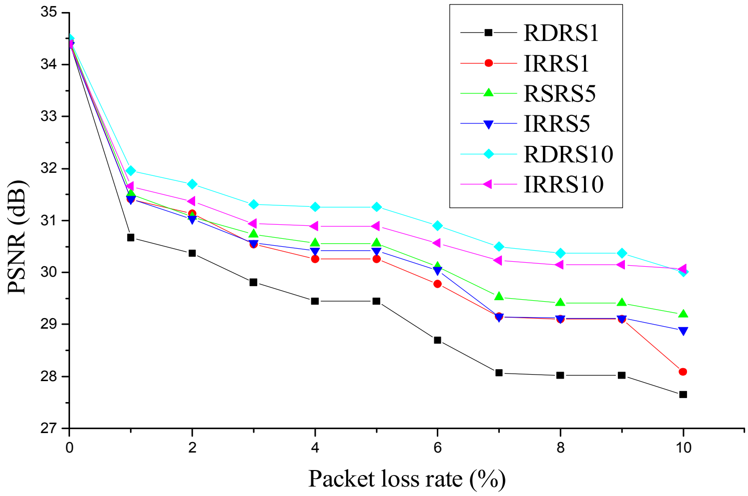


Fig. 5.14 Quality of tennis sequence with either RDRS or IRRS under different Intra Refresh.

Table 5.3 Performance gain (dB) of IRRS over RDRS method against different random Intra for football sequence

PLR (%)	$\Delta PSNR_1$	$\Delta PSNR_5$	$\Delta PSNR_{10}$
0	+0.04	+0.07	+0.07
1	+0.07	+0.76	0
2	+0.2	+0.72	+0.23
3	0	+0.67	+0.47
4	+0.03	+0.74	+0.59
5	-0.15	+0.45	+0.47
6	+0.45	+0.36	+1.1
7	+0.37	+0.49	+0.95
8	+0.43	+0.77	+1.14
9	+0.41	+0.82	+1.16
10	+1.14	+1.46	+1.18

Table 5.4 Performance gain (dB) of IRRS over RDRS method against different random Intra for stefan sequence

PLR (%)	$\Delta PSNR_1$	$\Delta PSNR_5$	$\Delta PSNR_{10}$
0	+0	+0.01	+0.01
1	+0.3	+0.23	+1.05
2	+0.11	+0.55	+0.96
3	+0.17	+0.36	+1.07
4	+0.16	+0.28	+0.92
5	+0.19	+0.72	+0.9
6	+0.18	+0.55	+0.85
7	+0.19	+0.43	+0.95
8	+0.23	+0.29	+0.71
9	+0.3	+0.13	+0.69
10	+0.22	+0.56	+0.64

Table 5.5 Performance gain (dB) of IRRS over RDRS method against different random Intra for soccer sequence

PLR (%)	$\Delta PSNR_1$	$\Delta PSNR_5$	$\Delta PSNR_{10}$
0	+0.02	+0.1	+0.06
1	+0.50	+1.83	+1.14
2	+0.52	+1.4	+1.53
3	+0.43	+1.81	+1.68
4	+0.5	+1.93	+1.62
5	+0.5	+1.93	+1.62
6	+0.69	+2.41	+2.14
7	+0.53	+2.26	+1.95
8	+0.62	+2.15	+1.99
9	+0.62	+2.15	+1.99
10	+0.69	+2.00	+1.8

Table 5.6 Performance gain (dB) of IRRS over RDRS method against different random Intra for tennis sequence

PLR (%)	$\Delta PSNR_1$	$\Delta PSNR_5$	$\Delta PSNR_{10}$
0	-0.01	-0.08	-0.11
1	+0.74	-0.10	-0.30
2	+0.76	-0.04	-0.33
3	+0.73	-0.16	-0.37
4	+0.81	-0.14	-0.37
5	+0.81	-0.14	-0.37
6	+1.08	-0.06	-0.33
7	+1.08	-0.38	-0.27
8	+1.08	-0.29	-0.22
9	+1.08	-0.29	-0.22
10	+0.44	-0.30	+0.06

Fig. 5.15 Original frame from tennis sequence

Fig. 5.16 Subjective video quality of a frame of tennis sequence by using RDRS method

Fig. 5.17 Subjective video quality of a frame of tennis sequence by using IRRS method

E. Performance Comparison

To evaluate performance of IRRS, two types of video sequences were used, video sequence with frequent scene-changes and video sequence with fast movement objects. Among the test video sequences used in the simulation, tennis sequence has some scene changes within the frames. In the case of scene-change between frames, encoder uses intra mode to encode the whole frame. The other video sequences used in the simulation have fast movement between frames.

In the figures for individual frame comparison, both RDRS and IRRS are observed showing the similar quality performance in error free case. But IRRS recovered the video better than RDRS in all the video sequences when packet loss starts. Here, in tennis sequences, it shows that from damages, there is sudden increase in PSNR at frame number 90. This is because the scene change occurred at the frame 90. As in the case of scene change, intra coding mode was used. In subjective video quality comparison, figures showing in Fig. 5.15 to Fig. 5.26, one particular frame is chosen from the corresponding video sequences, it is observed that the visual quality for IRRS method is better than RDRS in all video sequences.

From the figures of overall video quality over different packet loss rate, it can be concluded that IRRS performed better than RDRS for all test sequences. Among the video sequences, IRRS performs comparatively better for tennis sequence than the other sequences. The number of random intra refresh value has an impact on overall performance according to the different random intra refresh value 1, 5, and 10. In case of scene-change sequence, using single forced intra-coded MB performance, the IRRS performs better than RDRS in different packet loss scenario. But if more MBs are forced to be intra-coded like 5 or 10, IRRS performs slightly less than RDRS method.

This is because of the low coding efficiency obtained by forcing higher intra coded MB blindly at the frequent scene-change frames. On the other hand, in case of fast movement sequences, IRRS performs better than RDRS over different random intra refresh values. As the more MBs are forcefully intra-coded, the performance of IRRS is getting better than RDRS. From the performance gain tables, overall gain of IRRS over RDRS can be up to + 2dB in erroneous network for fast movement video sequences, and performance gain of IRRS can be up to +1dB in erroneous network for scene-change video sequence especially in lower intra refresh ratio.

VI. Conclusion

H.264/AVC is a standard for video compression and most commonly used formats in video coding. However, in erroneous network environment, error resilient methods are very important to provide robust video bitstream. In this thesis, standard error resilient methods available in H.264 video codec are briefly described. Some alternative error resilience methods are introduced. Which error resilience method should be used in encoding is depends on network environments, type of video sequences and other parameters like bit rate, complexity and optimization. It is preferable to use error resilience methods without adding redundancy because of the tradeoff between compression and redundancy. Random intra refresh is a good choice as it stops error propagation. Multiple reference frames has higher coding efficiency. But to use multiple reference frames along with random intra refresh has some performance degradation problem because of the traditional RD optimization as it always prefer to use inter coding over intra coding for their best RD cost. Also allowing more number of random intra refresh has a negative impact on overall video quality in erroneous network due to the RDO technique.

In this thesis, an error resilient reference selection method, IRRS is shown to improve overall video quality transferred in erroneous network by selecting more intra refreshed regions as reference block in inter prediction. Instead of choosing the best RD cost, the proposed method choose a reference block by defining a custom refresh ratio function that selects a reference block that has most intra refreshed region.

From simulation results, the overall gain (mean PSNR) of the proposed method could be achieved to +2dB depending on packet loss rate for fast movement video sequences using multiple random intra refresh, i.e. by forcing 10 MBs to be intra coded while + 1dB for scene change video sequences using single random intra refresh per frame. For fast movement video sequences, overall performance increased by introducing more number of random intra refresh. In case of performance gain in specific region like between frames where random packet loss happened, the proposed method recovers quality degradation much better than the traditional method.

This proposed error resilient method can be used in the most of the applications developed on consumer electronics, mobile, and vehicular devices.

References

[1]. A. Puri, X. Chen, A. Luthra, “Video coding using the H.264/AVC compression standard,” *Signal Processing: Image Communication*, Vol. 19, No. 9, pp 793-849, October 2004.

[2]. I. E. Richardson, “The H.264 advanced video compression standard,” *ISBN: 978-0-470-51692-8,* 2010.

[3]. M. Rup, “Video and multimedia transmission over celluar networks”, *ISBN: 978-0470699331,* 2007.

[4]. P. Nunes, L. D. Soarces, Fernando Pereira, “Error Resilient MB Rate Control for H.264/AVC video coding”, *IEEE International Conference on Image Processing,* pp. 2132-2135, 2008,

[5]. K. –P. Lim, G. Sullivan, T. Wiegand, “Text description of joint model reference encoding methods and decoding concealment methods”, Doc. JVT-X101, ITU-T VCEG Geneva metting, Switzerland, 2007.

[6]. T. H. Vu and S. Aramvith, “Joint FMO and Adaptive Intra Refresh for Error Resilience in H.264 Video Coding”, *Asia Pacific Signal and Information Processing Association,* pp. 449-452, October 2009.

[7]. N. Qadri, M. Altaf, M. Fleury, M. Ghanbari, H. Sammak, “Robust video Streaming over an Urban VANET”, *IEEE International Confernence on Wireless and Mobile Computing, Networking and Communications,* pp. 429-434, October 2009.

[8]. B. Katz, S. Greenberg, N. Yarkoni, N. Blaunstien, R. Giladi, “New Error Resilience Scheme Based on FMO and Dynamic Redundant Slices Allocation for Wireless Video Transmission”, *IEEE Transactions on Broadcasting,* Vol. 53, No. 1, March 2007.

[9]. N. Thomas, S. Argyropoulos, N. V. Boulgouris, M. G. Strintizis, “Error Resilient Transmission of H.264/AVC streams using Flexible MB Ordering”, *The 2nd European Workshop on the Integration of knowledge, Semantics and Digital Media Technology,* pp. 183-189, November 2005.

[10]. I. Radulovic, P. Frossard, Y.-K. Wang, M. M. Hannuksela, and A. Hallapuro, “Multiple Description Video Coding with H.264/AVC Redundant Pictures”, *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 20, No. 1, January 2010.

[11]. S. Wenger, “H.264/AVC over IP”, *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 13, No. 7, pp. 645-656, *July 2003.*

[12]. S. Moiron, I. Ali, M. Ghanbari, and M. Fleury, “Limitations of multiple reference frames with cyclic intra-refresh line for H.264/AVC”, *ELECTRONIC LETTERS,* Vol. 47, No. 2, pp. 103-104*,* January 2011.

[13]. H.264/AVC Reference Software, JM 17.2, available at http://iphome.hhi.de/suhering/tml/download.

[14]. T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, “Overview of the H.264/AVC Video Coding Standard”, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 13, No. 7, July 2003.

[15]. T. Stockhammer, M. M. Hannuksela, and T. Wiegand, “H.264/AVC in Wireless Environments”, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 13, No. 7, pp. 657-673, July 2003.

[16]. C.-C Su, H. H Chen, J. J. Yao, P. Huang, “H.264/AVC-based multiple description video coding using dynamic slice groups”, *Signal Processing: Image Communication*, pp. 677-691, July 2008.

[17]. P. Lambert, W. D. Neve, Y. Dhondt, R. V. de Walle, "Flexible macroblock ordering in H.264/AVC", *Elsevier Journal of Visual Communication and Image Representation*, pp. 358-375, August 2005.

[18]. Y. C. Chang, S. W. Lee and R. Komyia, "A Fast Forward Error Correction Allocation Algorithm for Unequal Error Protection of Video Transmission over Wireless Channels", *IEEE Transactions on Consumer Electronics*, Vol. 54, No. 3, pp. 1066-1073, August 2008.

[19]. L. D. Soares and F. Pereira. "Adaptive Shape and Texture Intra Refreshment Schemes for Improved Error Resilience in Object-Based Video Coding", *IEEE Transactions on Image Processing*, Vol. 13, No. 5, pp. 662-676, May 2004.

[20]. C.-C. Kuo, G.-S. Ma, M.-S. Chen, and J.-C. Chen, "Adaptive Transmission Control for Error-Resilient Multimedia Synchronization", *IEEE Transactions on Consumer Electronics*, Vol. 50, No. 2, pp. 700-708, May 2004.

[21]. S. Kumar, L. Xu, M. K. Mandal, and S. Panchanathan, "Error Resiliency Schemes in H.264/AVC Standard", *Elsevier J. of Visual Communication & Image Representation (Special issue on Emerging H.264/AVC Video Coding Standard),* Vol. 17, No. 2, April 2006.

[22]. Y. Guo, Y. Chen, Y.-K. Wang, H. Li, Miska M. Hannuksela, and M. Gabbouj, "Error Resilient Coding and Error Concealment in Scalable Video Coding", *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 19, No. 6, pp 781-795, June 2009.

[23]. Y. Zhang,W. Gao, Y. Lu, Q. Huang, and D. Zhao, "Joint source-channel rate distortion optimization for H.264 video coding over error-prone networks", IEEE Transactions on Multimedia, Vol. 9, No. 3, pp. 445–454, April 2007.

[24]. O. Harmanci and A. M. Tekalp, "A stochastic framework for rate-distortion optimized video coding over error-prone networks", IEEE Transactions on Image Processing, Vol. 16, No. 3, pp. 684–697, March 2007.

[25]. B. A. Heng, J.G. Apostolopoulos, J.S Lim, "End-to-end rate-distortion optimized mode selection for multiple description video coding", *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 5, pp. 905-908, 2005.

[26]. D. Li, J. Pan, "Performance Analysis and Evaluation of H.264 Video Streaming over Multi-hop Wireless Networks", *IEEE Global Telecommunications Conference*, pp. 1-5, 2008.

[27]. Y. Liao, J.D., Gibson, "Rate Distortion Based Mode Selection for Video Over Wireless Network with Burst Losses", *17th International Packet Video Workshop*, pp. 1-10, 2009

[28]. J. Mochnáč, S. Marchevský, P. Kocan, "Simulation of Packet Losses in Video Transfers Using Real-time Transport Protocol", 20th International Conference on Radioelektronika, pp. 1-4, 2010.

[29]. D. Li, J. Wu, "The test and analysis of FMO model in error concealment based on H.264", *International Conference on Computer Science & Education,* pp. 330-332, 2011.

[30]. P. Baccichet, S. Rane, A. Chimienti, and B. Girod, "Robust Low-Delay Video Transmission using H.264/AVC Redundant Slices and Flexible Macroblock Ordering", *IEEE International Conference on Image Processing,* Vol. 4, pp. IV93 – IV96, 2007.

[31]. F.-C. Chen, Y.-P. Hsu, "Rate-distortion optimization of H.264/AVC rate control with novel distortion prediction equation", *IEEE Transactions on Consumer Electronics*, Vol. 57, No. 3, pp. 1264-1270, 2011.

[32]. I. Ali, M. Fleury, S.Moiron, and M. Ghanbari, "Content dependent intra-refresh placement for video streaming", *3rd Computer Science and Electronic Engineering Conference,* pp. 35-39, 2011.

[33]. P. Nunes, L.D.Soares, and F. Pereira, "Automatic and adaptive network-aware macroblock intra refresh for error-resilient H.264/AVC video coding", *16th IEEE International Conference on Image Processing,* pp. 3073-3076, 2009.

[34]. I. Ali, Moiron, Sandro, Ghanbari, Mohammed, Fleury, and Martin, "Prioritized packetization for video with intra-refresh macroblock line", *IEEE International Conference on Multimedia and Expo,* pp.1-6, 2011.

[35]. Y. W. Huang, B.-Yu Hsieh, S.-Yi Chien, S.-Y. Ma and L.-G. Chen, "Analysis and complexity reduction of multiple reference frames motion estimation in H.264/AVC", *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 16, No. 4, pp. 507-522.

List of Publications

1. **Shaikhul Islam Chowdhury**, Jae-Young Pyun, "A Survey on Error Resilience Schemes of H.264/AVC for Erroneous Network", *Proceedings of KISM Spring Conference*, Korea, pp. 286- 289, 2012.

2. **Shaikhul Islam Chowdhury**, Won-Il Lee, Youn-Sang Choi, Guen-Young Kee, Jae-Young Pyun, "Performance evaluation of reactive routing protocols in VANET", *17th Asia-Pacific Conference on Communications,* pp. 559 - 564 , 2011.

3. Won-Il Lee, **Shaikhul Islam Chowdhury**, Guen-Young Kee, Woon-Sung Baek, Jae-Young Pyun, "Velocity Aware Multipath Distance Vector Routing Protocol for High Mobility over Vehicular Ad-Hoc Networks", *5th FTRA International Conference on Multimedia and Ubiquitous Engineering,* pp. 189-194, 2011.

ABSTRACT

Advanced H.264/AVC Error Resilient Video Coding over Erroneous Wireless Network

Shaikhul Islam Chowdhury
Advisor: Prof. Jae-Young Pyun, Ph. D.
Department of Information and
Communications Engineering,
Graduate School of Chosun University

With the increasing demands of real-time multimedia communications like video calling, video conferencing, and video streaming, both video compression and transmission robustness are getting important in every video codec standard. Internet of yesterday was full of static pages, but Internet of today is full of dynamic pages with embedded video streaming. Same case goes with mobile communications. For example, cellular phone of yesterday was used only for voice communication and it is more engaged with data services today. The demand for multimedia communication is vital in modern communication. Everybody wants crystal clear picture especially in their important broadcasting and/or streaming media programs like football world cup or national election. By the way, real time video services are particularly challenged over mobile communication systems due to the wireless channel quality variations.

In order to provide the video transmission service, video compression codecs are required. The combination of these compression methods provides high compression gain, but makes the encoded bit stream more error-prone at the same time. This is because recovering the distorted frame is difficult if errors happen in frames in erroneous network. Video quality suffers from significant degradation when video are transmitted over error-prone channel caused by packet loss, errors caused by fading in wireless channel. Therefore, error resilience techniques are required in the encoder to mitigate the video quality degradation. The preventive error resilience techniques based on the intelligent design of the encoder makes the task of the decoder much easier in terms of dealing with errors.

In this thesis, different standard and alternative error resilience techniques are described and compared for performance evaluation as videos are transferred over the erroneous network. Specially, multiple reference frame technique adopted in H.264/AVC is analyzed when random intra refresh is used for the strong error resilience role. Then, proposed intra-refresh based reference frame selection method is presented to overcome the shortcoming of multiple references frame combined with random intra refresh.

Acknowledgement

All praise is due to Allah, Lord of the worlds. I would like to thank all who have contributed one way or the other to make this thesis possible. The completion of this thesis owes a great deal to the people around me, my parents, and my family. I would like to express my sincere gratitude to my supervisor, Prof. Jae-Young Pyun for his invaluable support, encouragement, supervision, personal guidance, and useful suggestions throughout the course of my research work. I am also thankful to the entire Professors of Information and Communication Engineering department from whom I have learnt great deal of knowledge.

My sincere thanks to committee members, Prof. Jong An Park and Prof. Seung Jo Han, for their detail review, constructive criticism and excellent advice during the preparation of this thesis.

I also like to express my thanks to all my lab mates of Wireless and Mobile Computing Systems Lab and to all my friends and seniors in Korea for their cooperation and support and making my stay memorable.

I am as ever, especially indebted to my parents and family for their love, support, and encouragement in every moment of my life. I also wish to thank them for their understanding during my study.

The financial support of Korean Government Global IT scholarship committee members and Chosun University are greatly acknowledged.