國立交通大學

電子工程學系 電子研究所碩士班

碩 士 論 文

一個具有 R-D 最佳化內部更新的容錯 MPEG-4 編碼器

A Robust MPEG-4 Encoder with Rate-Distortion Optimized Intra Refreshment

研 究 生：蘇子良

指導教授：蔣迪豪 博士

中 華 民 國 九 十 三 年 七 月

一個具有 R-D 最佳化內部更新的容錯 MPEG-4 編碼器

A Robust MPEG-4 Encoder with Rate-Distortion Optimized Intra Refreshment

研究生: 蘇子良　　　　Student: Tzu-Liang Su
指導教授: 蔣迪豪　　　　Advisor: Tihao Chiang

國 立 交 通 大 學
電子工程學系 電子研究所碩士班

碩 士 論 文

A Thesis
Submitted to Department of Electronics Engineering & Institute of Electronics
College of Electrical Engineering and Computer Science
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master of Science
in
Electronics Engineering

July 2004

HsinChu, Taiwan, Republic of China

中華民國九十三年七月

一個具有 R-D 最佳化內部更新的容錯 MPEG-4 編碼器

研究生: 蘇子良　　　　　　　　　　　　　指導教授: 蔣迪豪 博士

國立交通大學

電子工程學系 電子研究所碩士班

摘要

MPEG-4 視訊編碼標準提供一個有效的方法在低位元率藉由網際網路或無線通道來傳送視訊內容以為串流或通訊之用。兩種傳輸通道皆採用封包轉換通訊協定作為資料傳輸用。當封包在有雜訊的通道上傳輸時，會發生封包遺失或位元錯誤。傳輸錯誤會顯著地降低視覺上的品質。為了使位元串流在不可信任的網路傳輸中更強韌，MPEG-4 編碼標準提供數種工具來支援容錯機制。特別的是，在編碼器端的適應性內部更新被用來減緩錯誤傳遞。為了更進一步地改善視覺品質並平衡編碼效率，我們提出一個編碼率及失真最佳化的內部更新以更有效率的方法來分配 intra 巨方塊(macroblock)。

我們所提出內部更新的方法將根據視訊特性以及封包轉換串流方式和網路狀況的性質來加入 intra 巨方塊。所考慮的視訊特性包含動作訊息、畫面間預測的時間、每個畫面中 intra 巨方塊的數量、以及編碼率和品質失真的總和成本。動作訊息和畫面間預測的關係可以用來計算每個巨方塊的參照記錄。我們使用每個畫面中 intra 巨方塊的數量來控制我們在一個位元流所加入的額外 intra 巨方塊在一定的數量之內。

為了減少一個封包內有著連續的 intra 巨方塊，一旦遺失之後所可能造成突發性的視訊品質下降，我們對於視訊封包化所造成的影響也加以詳查。根據我們創新的內部更新，我們可以改善位元流對於傳輸錯誤的強韌度來提供高品質的視訊內容。實驗結果顯示我們可以顯著地改善主觀上所看到經由網際網路和無線通道所傳輸的視訊內容。

A Robust MPEG-4 Encoder with Rate-Distortion Optimized Intra Refreshment

Student: Tzu-Liang Su Advisor: Dr. Tihao Chiang

Department of Electronic Engineering &
Institute of Electronics
National Chiao Tung University

Abstract

MPEG-4 video coding standard has provided an efficient way to deliver video contents at a low bit-rate for streaming or communications over the Internet or wireless channels. Both channels adopt a packet-switching protocol for data transmission. When the packets are transmitted via channels, the packet loss or bit error occurs. The occurrence of transmission errors significantly degrades the visual quality. To make the bitstream more robust for unreliable network transmission, MPEG-4 coding standard provides tools to support error resilience. In particular, Adaptive Intra Refreshment (AIR) at the encoder is used to alleviate error propagation. To further improve the visual quality and balance the coding efficiency, we propose a rate-distortion optimized intra refreshment (RDIR) to distribute intra macroblocks in a more efficient way.

The RDIR inserts intra MBs based on video characteristics, properties of packet-switched streaming schemes and network conditions. The video characteristics considered consist of motion information, duration of inter-frame prediction, the number of intra MBs per frame and R-D cost. The motion information for various bit rates and the R-D cost for different packet loss ratios are combined as a measure, which is referred to as a historical record. With the stored historical record for each MB, the duration of inter-frame prediction is obtained for various bit rates and different packet loss ratios. To reduce overhead, we use the number of intra MBs per frame to control the total number of intra MBs with constrained overhead in the bitstream.

To reduce burst quality degradation by loss of successive intra MBs in single stream packet, the impact of video packetization prior to delivering the bitstream via the packet-switched channels is investigated. We uniformly distribute the intra MBs in a frame.

Based on our innovative intra refreshment approaches, we can enhance bitstream robustness against the transmission error to provide high quality video contents. Our results show that we can significantly improve subjective quality of video contents that are delivered over the Internet and wireless channels.

誌謝

在研究所的這兩年當中，首先要感謝指導教授蔣迪豪教授，由教授在帶實驗室的衝勁、對學生的要求與期許、以及協助就業的過程當中，從教授身上學到、感受到了很多的東西，在此對蔣教授獻上最大的感謝之意。

感謝所有其他實驗室的學長與同學，尤其是王俊能與王士豪學長，嚴謹認真的研究態度，研究上需要注意的事項，與其他很多的協助與引導，都給了我在研究所生活當中很好的磨練。而曾在五樓實驗室的的孝強、耀諄、俊安、振章、岳賢、子瀚、明哲也在學習過程中提供了很多的幫助與一起學習的回憶，還要感謝在相同領域研究的名彥，總會適時的提醒我往正確的方向前進；還有五樓三位可愛的學弟妹，帶給我很多歡樂。以及其他通訊電子暨訊號處理實驗室的學長姐、同學、學弟妹，大家營造出了一個氣氛融洽的研究環境，陪伴我走過這兩年的研究生活。

感謝我兩年來的室友，包容我生活上缺點，讓我在研究上無後顧之憂。

感謝從小辛苦把我扶養長大的父母和一直照顧我的兄姐，有你們的用心，你們的付出，才能夠造就我今日的成就。我也會再接再厲，來報答你們的恩情。

其他還要感謝很多的親朋好友，有你們的扶持與鼓勵，讓我渡過了很多的艱難與考驗。最後，僅以這篇論文，獻給所有陪伴我走過這段日子的人，謝謝。

蘇子良

民國九十三年七月 於新竹

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Multimedia communications systems are more and more popular today. For video transmission or streaming over different networks, bitstreams transmitted over the error-prone environment may be corrupted by random error or packet loss of the channels. It's a challenge to provide a satisfactory quality of service for users. To address the problems in providing a proper quality of service, error resilience and robustness tools are used to recover the video contents from the corrupted bitstreams during transmission. The error resilience tools of MPEG-4 coding standard that has been adopted for applications over unreliable networks have been realized and applied to various multimedia players. The error recovery methods utilize all the useful information available at the receiver to re-synchronize the decoding processing.

The existing error resilience techniques can be separated into 4 major stages including the error detection, the error recovery, the error localization, and the error concealment as shown in **Fig. 1**. The error detection process prevents the decoder system from failure during a decoding process. The error detection process is critical for the next three steps. Once the errors are detected, the decoding process of the received bitstream will be stopped. The error localization helps locate the error and the error recovery resumes the decoding process based on header information from the partially damaged bitstream. Finally, the error concealment techniques enhance the visual quality of received video based on the reconstructed frames and MBs. Thus, with the error detection, the error localization, the error recovery and the error concealment, the MPEG-4 video can be applied for the multimedia applications in an error prone environment.

For video content delivery, video coding standards are commonly used. For compact storage and efficient transmission of video contents, video coding standards including MPEG-1/2/4 and H.263/H.264 have been widely used. Based on the motion information and temporal or spatial correlation, video coding standards can compress the video data to a smaller size for the limited storage space or transmission bandwidth. Without compression, a huge amount of transmission bandwidth is wasted for delivering video.

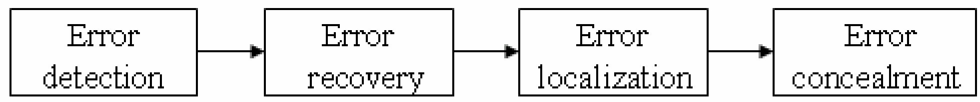


Fig. 1. Error resilience process

For the past decade, multimedia coding standards focused on the compression ratio and improved performance under some coding constraints. Many coding techniques have been employed to reduce temporal redundancies and spatial redundancies such as inter-picture prediction and discrete cosine transform (DCT). The inter-picture prediction takes advantage of temporal redundancies and only considers the difference or residual between the reference frame and the current frame with motion estimation. As for transmission over packet-switched network, prediction loop may propagate errors and the received video quality may deteriorate severely. The errors of the current frame will drift to the subsequent frames in temporal and spatial domains due to motion estimation and intra prediction.

Propagation of packet loss errors can only be stopped with intra coded macroblock (MB). Although intra coding is very efficient in preventing the error propagation, the increasing number of the intra MBs in the bitstream will take many bits, which decreases the coding efficiency. Video quality degradation depends on the number and location of intra MBs. Adaptive intra refreshment has been proposed in MPEG-4 standard to place intra MBs to proper positions [32] Some existing intra refreshment methods [34] -[40] have been adopted to stop the error propagation. To improve the coding efficiency, rate-distortion intra refreshment (RDIR) [25] has been used. To further balance the coding efficiency and bitstream robustness, a novel rate-distortion (R-D) optimized intra refreshment is presented.

The R-D optimized intra refreshment inserts intra MBs based on video characteristics, properties of packet-switched streaming schemes and network conditions. The video characteristics consist of the motion information, the duration of inter-frame prediction, the number of intra MBs per frame and the R-D cost. The two factors including the motion information for video coding at various bitrates and the R-D cost for different packet loss ratios are combined as a measure, which is called as a historical record. With the historical record for each MB, the proper duration of inter-frame prediction is obtained with investigations under various bit rates and different packet loss ratios. In addition, we use the number of intra MBs per frame to control the total number of intra MBs within the encoded bitstream, which can allocate overhead bits more uniformly into the encoding bitstream to increase the error robustness.

To retain the better quality of service in video content delivery, the methods of video packetization prior to bitstream streaming via the packet-switched channels to the decoding end are investigated. To avoid burst quality degradation by the loss of successive intra MBs in a single stream packet, a novel counter is used to record the run length of successive intra MBs. Thus, based on the innovative intra refreshment approaches, we can increase bitstream robustness against the transmission error to provide better quality of service over the Internet and wireless channels, which both adopt packet-switching protocols for data transmission.

To further address the issue for providing the best video quality, terminal capabilities of consumers are considered. As to the terminal capabilities, we will focus on error robustness and error recovery capabilities of the media encoder and decoder. In addition, we will analyze the impact of each tool used. The results show that we can significantly improve subjective quality of video contents that are delivered over the Internet and wireless channels.

The organization of this thesis is described as follows. Chapter 2 summarizes the related MPEG-4 video coding systems and error resilience techniques. In addition, we survey existing intra refreshment approaches for further performance comparisons. For practical performance evaluation, we introduce MPEG-21 part-12 multimedia test bed for resource delivery. Chapter 3 explains the proposed RDIR framework that is realized for the MPEG-4 reference encoder. Within the RDIR framework, we will explain the concept and rationale of each intra refreshment tool including the R-D cost, the historical record, the systematic insertion of intra MBs, and the run length of successive intra MBs. Chapter 4 gives the experimental results using the proposed RDIR algorithm under various bit rates and different network conditions. Chapter 5 draws the conclusions.

Chapter 2

Error Resilience and Intra Refreshment for MPEG-4 Video Coding

In this chapter, we describe related works on constructing a streaming video system with error resilience capabilities. The streaming video system is built based on a subset of visual coding tools of the MPEG-4 video coding standard. In addition, the streaming system provides error resilient services to serve robust bitstreams for video streaming or communications. The error resilience tools can recover the error occurred during the network transmission. To get better quality when we lose packets, the error concealment algorithms and error resilient schemes are employed. The streaming system is integrated into a multimedia test bed that can simulate a variety of network conditions and evaluate error resilience performance for on-line testing.

2.1 Overview of MPEG-4 Standard

MPEG-4 video coding standard provides a rich set of tools for audio-visual objects coding. We will provide an overview of the natural visual coding tools of video specification. In addition, we will introduce error resilience tools including Resynchronization Marker, data-partition, and reversible variable length coding (RVLC) defined in MPEG-4 video specification.

2.1.1 MPEG-4 Video Coding Standard

MPEG-4 video coding standard is developed to provide users a new level of performance for various video communications applications such as Internet streaming and mobile multimedia applications. For the applications, both device power and channel bandwidth are very limited. For handheld devices using the batteries, the encoding process of video contents is sped up for minimizing power consumption. Thus, it is important to realize a low-complexity real-time software or hardware video decoder. To address issues of low complexity and minimal power consumption, many fast algorithms have been proposed to reduce the encoding and decoding complexities.

MPEG-4 encoder as shown in **Fig. 2** consists of six common modules in encoding process. The six encoding modules are discrete cosine transform (DCT), motion estimation, motion compensation (MC), zigzag scan, quantization (Q) and variable length coding (VLC). With the DCT, a spatial domain data with high inter-pixel correlation is mapped into spectral coefficients that are approximately uncorrelated. The nearly uncorrelated spectral coefficients are then coded with the zigzag scan, Q and VLC to generate the bitstream. The bitstream is synthesized into a reconstruction of the spatial domain input data via the decoding modules in **Fig. 3**, which combines the inverse operations of the DCT, the zigzag scan, the Q, and VLC. The reconstructed data are stored and further used for removing temporal redundancy between the synthesized data and the immediately next input video data. With the reconstructed data, temporal redundancy can be removed by finding the best approximation under a specified distortion measure for each block of the current input data.

MPEG-4 decoder in **Fig. 3** consists of five modules including the variable length decoder (VLD), inverse zigzag scan, inverse quantization (Q^{-1}), inverse discrete cosine transform (IDCT) and motion compensation (MC).

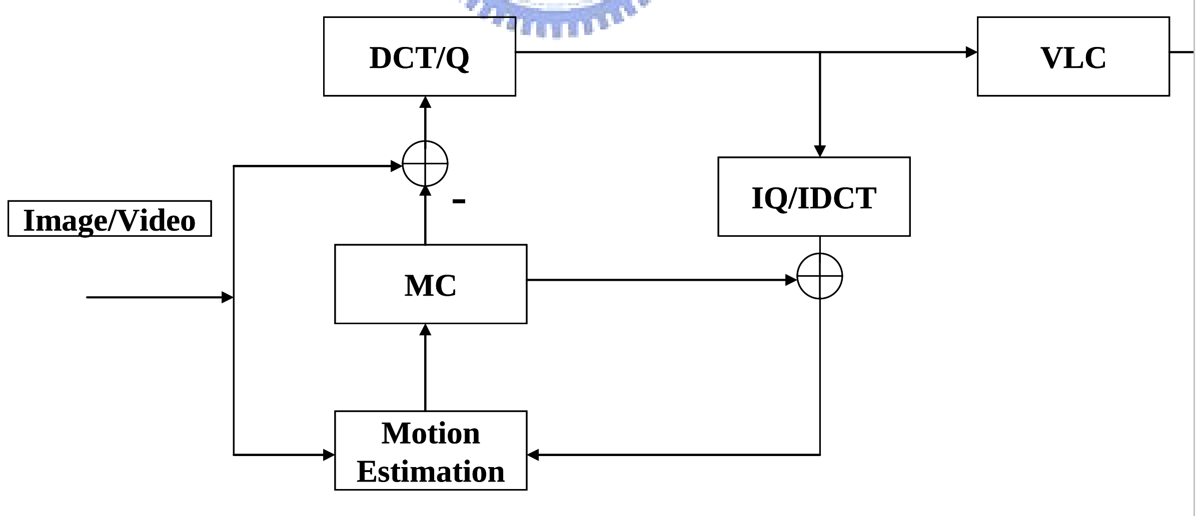


Fig. 2. Coding modules of an MPEG-4 Simple Profile encoder. The VLC indicates the variable length coding module.

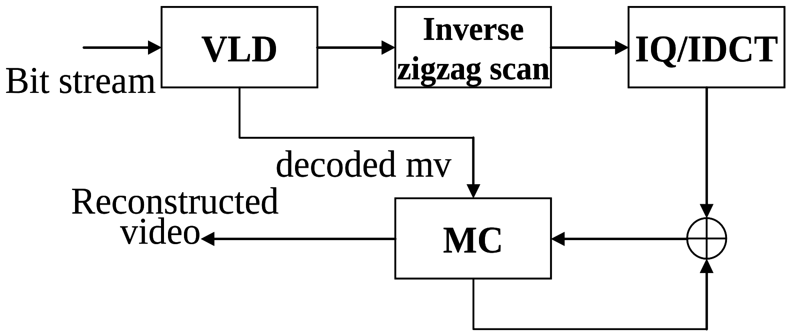


Fig. 3. Coding modules of an MPEG-4 Simple Profile decoder.

2.1.2 Overview of MPEG-4 Error Resilience

Error resilience techniques have been developed and applied to various multimedia players. The existing error resilience techniques can be separated into four major steps including the error detection, the error localization, the error recovery and the error concealment. The error detection process prevents the decoder system from failure during a decoding exception. The error detection process is critical for the subsequent steps. The error localization located the error and the error recovery process resumes the decoding process. A proper error concealment approach is used to amend the visual quality. Thus, with the error detection, the error localization, the error recovery and the error concealment, MPEG-4 video standards can be applied into the multimedia applications under the error prone environment.

2.1.2.1 Error Resilience Tools

In the MPEG-4 Simple Profile specification, the error resilience tools consist of re-synchronization marker (Resync_Marker), data partition and reversible variable length coding (RVLC). The error resilience tools can resolve the occurrence of errors. The Resync_Marker is used to stop error propagation, data partition is used to reduce the range of error corruption, and RVLC is used to decode partial information from the corrupted bitstream. As to the packet loss, the Resync_Marker can be used to detect errors and re-synchronize bitstream decoding. The three error-resilience tools provided by MPEG committee are summarized in the following paragraphs.

A. Resynchronization Markers

The resynchronization marker (Resync_Marker) placed at the start of a new video packet is distinguishable from all possible VLC code words. Header information of Resync_Marker is provided at the start of a video packet. The header contains the information necessary to restart the decoding process. The header information covers the address of the first MB contained in the current packet and the quantization parameter (QP) necessary to decode that first MB. The MB number provides the necessary spatial resynchronization and the QP allows the differential decoding process to be resynchronized. Following the QP is the header extension code (HEC). As the name implies, the HEC is a single bit to indicate whether additional video object plane (VOP) level information is available in the header. If the HEC is equal to one, the additional information (which is constant for each VOP and transmitted with the VOP header) is available in the packet header. The information means timing information, temporal reference, VOP prediction type and some other information. The header extension feature enables the decoder to correctly utilize data contained in the current packet without reference to the packet containing the VOP header, which can help error detection based on crosschecking capability. In MPEG-4 specification, all packets in the same VOP share the same QP, time stamp, etc.

When there is an error detected, we can seek for the next Resync_Marker and mark all MBs before the next Resync_Marker as errors, which can skip the corrupted MBs and enhance the crash proof.

B. Data Partitioning

Data partitioning can avoid discarding both texture and motion information of the MBs at the occurrence of any packet loss by separating the information into the motion and texture parts. The separator is called as motion markers as shown in **Fig. 4** to **Fig. 8**. Data partitioning puts all the motion vectors of MBs in the handling packet in front of a packet, inserts a motion marker and then puts the residuals of MBs after the motion markers. When one of the motion and texture parts is corrupted, the information of the correct part can be used for error recovery of the handling MBs. The motion information is more important than texture part in recovering the corrupted bitstream to a video sequence with satisfactory quality. When any error occurs in the texture residuals, we can use the correct motion vector to do the motion compensation and get the reconstructed MB. When an error occurs in the motion vector part, the correct residual may be used for error recovery. To model the effectiveness of using data partitioning, the authors in [26] have proposed 4 approaches for preventing errors.

(a) When an error is detected in the motion section, the decoder signals an error and losses all the MBs in the current packet

(b) When there is no error in the motion section and an error detected in the texture section, the decoded motion vectors are used to perform motion compensation.

(c) If no error is detected either in the motion section or the texture section but the Resync_Marker is not found at the end of decoding all the MBs in the current packet, an error is flagged and only the texture part is discarded.

(d) If no error exists in the motion part and the texture part, we will detect the occurrence of semantic errors. For example, when the Resync_Marker is found and the MB number is inconsistent, the number of MBs in the packet from the correct motion part is correct and can be use for advanced processing.

In addition, since the DC components are more important than the AC components in recovering the picture quality, the texture information is separated by a DC marker into DC and AC parts. With the DC markers, we can reduce the impact of random bit errors in transmission channels on the quality of the received video.

Resync. Marker	MB. No.	QP	Combined Motion and DCT data

Fig. 4. Traditional bit stream organization within the video packet [19].

COD	MCBPC	CBPY	DQUANT	Encoded MV(s)	DCT data	.....

Fig. 5. Bit stream components of each of the macro blocks within the video packet [19].

Resync. Marker	MB. No.	QP	Motion Data	MBM	DCT Data

Fig. 6. Bit stream organization with data partitioning for the motion and DCT data [19].

COD_1	$MCBPC_1$	Encoded $MV(s)_1$	COD_2	$MCBPC_2$	Encoded $MV(s)_2$	. . .

Fig. 7. Bit stream components of motion data [19].

$CBPY_1$	$DQUANT_1$	$CBPY_2$	$DQUANT_2$	. . .	DCT $Data_1$	DCT $Data_2$	. . .

Fig. 8: Bit stream components of DCT data [19].

C. Reversible variable length coding (RVLC)

RVLC is designed for inconsistent decoding from forward and backward direction as illustrated in **Fig. 9**. When any errors caused by decoding failure in a normal VLC decoding process, the backward VLD in RVLC decoding process is adopted. RVLC can enhance the decoding ability to obtain correct reconstructed values.

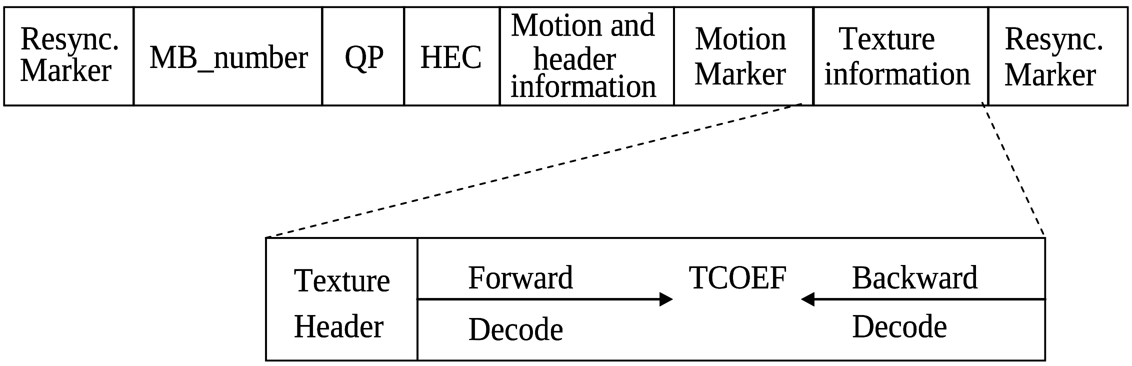


Fig. 9: Reversible variable length coding structure.

2.1.2.2 Error Detection and Recovery

The transmission errors are typically caused by random bit error or packet loss. The burst bit error can be categorized as the special case of the random error. The random bit error could be checked with syntax and semantic errors. As for the packet loss, we only need to resynchronize the decoding processes and continue the processing. The following will introduce the error detection and recovery algorithms for random bit error and packet loss, respectively.

(A) Packet loss

The packet loss is often detected as semantic errors. For examples, when nearby MBs are inconsistent, there is a high probability that packets are lost over transmission channels. There are two approaches to detect the packet loss. One is by checking the consistence of the MB numbers of the nearby blocks, which can be used to detect the packet loss for the texture part of the bitstream. The other is by checking if the neighboring MBs have the numbers increased monotonically. When the second type of the packet loss is detected, the resynchronization process is required to resume the decoding process. Seeking for the first MB of the first correctly decoded video packet can recover the synchronization.

(B) Random Bit Error

Bits error is occurred when there is a mismatch found during the decoding. The error is called as either syntactic or semantic error. The instances of the syntactic error consist of the VOP start code error, the video packet start code (Resync_Marker) error, and the VOP header error. The semantic error is detected when any of the following cases is found in decoding. marker_bit does not equal to the unity, the stuffing_bit has wrong length, more than 64 coefficients are in a decoded block, the number of MB in a VOP is larger than the maximum MB_number, illegal VLC symbols are parsed, the quantization scale is out of range, and the motion vector is out of search range, etc. With the error detection techniques, we can find the error and avoid the decoding failure. For those undetected errors, a crash proof decoder is necessary to prevent the crash.

In [8] , the authors also proposed a statistical method that can resume decoding processing.

The basic idea is that the error may be found previously, we can back-track the decoding process for each MB and decide if the processing MB is wrong based on the statistical information. The statistical information used comes from the instances that natural scene will be smooth in general, boundaries across a corrupted block will be irregular, and the statistical properties of a corrupted block is very different from the properties of the reliable neighboring blocks.

2.1.3 Summary

For better video service over the Internet or wireless channels, transmission error as bits error or packet loss shall be addressed. To retain the better error recovery, MPEG-4 video coding standard provides three error resilience tools. Use of resynchronization maker can deal with the bits error or packet loss by resuming the decoding process from next resynchronization marker recognized. With the resynchronization markers, data partitioning and RVLC can handle random bit error to avoid the discarding of correct motion or texture information and localize the positions of wrong data precisely. With error resilience tools provided by MPEG-4 specification, some error concealment can improve the error recovery to provide video content of better quality.

2.2 Overview of Intra Refreshment Techniques

2.2.1 Basic Idea

As the corrupted bitstream is received, only the information within the bitstream can be used to reconstruct the original video sequence. Once the error presented in the bitstream, the error will propagate in both spatial and temporal domain due to predictive coding. Inter block uses the spatial and temporal redundancy to get the most efficient way to compress the original video sequence. However, the error drift between blocks can not be stopped without an intra block that takes no previous information for coding. When we have to transmit the compressed bitstream over non-QoS network, we have to pay some coding inefficiency with better visual quality in the client end. Since intra blocks may lower the coding performance, we have to insert them in a suitable way to get better quality of the received video. Some methods choose intra blocks periodically in several frames, randomly in one frame, or according to the motion information. We will briefly introduce existing techniques.

2.2.2 Adaptive Intra Refreshment (AIR) in MPEG-4

Adaptive Intra Refreshment (AIR) is one of the intra refreshment methods for the error resilience. In the AIR, MBs in the motion area is encoded frequently in Intra mode. Therefore, it is possible to recover the corrupted MBs in the motion area rapidly.

AIR that is an informative technique is used for the rectangular VOP. It was designed in order to extract the MBs in motion area from the rectangular VOP at the encoder and encode the MBs in intra mode frequently. As the objects are extracted in advance in the arbitrary shape coding mode, it is possible to obtain the same performance by using the conventional cyclic intra refreshment within the arbitrary shape.

The number of intra MBs in a VOP is fixed and pre-determined. It depends on bitrates and frame rate and so on. In addition, the encoder estimates motion of each MB and the only MBs in motion area is encoded in Intra mode. The estimation values are recorded to the refreshment map MB by MB. The encoder refers to the refreshment map and decides to encode the current MB in Intra or Inter mode. The estimation of motion is performed by the comparison between SAD and SAD_th. SAD is the Sum of the Absolute Differential value between the current MB and the MB in same location of the previous VOP. The SAD has been already calculated in the Motion Estimation part. Therefore, additional calculation for the AIR is not needed. SAD_th is the threshold value. If the SAD of the current MB is larger than the SAD_th, the current MB is regarded as in motion area. Once a MB is regarded as in the motion area, the following MBs are regarded as in the motion area until the MBs are encoded in Intra mode by predetermined times. The predetermined value “1” as illustrated in **Fig. 10** is recorded to the refreshment map.

The horizontal scan is used to determine the MBs to be encoded in Intra mode within the moving area (*see* **Fig. 11**).

0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	1	0	0	0	0	0	0
0	0	1	1	1	1	0	0	0	0	0
0	0	0	1	1	1	1	0	1	0	0
0	0	0	1	1	1	1	1	1	1	0
0	0	0	0	0	1	1	1	1	1	0
0	0	0	0	0	0	1	1	1	0	0
0	0	0	0	0	0	0	0	0	0	0

Fig. 10. Refreshment map for a frame in QCIF resolution.

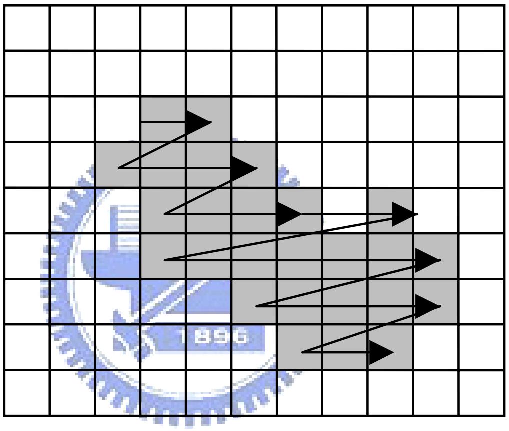


Fig. 11. Scanning order for the adaptive refreshment method [32]

2.2.3 Cyclic Intra Refreshment (CIR)

The mismatch noise is propagated and accumulated temporally, and the reconstructed image will suffer from catastrophic degradation eventually. In order to eliminate the accumulation of mismatch noise, the refreshment is essential to the self error resilient video coding. The refreshment can be easily achieved by forced intra coding, such as periodic intra frame, cyclic intra slices and cyclic intra MBs. The cyclic intra refreshment technique basically cyclically refreshes the whole image to avoid the error propagation for a long reference chain, which is done by intra coding a few MBs per frame.

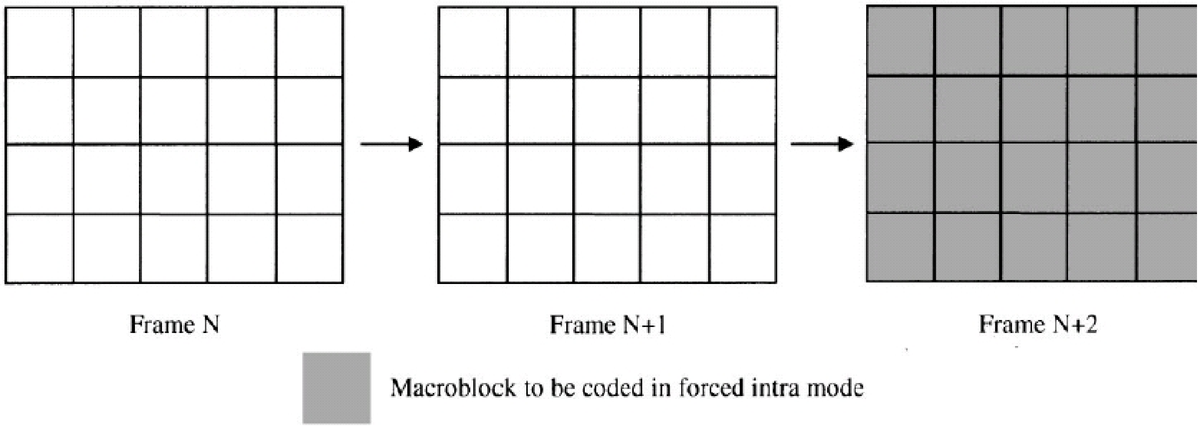


Fig. 12. Periodic intra frame

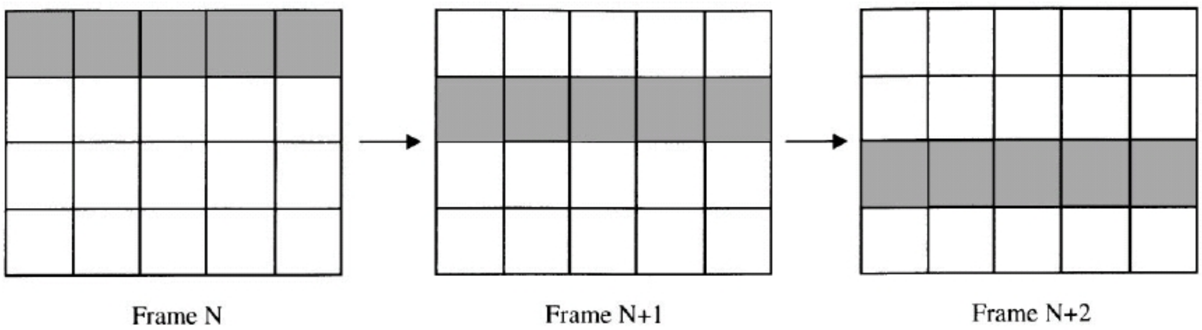


Fig. 13. Cyclic intra slices

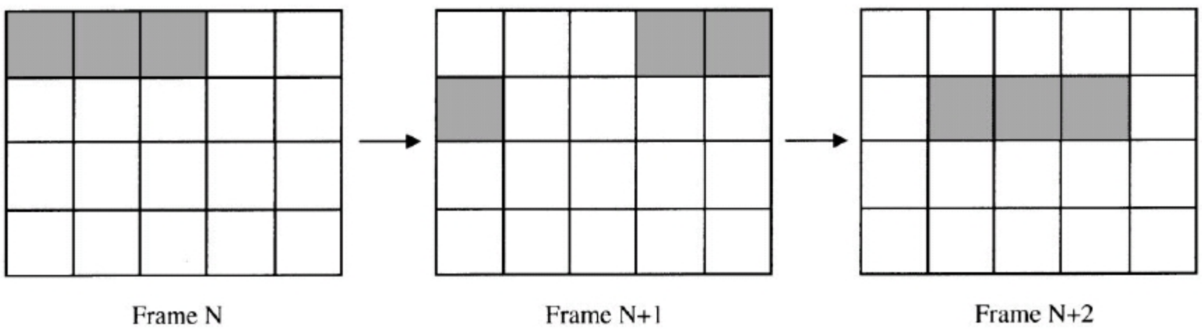


Fig. 14. Cyclic intra MBs

Periodically refreshing whole frame with all MBs in intra-coding is a straightforward idea to execute intra refreshment, but it costs too much bits and also causes large degradation in coding performance. The cyclic mechanism of intra refreshment is simple but not efficient because the decision of updated intra MBs is blind without considering any information.

2.2.4 Other Intra Refreshment Techniques

2.2.4.1 Random Intra Coding of Blocks

The intra refreshment that randomly updates intra coded blocks decides the portion of blocks to be intra updated in a coded frame based on the expectancy of the errors [39] . The expectancy depends on the intra block refreshment rate but is fairly independent of the probability p that a block is in error.

2.2.4.2 Periodic Refreshment

Some other methods in MB level have been proposed: "Scattered-Block Intra Update" (SB-IU), and "Contiguous-Block Intra Update (CB-IU)". The SB-IU method arbitrarily assigns MB's to $1/p$ groups and cycles through the groups updating one group per frame. The intra-updating frequency for each MB is $1/p$. The similar intra updating approach was proposed in [36] . In the CB-IU method, contiguous-block patterns are recommended depending on the packet loss rate. Specifically, we use sizes of 2x2 for p = 5%, 3×3 for p = 10%, 4×4 for p = 15%, 5×5 for p = 20%. "Block-weighted distortion estimate" (BWDE) estimates the decoder distortion via the formula $\hat{D} = pD_c + (1-p)D_q$, which means that for each block of the previous frame, we compute the distortion that is from lost and concealed MBs and quantization distortion of block residuals. Concealment error of a block in the current frame is defined as the weighted average of the concealment distortion of blocks in the previous frame. In addition, the concealment error calculation considers the motion compensation of blocks is considered and the weights correspond to their relative coverage of the block area. D_q indicates only the quantization distortion of the residuals at the current MB. Note that the approach assumes that the current block is received correctly and considers two cases based on status of previous frame blocks.

2.2.5 Summary

The intra refreshment techniques mentioned are simple and straightforward. The AIR method considers the motion area as the important part that needs to be intra refreshed frequently. The gain obtained from intra refreshment is limited since the number of intra MBs in a frame is fixed. The CIR is an inefficient way to intra refresh sequence especially when we use periodic intra frame because the overhead of intra MBs is added without considering any coding information. The CIR pays more overhead but loses more coding efficiency. The RIR uses the simplest way to implement intra refreshment but the improvement is also limited. We need to develop more efficient ways to improve performance and retain the R-D performance at the same time.

2.3 Error Concealment Methods

After detecting and localizing the lost MBs in a frame, we can use concealment methods to recover the damaged frame to a proper quality. With error concealment, the error in the current frame will not be propagated to the subsequent frames or MBs via the spatial and temporal prediction loops defined in the video compression standard.

Error concealment techniques can be basically classified into three categories including forward error concealment, error concealment by post-processing and interactive error concealment. The forward error concealment refers to the techniques in which the encoder adds redundancy to the bitstream to enhance the error resilience at the decoding process. The existing forward error concealment includes adaptive intra refresh [23] or forward error control (FEC) coding [27] . Error concealment by post-processing refers to decoding operations that can deal with the MBs recognized as being corrupted under the transmission channels [1] - [20] . Error concealment by post-processing is a major research direction. In interactive error concealment, the transmitter and receiver cooperate to minimize the effect of transmission errors. The interactive methods heavily utilize feedback information provided by the receivers. Error concealment by post-processing can also be referred to as passive error concealment whereas the other two categories represent forms of active error concealment.

In addition, based on the processing domain, error concealment methods can be classified into spatial methods and temporal methods. Some hybrid approaches that combine both methods adaptively to fit video characteristics can get better performance. Based on considerations on realization complexity and computation cost, error concealment methods that are simple and can provide proper reconstruction quality are demanded in the decoder for practical usage.

2.3.1 Spatial Error Concealment

Spatial error concealment (SEC) techniques exploit the spatial redundancy within a picture. SEC techniques are mainly devoted to conceal I-pictures that have no motion information to be used for error concealment. For example, SEC techniques are applied to the first I-frame of a sequence or I-frames after scene changes. Considering the spatial correlation, the SEC approaches can provide better reconstructed video quality for the smooth or regular image blocks. For the irregular image blocks, SEC will produce the blurred image blocks, which is undesirable for subjective quality. The following paragraphs will introduce various SEC approaches and the impact on the MPEG-4 video.

2.3.1.1 Weighted Interpolation

Weighted interpolation interpolates each pixel of the entire $2N*2N$ erroneous MBs with the adjacent pixels of the four adjacent MBs [6] . The interpolation operation is defined by

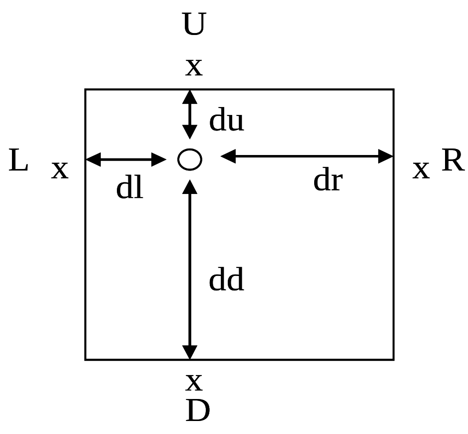


Fig. 15. Illustration of the weighted interpolation.

$$value = \frac{U * dd + D * du + L * dr + R * dl}{dd + du + dr + dl}$$

where U, D, L and R stand for the nearest correct pixels from the pixel that is concealed in the directions from the top, down, left and right directions respectively. The du, dd, dl, and dr are the distances between the current pixel and the pixel in the associated direction respectively.

When some of the MBs do not exist during interpolation, the distances at the related interpolation directions will be set to zero. In addition, weighted interpolation techniques have the best performance when all surrounding MBs exist. For existing video coding standards covering H.263 and MPEG-4, the casual coding scheme provides only the top and the left MBs for weighted interpolation.

Usually, with this technique, no abrupt transition will occur in an MB. However, the details and edges of the MB become blurred, which may result in visible blocking artifact between the adjacent blocks.

2.3.1.2 Quadrilinear Border Interpolation

In [4] the authors interpolate the value of a missing pixel from the closest top, left, bottom and right closest pixels surrounding the handling pixel. Quadrilinear border interpolation is a little different from general methods that use MB based interpolation. In addition, to improve the quality, the methods need some refinements. Some modify the weights for the interpolation, which increases weighting to the pixel at the direction that has less distance. Some uses the average values of the nearest pixels and the two neighbors instead of the nearest pixels for the interpolation.

2.3.1.3 Other Interpolation Methods

To preserve the detail of the missing blocks, some other interpolation methods covering multi-direction interpolation [17] and adaptive recursive interpolation algorithm [25] have been proposed. To apply the detail preservation methods to the MPEG-4 video, some refinement to fit the casual coding procedure is required. For example, multi-directional filtering shall be reduced to the limited directions, which decreases the interpolation efficiency in reproducing the spatial details for error recovery.

2.3.2 Temporal Error Concealment

Temporal error concealment techniques exploit the temporal similarity of frames in a sequence. Temporal approaches are usually applied to P-pictures and B-pictures that have motion vectors. For a slow motion sequence, the motion information for the collocated MBs at two adjacent pictures is highly correlated, which indicates the motion vectors of the previous reference picture can also be used at concealment of the immediately following I-pictures. In addition, inconsistent motion vectors of the corrupted MB and the surrounding MBs may decrease the concealment efficiency of the temporal error concealment methods. The temporal algorithms are summarized in the following.

2.3.2.1 Direct Copying

When the current blocks are lost, the simplest way of error concealment is by copying the block data at the same positions of the reference frame [6] . The direct copy method needs no additional memory and suffers from visible shifts in the reconstructed picture when motion is present.

2.3.2.2 Motion Compensated Temporal Error Concealment (MC-TEC)

To improve the visual quality of direct copy approach, MC-TEC considers the motion information for error concealment. Based on the utilization of temporal correlation to compress the video sequence, motion vectors are introduced in video compression standard. For the lost MBs, we can assume that the motion information of the neighboring MBs has a close relationship. Based the relationship, we can use some motion vectors in the motion compensation to rebuild the lost MBs. The remaining issue is to derive the motion vectors of lost MBs. We can simply derive motion vectors of handling MB as the mean or median value of the motion vectors from the surrounding MBs [6] . Another way to choose a motion vector is to select a motion vector in a set of candidate motion vectors that have the minimum mean square error (MMSE) of the border pixels [6] .

In [4] the authors have introduced two temporal concealment methods covering concealment by selection and concealment by search. Concealment by selection rebuilds the lost MBs based on the NEighbors MAtching criterion (NEMA). For each candidate motion vector, an error function is evaluated between the surrounding pixels and the motion compensated pixels of the extended area. Concealment by search is also called as Dynamic Motion Vector Estimation (DMVE). Each candidate motion vector is refined by examining all motion vectors in a specified search window surrounding the current MB under a NEMA-like criterion. DMVE can retain better recovered video with much more complexity as compared to the performance of the concealment by selection.

2.3.3 Hybrid Error Concealment

For better error concealment, a straightforward concept is to adapt the best concealment technique based on video properties, which is called as hybrid error concealment. At frame level, we can use spatial error concealment for I-frames and temporal error concealment for P-frames and B-frames. At MB level, spatial and temporal concealment are adopted based on the properties of the handling frame. The remaining issue of realizing optimized hybrid error concealment schemes is to derive the criterion of choosing the best concealment technique for each MB.

The selection criteria of the best concealment method are based on the measurement of the image activities including spatial activities and temporal activities.

In [6] , the authors have provided that the amount of the temporal activities is derived from the mean square error (MSE) between two MBs of adjacent pictures. The spatial activities can be computed by measuring the variance (VAR) of the nearest neighboring MB (top or bottom) of the current picture. When the amount of temporal activities in MSE is smaller than a certain threshold T1=VAR, the temporal concealment is applied. Otherwise, the spatial concealment is used.

In [24] , the authors have proposed another decision method based on prediction error statistics measured in the neighborhood. The decision region is shown in **Fig. 16** by

$$VAR = E[(X - \overline{X})^2], VAROR = E[X^2] - \mu^2,$$

where X is the neighboring good MB data, $\overline{X}$ is the data of the corresponding MB in the previously decoded frame at the co-located position and μ is the average value of the neighboring good MB in the current frame. One can appreciate that VAR may indicate the local motion and VAROR can present the local spatial details. If VAR > VAROR, the better concealment method is spatial interpolation. When VAR ≤ VAROR or VAR ≤ T, where T is a preset threshold, the better concealment method is temporal replacement.

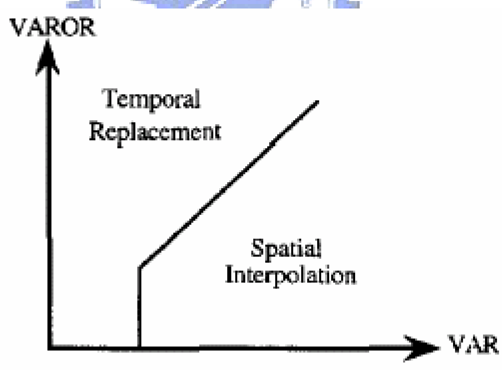


Fig. 16. Adaptive error concealment strategy

In [8] , the authors have given a set of concealed block candidates for the corrupted block and proposed a fitness function for error concealment to select the “best” concealed block candidate among the set. The set includes several types of concealed block candidates covering neighboring blocks, average or median of neighboring blocks, motion compensation blocks with motion vectors from neighboring blocks, average or median of motion compensation blocks with motion vectors from neighboring blocks…etc. To select the best concealed block, some parameters are calculated based on the following information consisting of smooth boundary, similar statistical properties, and temporal similarity with the motion compensated block in the previous picture. Finally, with the parameters, a fitness function is used to decide the best candidate for concealing each block.

In [9] the authors have adopted boundary matching errors (BME) for various combinations of MBs. The MB combinations cover the top and top-left, the top and top-right, the top and bottom-left, top and bottom-right, the temporal concealed and top, and the temporal concealed and bottom. With MBE values, the smoothness and similar property between concealed MB and neighboring MBs can be verified.

In [4] some strategies using temporal concealment in I-VOP or spatial concealment in P-VOP are discussed. At the scene change or fading pictures, the spatial redundancy may be more than the temporal redundancy. The majority of MBs in I-VOPs are referred as INTRAs. When the INTRAs have some temporal redundancy between the temporally successive VOPs, temporal methods will achieve better visual quality than the spatial algorithms.

2.3.4 Summary

The general error concealment methods are divided into two types including spatial error concealment and temporal error concealment. Spatial methods based on spatial correlation recover the picture quality. For spatial concealment, interpolation is simple, easy to implement, and has good performance. Temporal concealment using temporal correlation recovers the visual quality. The basic methods for temporal concealment use some candidate motion vectors and some cost functions to decide a best motion vector. In addition, to have a better performance based on the video characteristics, some hybrid methods are investigated. To further avoid the blocking artifacts, the smoothness of the concealed video frames is judged with boundary matching or statistical similarity.

2.4 Overview of MPEG-21 Standard

MPEG-21 is an open standards-based framework for multimedia delivery and consumption [28] ,[29] . The MPEG-21 vision is to define a multimedia framework to enable transparent and augmented use of multimedia resources across a wide range of networks and devices used by the different communities. The concerns of MPEG-21 are: management of content, repurposing content based on user preferences and device capabilities, protection of rights, protection from unauthorized access/modification, protection of privacy of providers and consumers, and so on. Universal Multimedia Access (UMA) deals with delivery of audio, images, video, and multimedia content under different network conditions, user and publisher preferences, and capabilities of terminal devices. The primary motivation of UMA is to enable terminals with limited communication, processing, storage and display capabilities to access rich multimedia content.

The basic concepts in MPEG-21 relate to what and who within the multimedia framework. What is a Digital Item that's a structured digital object with a standard representation, identification, and metadata within the MPEG-21 framework. The who is a user who interacts in the MPEG-21 environment or uses a Digital Item, including individuals, consumers, communities, organizations, corporations, consortia, governments and other standards bodies, and initiatives around the world. The user roles include creators, consumers, rights holders, content providers, distributors, and so on—there's no technical distinction between providers and consumers. All parties that must interact within MPEG-21 are categorized equally as users. They assume specific rights and responsibilities according to their interaction with other users. All users must also express and manage their interests in Digital Items.

In practice, a Digital Item is a combination of resources, metadata, and structure. The resources are the individual assets or (distributed) content. The metadata describes (distributed) data about or pertaining to the Digital Item as a whole or also to the individual resources in the Digital Item. Finally, the structure relates to the relationships among the parts of the Digital Item—both resources and metadata. An example of a Digital Item might be a presentation of a university, including photos, videos, animation graphics, textual information, news related to the university's research activities, e-learning material, navigational information driven by user preferences, and so on. The Digital Item is thus the fundamental unit for distribution and transaction within the MPEG-21 framework.

Although electronic content creation, distribution, consumption, and trade are already possible today, no current end-to-end solutions let different user communities interact in an interoperable and efficient way. MPEG-21 guarantees such interoperability by focusing on how the elements of a multimedia application infrastructure should relate, integrate, and interact. The strength of MPEG-21 is that it's a standard grown from clear multimedia industry requirements for interoperability. As such, the requirements, existing standards, and current work are firmly based on usage cases and scenarios provided by that industry.

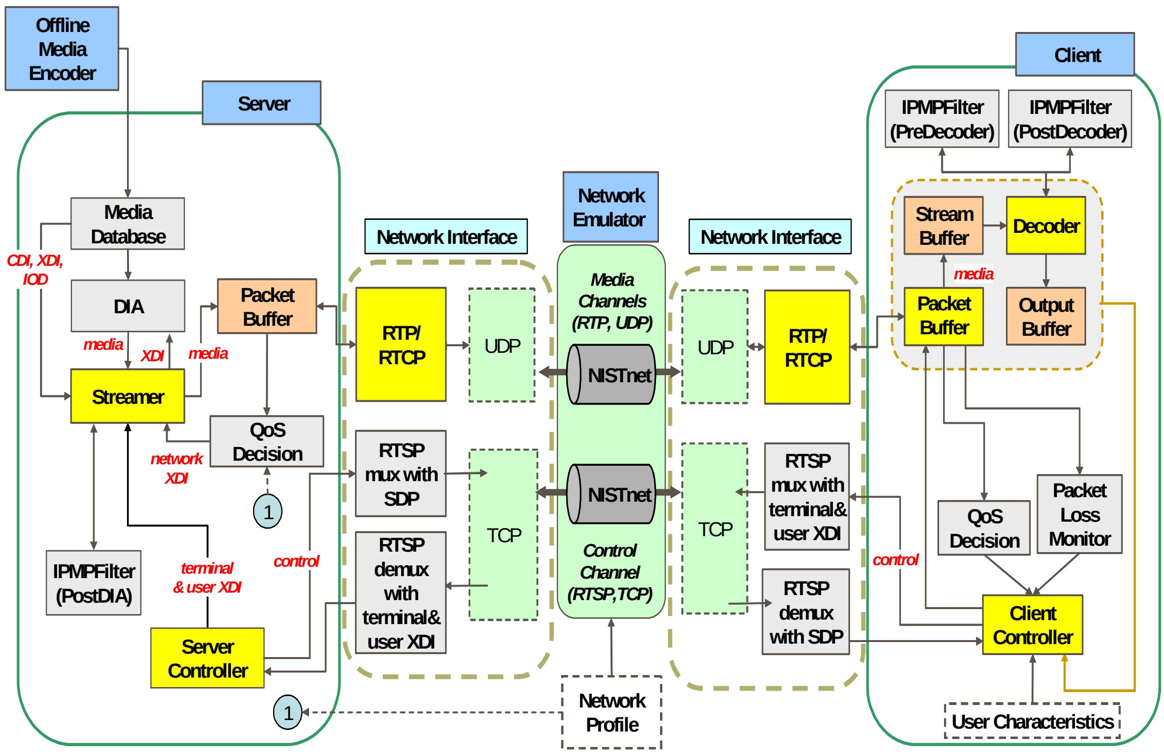


Fig. 17. Architecture of the proposed FGS-based video streaming test bed [31] .

2.4.1 MPEG-21 Part-12 Multimedia Test Bed for Resource Delivery

The goal of FGS-based streaming test bed is to support MPEG-21 DIA scheme with a more strict evaluation methodology according to the specified common conditions for scalable coding. **Fig. 17** shows the NCTU test bed system architecture, which covers four key modules including the FGS-based Video Content Server, Video Clients, Network Interface, and Network Simulator. The detailed description of each module can be found at [31] .

2.4.2 Summary

MPEG-21 gives a provision to realize a digital life with various multimedia data by network or communications to deliver multimedia everywhere anytime. Digital item adaptation makes the digital items adaptive to the various service conditions covering terminal capability, network condition, user preference, etc. The multimedia test bed provides a platform to simulate real network conditions for video content delivery. With the digital item adaptation, multimedia test bed, and other parts of MPEG-21, we could realize universal multimedia access environment in a near future.

2.5 Remarks

We have described existing works on constructing a streaming video system with error resilience techniques. To get better quality when packets are lost, error concealment algorithms and error robustness schemes are employed. The streaming system is integrated into a multimedia test bed that can simulate a variety of network conditions and evaluate error resilience performance for on-line testing.

Chapter 3

Rate-Distortion Optimized Intra Refreshment

The proposed Rate-Distortion Intra Refreshment (RDIR) framework is realized based on MPEG-4 reference encoder. Within the RDIR framework, we will introduce the concept and rationale of intra refreshment tools including the R-D cost, the historical record, the systematic insertion of intra MBs and the run length of successive intra MBs.

3.1 Encoder with Intra Refreshment

Fig. 18 shows the whole architecture of MPEG-4 simple profile encoder with intra refreshment mechanism RDIR, which considers the mode decision of motion estimation (ME). The module of mode modification will collect information from R-D cost, history record, and intra run to choose a better mode in transmitting video content over the channels with packet loss. The R-D cost module calculates the Lagragian cost of inter and intra mode based on the information from two more encoding processes, which mean inter coding and intra coding. The derivation of the R-D cost is based on real bits and distortion. Two other coding processes that are associated with intra and inter modes for the whole frame are executed before the R-D cost calculation. During the third coding process, conditions about history record and intra run will also be checked. History record will help stop the long reference chain and the deterioration of error drift. Intra run will help distribute the overhead to different frames in the whole sequence, which can improve concealment efficiency since the intra MBs contain no motion information that is used in the temporal concealment process.

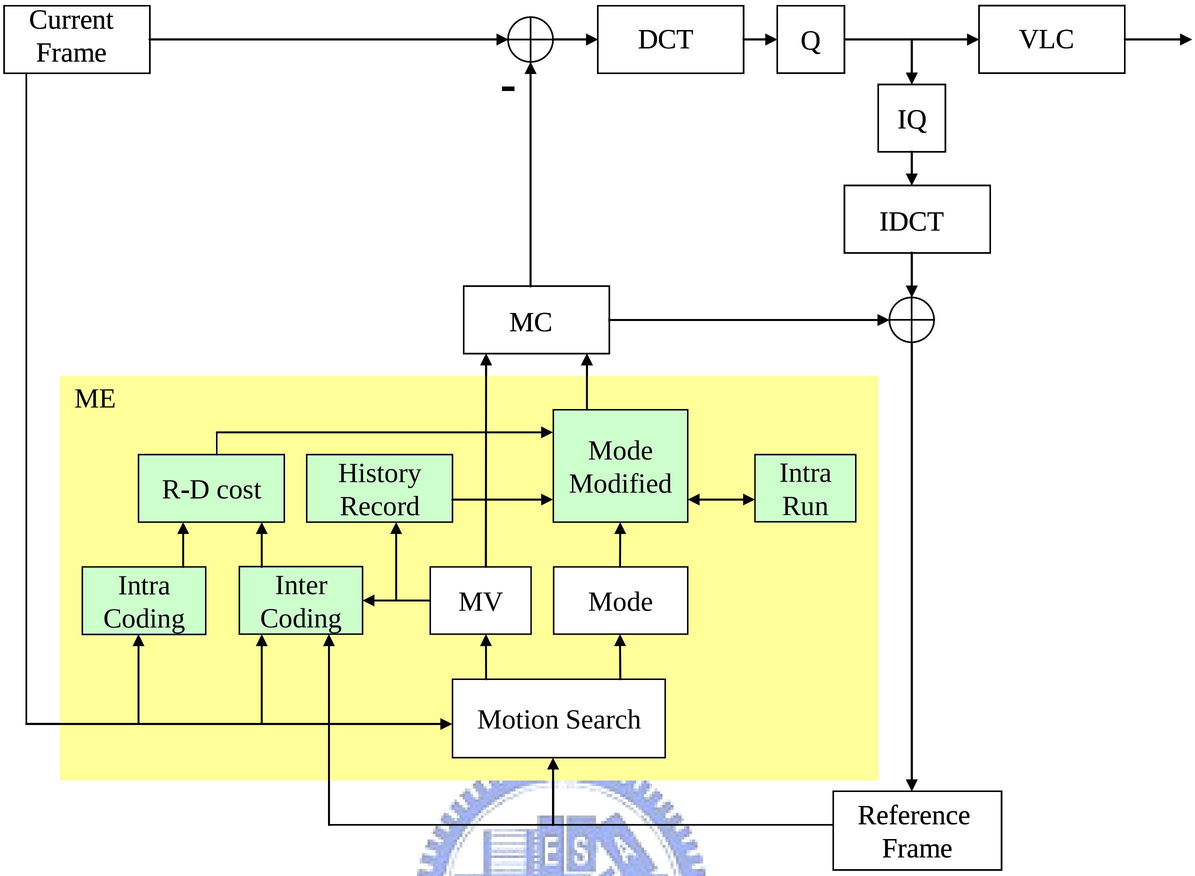


Fig. 18. Architecture of encoder with intra refreshment

3.2 Mode Decision for Intra Refreshment

Considering packet loss during transmission, our coding system makes intra refreshment mode decision based on the error location. As shown in **Table 1**, there are four conditions when encoding each pair of frames. To have a better mode decision by inserting a proper number of intra MBs to the bitstream of the processing frame, we simply the mode decision by assuming that only the previous frame has packet lost MBs and the current frame has no packet loss. Under the assumption, both types of real distortion including the quantization error and the concealed error from the previous frame are used for refining the mode decision for vide delivery over the packet-switching network.

Table 1. The four conditions for the mode decision of intra refreshment.

Current frame / Previous frame	Packet loss	No packet loss
Packet loss	**X**	**V**
No packet loss	**X**	**X**

3.3 Algorithm Descriptions

The intra refreshment technique is to insert intra coding block instead of inter coding block in P frames to prevent serious error propagation caused by packet loss due to transmission over error-prone network. Since the intra coding block uses more bits, the coding efficiency is poor when the packet loss rates decrease. To balance the coding efficiency and error robustness, intra block insertion with rate-distortion optimization adaptive to channel conditions can reduce the total amount of overhead and retain the coding performance at the same time.

The RDIR can enhance the robustness of the encoding bitstreams based on the following approaches, which indicate the use of history record, the inclusion of motion and texture information for record update, the coding bitrate of previous frame, and the successive number of intra MBs.

3.3.1 Rate-Distortion Model

Rate-distortion (R-D) model has been proven to effectively fit mode decision in coding system for error-free channel [33] In error-prone environments, the mode selection based on an R-D framework to convey packet loss has been proposed in [25] .

We add a R-D model in our coding system that derives the distortion with both concealment error and quantization error. Inter and intra modes are used in our mode decision framework. The intra mode represents coding without temporal prediction and the inter mode uses motion compensated temporal prediction. For each MB, we collect the information about the coding bits and quantization distortion as for inter and intra coding individually. For each MB, we calculate the cost for intra and inter blocks, which are indicated as J_{intra} and J_{inter} respectively, by a Lagrangian formula.

$$J = D_q + \lambda \cdot R .$$

Where J indicates the Lagrangian cost. For each block, the cost $J_{intra} = J$ is for intra coding mode and $J_{inter} = J$ is for inter coding mode. λ is the parameter used to control coding bit rate in encoding process. D_q means distortion induced from residual quantization and R presents bits used in coding a MB.

After the cost *J* is decided, the mode with the minimal value of *J* is chosen as the current MB coding mode. For error-prone environment, the distortion D is increased by lost packets. When any error occurs, temporal prediction will allow the error to propagate if the inter mode is chosen. Using the intra coding mode will stop error propagation and decrease the coding efficiency simultaneously. To achieve the R-D optimization under the proposed intra refreshment encoding, the parameter of λ needs to be updated every frame to control the used bits under the same distortion by

$$\lambda_{n+1} = \lambda_n \left(1 + \alpha \left(\sum_{i=1}^{n} R_i - nR_{target} \right) \right), \quad \alpha = \frac{1}{20 \cdot R_{target}}$$

Where the parameter α is decided with massive simulations of buffer control. The packet loss rate is used to model the Internet protocol. Using network conditions to model the situation in decoder end is expected to reconstruct better image quality. As the modeling is 100% matched, we will get the same quality as transmitted one in error prone environment.

In our proposed RDIR, for each MB, we compute the cost for intra and inter blocks, J_{intra} and J_{inter} by the Lagrangian formula $J = D + \lambda \cdot R$ where J is the Lagrangian cost. λ is the parameter used to control coding bit rate in encoding process. The D means the distortion induced from residue quantization and R equals to the number of bits used in coding a MB. We take the different D for each mode. For the inter mode, we have to take concealment distortion due to a concealment block with variable pack-loss rate by

$$D = D_q \times (1 - p) + D_c \times p\,,$$

where D_q is the quantization and D_c is the concealment error from previous frame; for intra mode, there is something difference in the distortion, $D = D_q$. Since intra mode doesn't refer to the previous frame and concealment error is out of consideration. The parameter p is the estimated channel packet loss rate, which is 10 % here.

3.3.1.1 Use of Error Concealed Data

The concealment error will be calculated when we get the motion information. We will reconstruct one 'concealed' frame by assuming one MB lost each time and doing frame reconstruction with the same concealment strategy in decoder end. The concealment error will be derived by comparing the 'concealed' frame with the original reconstructed frame.

After the cost of *J* is decided, the mode with minimal value of *J* is chosen as the current MB coding mode. For error prone environment, the distortion of D will suffer more serious quality loss. The quality loss comes from the original quantization error and the errors introduced when concealing the lost MB from nearby MBs. To consider the error concealed distortion, the metric *J* is quantified by

$$J = \left(D_q \cdot (1 - p) + D_c \cdot p\right) + \lambda \cdot R\,.$$

Where D_q presents the distortion induced from residual quantization and D_c is distortion induced from a specified concealment algorithm. p indicates the packet loss rate.

3.3.2 Historical Record

The motion information is used for selecting better coding modes after R-D cost calculation. As in [38] , pixel level inter-frame dependence history has been adopted to enhance the intra refreshment. We take some record about reference times as another intra refreshment criterion for the MBs that are assigned as inter-mode by R-D decision.

Initially, we set the value of all MB's reference history chains by 0. The setting is starting from the 1st P-frame and the record values will be updated for each subsequent frame under two major considerations, which include the weighting summation of the record values of previous frame and the use of current motion and coding information. The update is continued until the last coding frame.

The record will be updated while encoding each frame. If the mode of MB is intra, corresponding record will be refreshed to 0; otherwise, we will use motion vector to count the record with average weighting in the reference frame by

$$HR_{new} = \frac{A_1 * HR_1 + A_2 * HR_2 + A_3 * HR_3 + A_4 * HR_4}{A_1 + A_2 + A_3 + A_4}$$

where A_1-A_4 are areas of each MB in the previous frame which are referenced by current MB in the current frame, and HR_1-HR_4 are records of each referenced MB. Intra mode will be used if the record has shown that current MB is in a long reference chain.

We take the moving-intensive area as a specific condition that we shall refresh more frequently because there will be more severe concealment distortion when the area is lost. The ratio of motion vector amplitude (mv) over the specified search range (SR) will be added into the calculation of current MB's history record.

$$HR'_{new} = \frac{A_1 * HR_1 + A_2 * HR_2 + A_3 * HR_3 + A_4 * HR_4}{A_1 + A_2 + A_3 + A_4} + \frac{mv}{SR} = HR_{new} + \frac{mv}{SR}$$

We will reset the record to 0 once the processing MB is selected to be intra coded. As the record of each MB is larger than a specified threshold and the current MB is set as inter coded by the mode decision module, the processing MB is changed to intra coded. The threshold for the better performance of intra refreshment is found empirically.

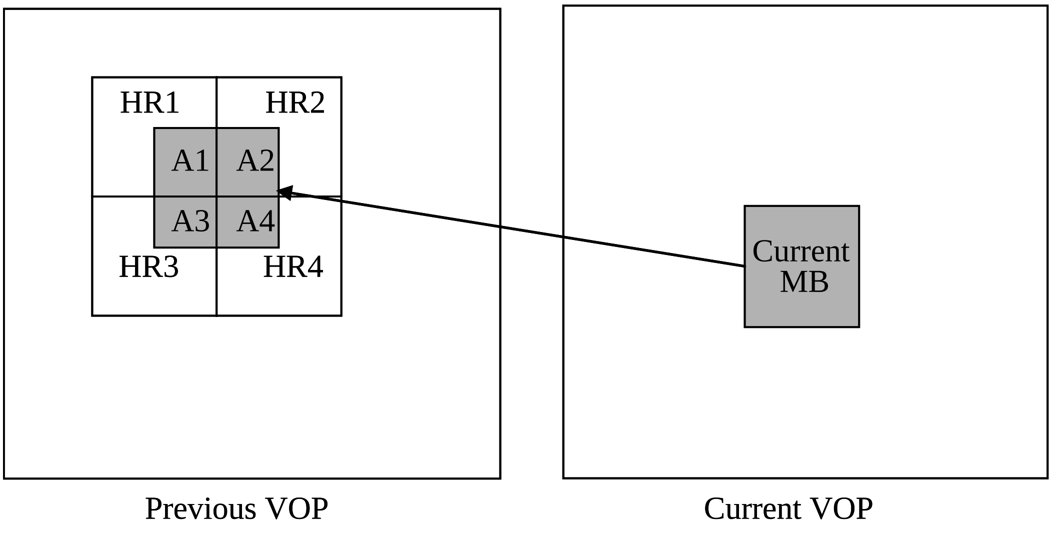


Fig. 19. History record update.

3.3.3 Uniform Allocation of Overhead Bits

The bits used for coding previous frames are considered to spread the overhead bits of intra refreshed MBs more uniformly for the handling sequence. To alleviate the increase of the overall bit budget, we multiply the history record value of the current frame with a fractional number when the used bits exceed a specified threshold. The fractional value that is set as a floating number less than the unity is to control the number of spending bits. In our simulations, the record values are recalculated by

$$HR'_n = \begin{cases} \alpha \cdot HR_n, & \text{if } R_{n-1} > R_T + \Delta R \\ HR_n, & otherwise \end{cases},$$

where HR_n indicates the original history record and HR'_n is the reduced history record. R_{n-1} means the total bits used for coding previous (*n-1*) frames. R_T is the target bit rate in bits/sec and ΔR is used to control the increase of overhead bits from the insertion of more intra MBs. The floating number α is set as 0.9 in our simulations.

3.3.4 Uniform Location of Successive Intra MBs

For streaming over the Internet and wireless channels, the video bitstreams are packed into packets. The packet loss may occur when transmitting via the channels, which may drop successive intra MBs. To avoid quality degradation of received video by the burst packet loss, when successive number of intra MBs are detected, the intra refreshment will not change the inter MBs to the intra MBs. The record value is retained for the following update. It will help distribute the overall overhead to different frames and improve error concealment. The successive intra MBs that contain no motion information will degrade the error concealment performance especially in temporal concealment.

3.3.5 Application to MPEG-4 Video Reference Software

In RDIR, we incorporate the RD-framework, concealment error, and the historical record into the mode decision process of the coding standard. In addition, we consider packet loss rate for deriving the distortion of the inter coding MBs. R-D framework with concealment error will be more efficient in mode decision. We try to stop the error propagation caused by packet loss since the concealment error is usually larger than the quantization distortion. In enhancing the quality of decoded video, historical record is recommended. The long series of P-frame will be stopped once the historical record is larger than the threshold. As in **Fig. 20,** the RDIR will be performed in each MB per frame. The MB based RDIR can fit to the video content delivery over an error-prone transmission channels.

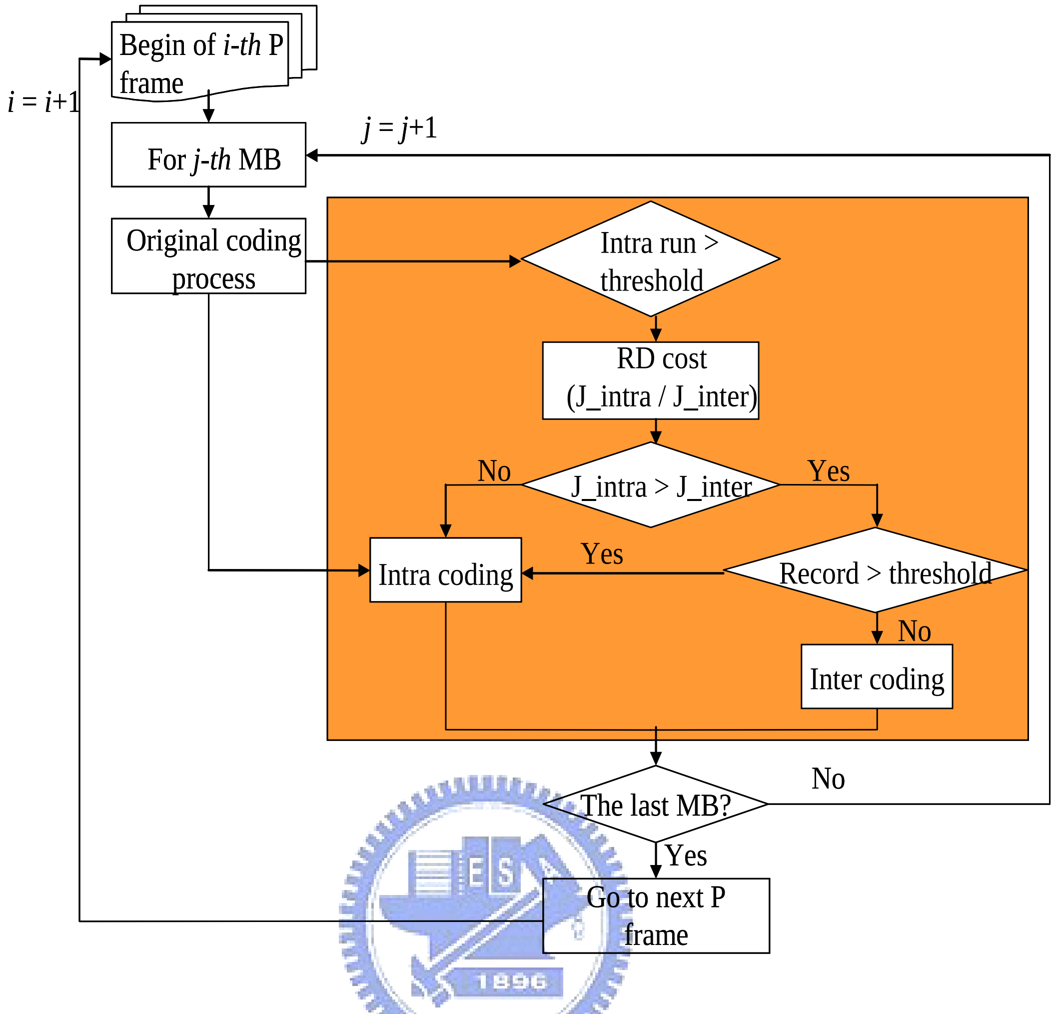


Fig. 20. The flow chart of RDIR encoding

3.4 Summary

RDIR inserts intra MBs based on video characteristics, properties of packet-switched streaming schemes and network conditions, which can further improves the visual quality and balances the coding efficiency with error robustness of bitstream.

Chapter 4

Experimental Results

Experimental results using the proposed RDIR algorithm under various bit rates and different network conditions are provided. We will perform both off-line and on-line testing. As for off-line testing, a decoding process with the hybrid concealment algorithms is used to investigate the overall performance of an error resilience system. For on-line testing, retransmission scheme in the steaming system is taken into consideration for providing the better quality of service.

4.1 Off-line Testing

4.1.1 Experimental environment

We use MPEG-4 Simple Profile reference software of MoMuSys version to realize the RDIR algorithm. We analyze the effect of the RDIR algorithm with different test sequences of slow motion, fast motion, and D1 (720x480) resolution. For simulations, the encoding frame rate is 30 frames per second (fps), decoding frame rate is 10 ftp, GOP structure is I-P-P…, and packet size is 1000 bits. To off-line simulate packet loss conditions and observe the impact of packet loss rates on the visual quality of reconstructed video sequences, we randomly drop the MBs to simulate different packet loss rates. The dropping probability is in a uniform distribution. We average ten simulation results to get the average performance because different lost packets in the bitstream will cause different quality degradation.

4.1.2 RDIR performance

We evaluate the RDIR performance with different tools and compare with the encoding schemes without intra refreshment. The test sequence is Foreman with CIF (352x288) resolution under 10% packet loss rate. The concealment algorithm adopted in the decoding loop of the encoder is spatial copy in I-VOPs and zero motion in P-VOPs.

Table 2. PSNR improvement of different tools at different bitrate

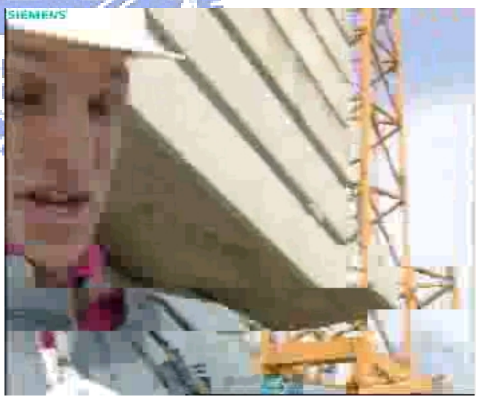
	Error Free	NoIntra	RD	RD+HR	RD+HR+MV	Total gain
128k	30.19	19.25	21.79	24.47	24.78	
Gain		--	2.54	2.68	0.31	5.53
256k	33.58	19.16	22.26	25.43	25.63	
Gain		--	3.1	3.17	0.2	6.47
512k	36.82	19.03	22.44	26.02	26.20	
Gain		--	3.41	3.58	0.18	7.17

The column ‘Error Free’ shows the original bitstream without intra refreshment under error-free environment and the column ‘NoIntra’ presents the original bitstream transmitted over the channels with packet loss. The next three columns indicate the RDIR algorithms that include the different additional tools are used for simulations. As for different bitrates, the gain as compared with the ‘NoIntra’ condition is about 5.53 ~ 7.17 dB. The visual quality of the 61st frame with different tools at 128k bits per second (bps) is shown in **Fig. 21**.

NoIntra

RD

RD+HR



RD + HR + mv

Fig. 21. Visual quality improvement of RDIR

As for slow motion video sequences, Akiyo sequence is used for simulations at different bitrates and packet loss rates. We can get 0.88~6.32 dB gain in PSNR.

Table 3. PSNR improvement of slow motion simulation

Akiyo									
bitrate(bits/sec)	**50k**			**100k**			**200k**		
	Proposed	**NoIntra**	**Gain**	**Proposed**	**NoIntra**	**Gain**	**Proposed**	**NoIntra**	**Gain**
Error Free	**37.63**	**--**	**--**	**40.34**	**--**	**--**	**42.76**	**--**	**--**
PLR 1%	**36.05**	**35.17**	**0.88**	**37.22**	**36.15**	**1.07**	**38.96**	**36.85**	**2.11**
PLR 5%	**31.75**	**27.16**	**4.59**	**33.30**	**29.82**	**3.48**	**33.60**	**29.18**	**6.32**
PLR 10%	**28.61**	**23.15**	**5.46**	**29.52**	**24.22**	**5.3**	**30.87**	**24.88**	**5.99**

As for fast motion video sequences, Foreman sequence is used for simulations at different bitrates and packet loss rates. The RDIR has improved the PSNR of the ‘No intra’ scheme by 1.21~6.95 dB gain.

Table 4. PSNR improvement of fast motion simulation

Foreman									
bitrate(bits/sec)	**128k**			**256k**			**512k**		
	Proposed	**NoIntra**	**Gain**	**Proposed**	**NoIntra**	**Gain**	**Proposed**	**NoIntra**	**Gain**
Error Free	**30.19**	**--**	**--**	**33.58**	**--**	**--**	**36.82**	**--**	**--**
PLR 1%	**28.26**	**27.05**	**1.21**	**31.00**	**27.95**	**3.05**	**33.55**	**28.87**	**4.68**
PLR 5%	**26.57**	**20.84**	**5.73**	**28.10**	**22.35**	**5.75**	**28.98**	**22.84**	**6.14**
PLR 10%	**25.01**	**19.10**	**4.91**	**25.77**	**19.16**	**6.61**	**26.37**	**19.42**	**6.95**

As for D1 resolution video sequences, Crew sequence is used for simulations at different bitrates and packet loss rates. The results show that the RDIR can get 1.65~4.55 dB gain in PSNR.

Table 5. PSNR improvement of D1 resolution simulation

Crew									
bitrate(bits/sec)	**600k**			**900k**			**1200k**		
	Proposed	**NoIntra**	**Gain**	**Proposed**	**NoIntra**	**Gain**	**Proposed**	**NoIntra**	**Gain**
Error Free	**32.11**	–	–	**34.40**	–	–	**35.79**	–	–
PLR 1%	**31.21**	**29.56**	**1.65**	**32.86**	**30.53**	**2.33**	**33.92**	**31.51**	**2.41**
PLR 5%	**28.70**	**25.31**	**3.39**	**29.62**	**25.38**	**4.24**	**29.82**	**25.47**	**4.35**
PLR 10%	**26.65**	**22.95**	**3.60**	**26.88**	**22.96**	**3.92**	**27.34**	**22.79**	**4.55**

4.1.3 Combination with hybrid concealment

To evaluate the performance of the error resilience system, 3 series of experiments are adopted. For VOD applications under various network conditions, the error resilient capabilities are examined at 3 bitrates including 256, 550 and 700 kilo-bits per second (kbps). In each experiment, 4 error resilience systems as listed in **Table 6** are compared. Type 1 represents the original reference encoder and decoder system with default error resilience tools including Resync_Markers and zero-motion-vector spatial copy. Type 2 enables RDIR in the encoder. Type 3 enables the error robustness and hybrid concealment algorithm. Type 4 is the error system with RDIR in the encoder and hybrid concealment in the decoder. 3 different network conditions including packet loss rate (PLR) of 1%, 5% and 10% with uniform dropping probability model are used for testing. The test sequence is encoded with one I-VOP's and 99 P-VOP's.

The results are shown in **Table 7** to **Table 12**. We can find that the Type 2 can get a gain over Type 1 with 3.64-8.45 dB in PSNR. Type 3 can achieve a gain over Type 1 with 5.16-10.62 dB in PSNR. Based on error resilient decoder and rate-distortion intra refresh encoder, we can get a gain over Type 1 with 6.33-13.22 dB in PSNR. Fig. 22 to Fig. 27 demonstrates rate-distortion (R-D) curves to compare performance of the four error resilience systems with the different packet loss rates (PLR). **Fig. 28** shows that Type 4 has the quality improvement and the objective quality over the other 3 types.

Table 6. The 4 types of system with embedded error resilience tools

Type	Encoder	Decoder
1	Resynchronization marker	Zero motion for P-VOPs and spatial copy for I-VOPs
2	Intra-refresh	Zero motion for P-VOPs and spatial copy for I-VOPs
3	Resynchronization marker	Proposed hybrid concealment
4	Intra-refresh	Proposed hybrid concealment

Table 7. The reconstructed image quality for Foreman with 260 kbps

		Type 1	Type 2	Type 3	Type 4
PLR[2]	PSNR[3]	33.59	-	-	-
	Gain[1]	-	-	-	-
1%	PSNR	28.26	30.22	30.92	31.70
	Gain[1]	0	1.96	2.66	3.44
5%	PSNR	21.80	26.69	26.60	28.65
	Gain[1]	0	4.89	4.8	6.85
10%	PSNR	19.03	23.67	23.78	25.78
	Gain[1]	0	4.64	4.75	6.75

[1]Gain: the difference compared to PSNR of type 1 (unit:dB)

[2]PLR : packet loss rate

[3]unit: dB

Table 8. The reconstructed image quality for Foreman with 550 kbps

PLR2		Type 1	Type 2	Type 3	Type 4
0%	PSNR3	36.94	-	-	-
	Gain1	-	-	-	-
1%	PSNR	28.87	32.78	31.90	34.55
	Gain1	0	3.91	3.03	5.68
5%	PSNR	22.00	28.00	26.94	30.75
	Gain1	0	6.00	4.94	8.75
10%	PSNR	19.64	26.32	24.55	28.75
	Gain1	0	6.68	4.91	9.23

Table 9. The reconstructed image quality for Foreman with 700 kbps

PLR2		Type 1	Type 2	Type 3	Type 4
0%	PSNR3	38.04	-	-	-
	Gain1	-	-	-	-
1%	PSNR	30.10	34.29	34.03	35.70
	Gain1	0	4.19	3.93	5.6
5%	PSNR	21.60	28.46	27.05	31.63
	Gain1	0	6.86	5.45	10.03
10%	PSNR	19.41	26.74	25.23	29.54
	Gain1	0	7.33	5.82	10.13

Table 10. The reconstructed image quality for Akiyo with 130 kbps

PLR[2]		Type 1	Type 2	Type 3	Type 4
0%	PSNR[3]	41.2	-	-	-
	Gain[1]	-	-	-	-
1%	PSNR	22.57	39.83	33.86	39.95
	Gain[1]	0	17.26	11.29	17.38
5%	PSNR	21.68	37.5	30.18	39.21
	Gain[1]	0	15.82	8.5	17.53
10%	PSNR	20.62	35.55	28.03	36.6
	Gain[1]	0	14.93	7.41	15.98

Table 11. The reconstructed image quality for Akiyo with 200 kbps

PLR[2]		Type 1	Type 2	Type 3	Type 4
0%	PSNR[3]	43.15	-	-	-
	Gain[1]	-	-	-	-
1%	PSNR	22.61	41.34	33.93	42.1
	Gain[1]	0	18.73	11.32	19.49
5%	PSNR	21.7	38.68	30.12	40.13
	Gain[1]	0	16.98	8.42	18.43
10%	PSNR	20.63	36.28	28.78	37.38
	Gain[1]	0	15.65	8.15	16.75

Table 12. The reconstructed image quality for Akiyo with 310 kbps

PLR[2]		Type 1	Type 2	Type 3	Type 4
0%	PSNR[3]	44.89	-	-	-
	Gain[1]	-	-	-	-
1%	PSNR	22.55	39.83	34.48	43.74
	Gain[1]	0	17.28	11.93	21.19
5%	PSNR	21.72	37.5	31.02	40.59
	Gain[1]	0	15.78	9.3	18.87
10%	PSNR	20.64	35.55	29.26	37.75
	Gain[1]	0	14.91	8.62	17.11

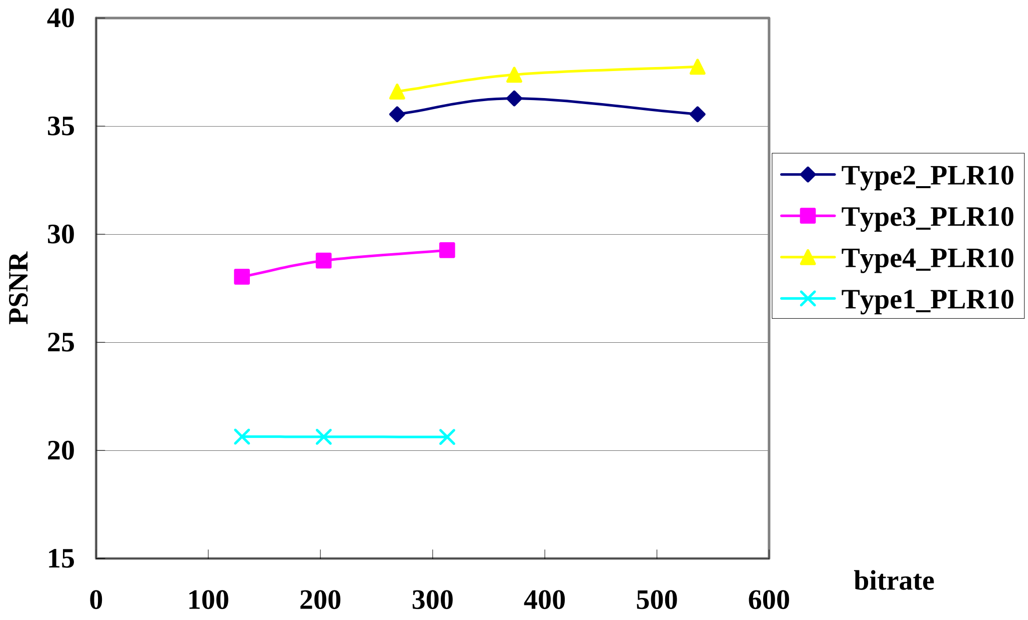


Fig. 22. R-D curve of reconstructed image quality for Akiyo sequence with PLR 10%

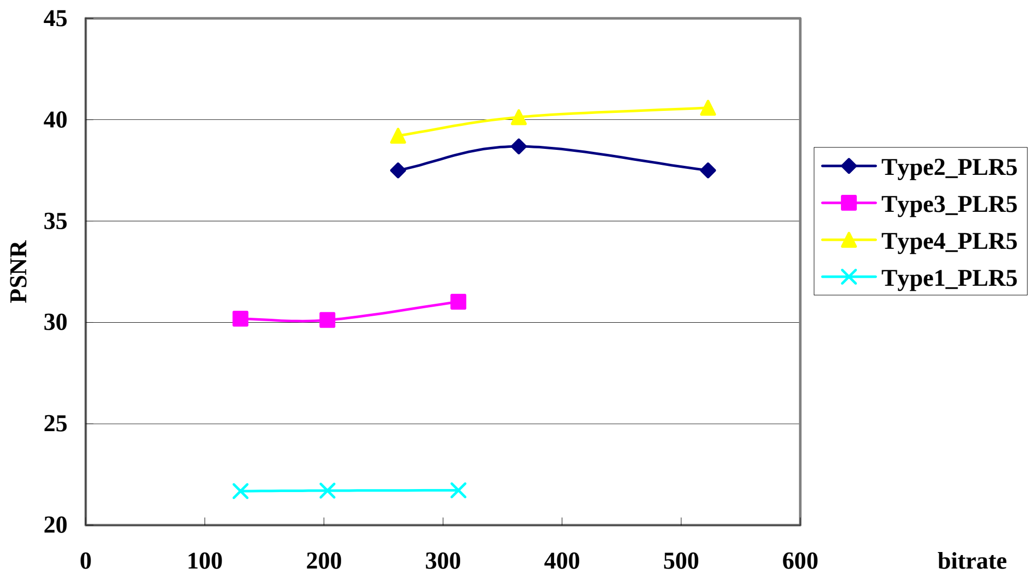


Fig. 23. R-D curve of reconstructed image quality for Akiyo sequence with PLR 5%

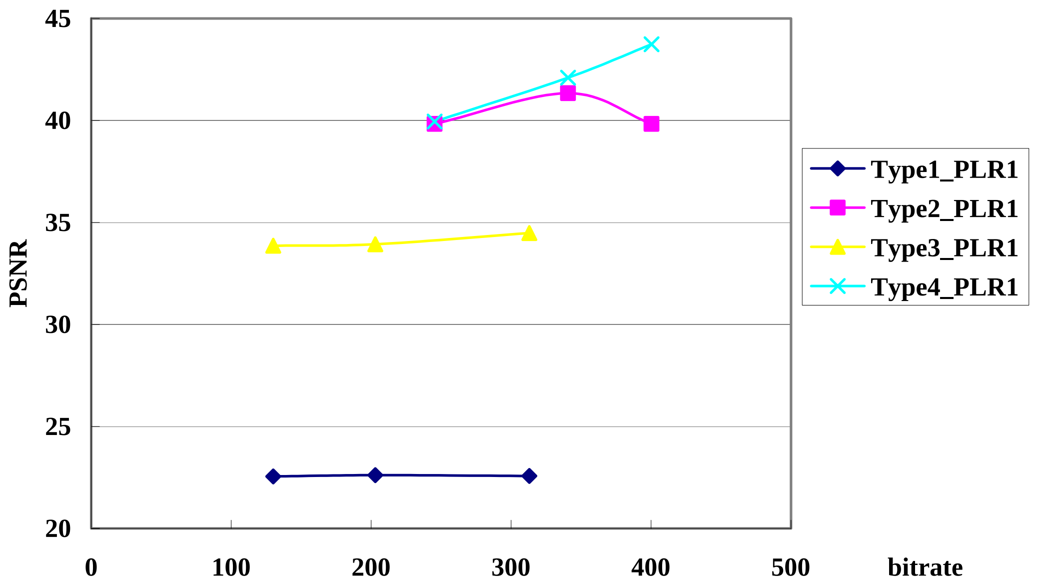


Fig. 24. R-D curve of reconstructed image quality for Akiyo sequence with PLR 1%

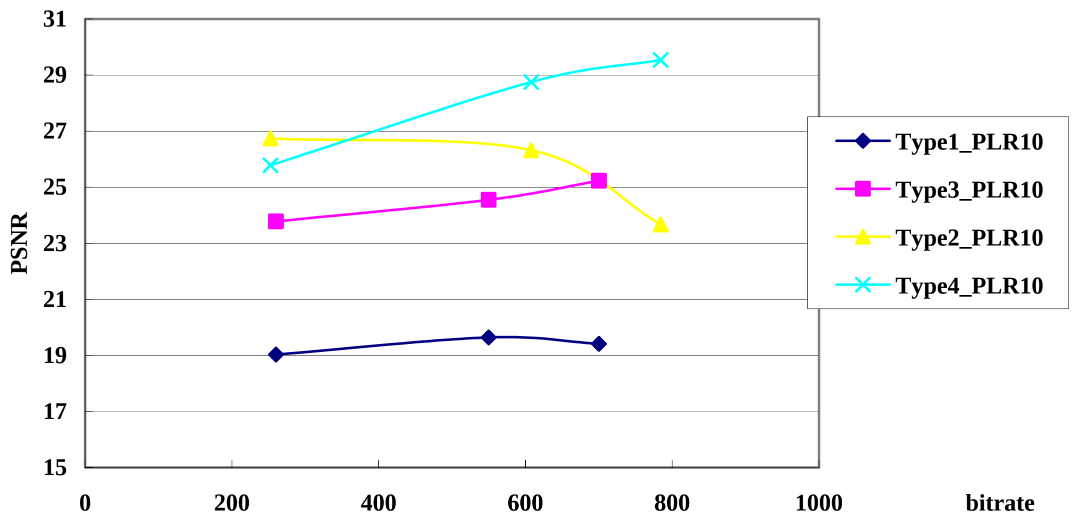


Fig. 25. R-D curve of reconstructed image quality for Foreman sequence with PLR 10%

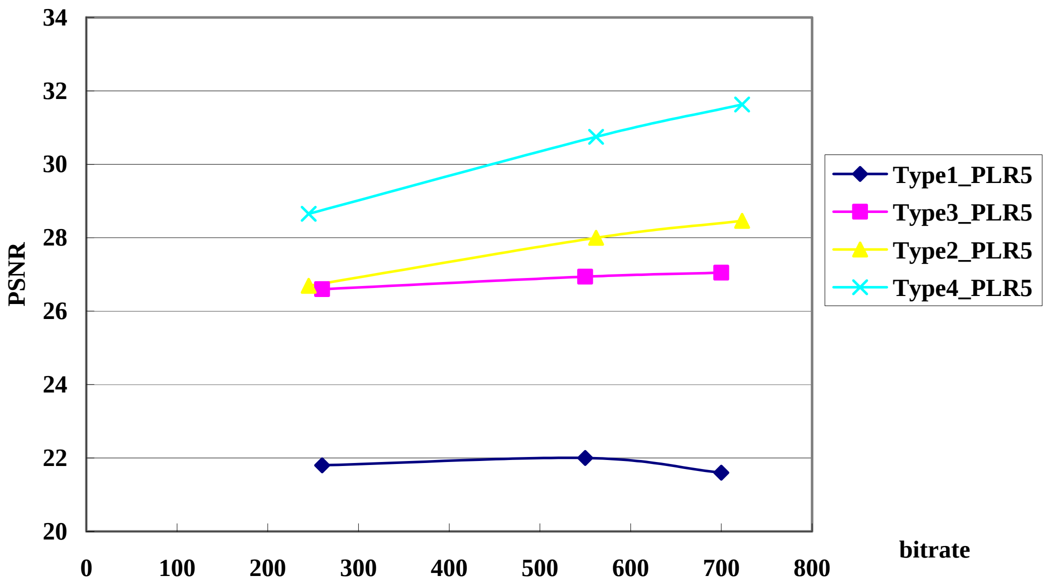


Fig. 26. R-D curve of reconstructed image quality for Foreman with PLR 5%

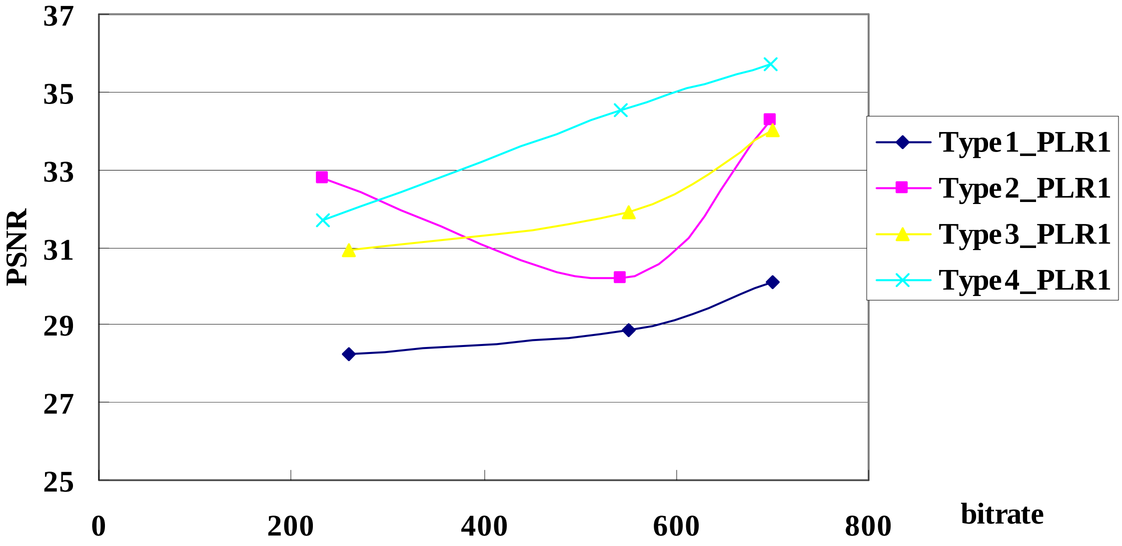


Fig. 27. R-D curve of reconstructed image quality for Foreman with PLR 1%

Fig. 22 to **Fig. 27** show a tendency that the PSNR values at high bitrate are lower than the PSNR values at low bitrates under identical error resilient tools and the same packet loss rate. For off-line testing, we randomly drop the packets based on a uniform probability model. The concealment performance is related with the lost area. When dropping the detailed or important areas, we may get the worse concealed video as any packet loss.

(a)

(b)

(c)

(d)

Fig. 28. The image quality comparison for the 4 types of error resilience system. (a) Type 1; (b) Type 2 ;(c) Type 3 ;(d) Type 4. (No. 8th frame, PLR 10%, bitrate 700kbps)

4.2 On-line Testing

With the on-line tests on the multimedia test bed, the simulations of video content delivery with the proposed robust streaming system without use of motion information over wireless channels are shown in Fig. 29 and Fig. 30. We test each case with different combinations of a retransmission scheme and Resync_Markers to evaluate the visual quality of the reconstructed video sequences. There are three combinations of the two techniques. Tool 1 disables simultaneously the retransmission scheme of the test bed and the use of Resync_Markers in the coded bitstream. Tool 2 adopts the same techniques as in the Tool 1 except the use of Resync_Markers. Tool 3 enables both tools at the same time, which can have the best subjective quality on the average for streaming video contents over the wireless channels. The simulation results show that the Tool 3 has the best visual quality. Observing the reconstructed video frames, the proposed video streaming system can provide better subjective visual quality of received videos for the consumers.

Retransmission can reduce the real packet loss rate for the decoder. Deciding which packets should be retransmitted first or using efficient retransmission can make the results better. In addition, with enough bandwidth to retransmit all the lost packets, we can almost make the decoder to achieve the picture quality of error free bitstreams.

	
(a)	(a)

	
(b)	(b)
	
(c)	(c)
Fig. 29. The visual quality comparisons for the 3 combinations of the tools within the error resilience system. (a) Tool 1 (b) Tool 2 (c) Tool 3.	**Fig. 30.** The visual quality comparisons for the 3 combinations of the tools within the error resilience system (a) Tool 1 (b) Tool 2 (c) Tool 3.

4.3 Summary

The experiment results show that the proposed RDIR method can improve the video quality under the packet-switching transmission channel. In addition, the hybrid concealment in decoder for off-line testing or retransmission scheme in streaming system for on-line testing can improve the overall visual quality of recovered video sequences. By combining with hybrid concealment in the decoder end or retransmission in the streaming system, we can increase the reconstructed video quality when any the packet loss occurs during transmission.

Chapter 5

Conclusion

5.1 Contributions

In this thesis, we have presented a novel R-D optimized intra refreshment, which inserts proper number of intra MBs to the encoding bitstream based on characteristics of video contents, packet-switched streaming schemes and network conditions. The video characteristics are consisting of the motion information, the duration of inter-frame prediction, the number of intra MBs per frame, and the R-D cost is used for deriving and updating the history record of each MB. The history record is adapted at various bitrates and for different packet loss ratios with proper number of overhead bits. The overhead bits are controlled by uniformly allocating bits to each frame in the current sequence. Thus, based on the innovative intra refreshment approaches, we can increase bitstream robustness against the transmission error to provide better quality of service over the Internet and wireless channels, which are both adopting packet-switching protocol for data transmission.

5.2 Future Works

The current intra refreshment methods could provide much better video quality for video streaming and communications with further investigation on the characteristic of rapidly varying network environments. In addition, we will try to fine-tune the proposed intra refreshment algorithms for the video coding standard MPEG-4 Advanced Video Coding/H.264 that can provide high coding efficiency with high computational cost.

Reference

[1] J. Zhang, J. F. Arnold and M. R. Frater, "A cell-loss concealment technique for MPEG-2 coded video", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 10, No. 4, pp.659-665, June 2000

[2] B. Yan, and K. W. Ng, "A Novel Selective Motion Vector Matching Algorithm for Error Concealment in MPEG-4 Video Transmission over Error-Prone Channels", *IEEE Transactions on Consumer Electronics*, Vol. 49, No. 4, pp.1416-1423, November 2003

[3] L. Atzori, F. G. B. De Natale, and C. Perra, "A Spatio-Temporal Concealment Technique Using Boundary Matching Algorithm and Mesh-Based Warping (BMA–MBW)", *IEEE Transactions On Multimedia*, Vol. 3, No. 3, pp.326-338, September 2001

[4] S. Valente, C. Dufour, F. Groliere, and D. Snook, "An Efficient Error Concealment Implementation for MPEG-4 Video Streams", *IEEE Transactions on Consumer Electronics*, Vol.47, No.3, pp.568-578, August 2001

[5] M. Ghanbari and V. Seferidis, "Cell-Loss Concealment in ATM Video", *IEEE transactions on circuits and systems for video technology*, Vol.3, No.3, pp.238-247, June 1993

[6] S. Kaiser, K. Fazel, "Comparison of error concealment techniques for an MPEG-2 video decoder in terrestrial TV-broadcasting", *Signal Processing: Image Communication*, Vol. 14, No.6-8, pp. 655-676, May 1999

[7] S. Cen and P. C. Cosman, "Decision Trees for Error Concealment in Video Decoding", *IEEE Transactions On Multimedia*, Vol. 5, No. 1, pp.1-7, March 2003

[8] W.-J. Chu and J.-J. Leou, "Detection and Concealment of Transmission Errors in H.261 Images", *IEEE Transactions On Circuits And Systems For Video Technology*, Vol. 8, No. 1, pp.74-84, February 1998

[9] M.-H. Jo and W.-J. Song, "Error Concealment for MPEG-2 Video Decoders with Enhanced Coding Mode Estimation", *IEEE Transactions on Consumer Electronics*, Vol.46, No.4, pp.962-969, November 2000

[10] P. Salama, N. B. Shroff, and E. J. Delp, "Error Concealment in MPEG Video Streams Over ATM Networks", *IEEE Journal On Selected Areas In Communications*, Vol. 18, No. 6, pp.1129-1144, June 2000

[11] J.-W. Suh and Y.-S. Ho, "Error Concealment Techniques for Digital TV", *IEEE Transactions On Broadcasting*, Vol. 48, No. 4, pp.299-306, December 2002

[12] L. D.-Soares, F. Pereira, "Error resilience and concealment performance for MPEG-4 frame-based video coding", *Signal Processing: Image Communication*, Vol. 14, No.6-8, pp.447-472, May 1999

[13] J. F. Arnold, M. R. Frater, J. Zhang, "Error resilience in the MPEG-2 video coding standard for cell based networks - A review", *Signal Processing: Image Communication*, Vol.14, No.6-8, pp.607-633, May 1999

[14] Y. Wang, S. Wenger, J. Wen, and A. K. Katsaggelos, "Error Resilient Video Coding Techniques ", *IEEE Signal Processing Magazine*, Vol.17, No.4, pp.61-82, July 2000

[15] M. E. A.-Mualla, C.N. Canagarajah and D. R. Bull, "Motion Field Interpolation for Temporal Error Concealment", *IEE Processing Vision, Image and Signal Processing*, Vol.147, No.5, pp.445-453, October 2000

[16] S. Tsekeridou, and I. Pitas, "MPEG-2 Error Concealment Based on Block-Matching Principles", *IEEE Transactions On Circuits And Systems For Video Technology*, Vol. 10, No. 4, pp.646-658, June 2000

[17] W. Kwok and H. Sun, "Multi-Directional Interpolation for Spatial Error Concealment", *IEEE Transactions on Consumer Electronics*, Vol.39, No.3, pp.455-460, August 1993

[18] J.-Y. Pyun, J.-S. Lee, J.-W. Jeong, J.-H. Jeong, and S.-J. Ko, "Robust Error Concealment for Visual Communications in Burst-Packet-Loss Networks", *IEEE Transactions on Consumer Electronics*, Vol. 49, No. 4, pp.1013-1019, November 2003

[19] S. Aign, K. Fazel, "Temporal and Spatial Error Concealment Techniques for Hierarchical MPEG-2 Video Codec", *1995 IEEE International Conference on Communications*, Vol.3, pp.1778-1783, 18-22 June 1995

[20] E. T. Kim, S.-J. Choi, H.-M. Kim, "Weighted boundary matching algorithm for error concealment in the MPEG-2 video bit stream", *Signal Processing*, Vol.73, No.3, pp.291-295, January 3,1999

[21] Y.-L. Lee, C.-N. Wang and T. Chiang, "An MPEG-4 error resilient decoder", *2002 Taiwan Workshop on Consume Electronics*, October 2002

[22] M.-Y. Huang, T.-L. Su, S.-H. Wang, C.-N. Wang, and T. Chiang, "An Error Resilient System for Streaming MPEG-4 Video over the Internet", *2003 Taiwan*

Workshop on Consume Electronics, November 27-28, 2003

[23] T.-L. Su, M.-Y. Huang, H.-C. Chuang, C.-N. Wang, and T. Chiang, "A Robust MPEG-21 Media Delivery Video Streaming System", *8th World Multiconference on Systemics, Cybernetics and Informatics (SCI2004)*, July 18-21, 2004

[24] H. Sun, J. W. Zdepski, W. Kwok, and D. Raychaudhuri, "Error Concealment Algorithm for Robust Decoding of MPEG Compressed Video", *Signal Processing：Image Communication*, Vol.10, No.4, pp.249-268, September 1997

[25] J. Y. Liao and J. Vilasenor, "Adaptive intra block update for robust transmission of H.263," IEEE Transactions on Circuits and Systems for Video Technology, vol. 10, pp. 30 -35, Feb. 2000

[26] R. Talluri, I. Moccagatta, Y. Nag, and G. Cheung, "Error concealment by data partitioning", *Signal Processing: Image Communication*, Vol.14, No.6-8, pp.505-518, May 1999

[27] M. Bystrom, V. Parthasarathy, and J. W. Modestino, "Hybrid Error Concealment Schemes for Broadcast Video Transmission over ATM Networks", *IEEE Transactions On Circuits And Systems For Video Technology*, Vol. 9, No. 6, pp.868-881, September 1999

[28] I. Burnett, R. V. de Walle, K. Hill, J. Bormans, and F. Pereira, "MPEG-21: Goals and Achievements", *IEEE Multimedia*, Vol.10, No.4, pp.60-70, October–December 2003

[29] J. Bormans, J. Gelissen, and A. Perkis, "MPEG-21：The 21st Century Multimedia Framework", *IEEE Signal Processing Magazine*, Vol.20, No.2, pp. 53-62, March 2003

[30] J. R. Smith, "MPEG-21 Digital Item Adaptation : Enabling Universal Multimedia Access", *IEEE Multimedia*, Vol.11, No.1, pp.84-87, January–March 2004

[31] C.-N. Wang, C.-Y. Tsai, H.-C. Lin, H.-C. Chuang, Y.-C. Lin, J.-H. Chen, K. L. Tong, F.-C. Chang, C.-J. Tsai, S.-Y. Lee, T. Chiang, and H.-M. Hang, "FGS-Based Video Streaming Test Bed for MPEG-21 Universal Multimedia Access with Digital Item Adaptation", *in Proc. of ISCAS2001*, 2001.

[32] Coding of Audio-Visual Objects, Part-2 Visual, Amendment 1: Visual extensions, ISO/IEC 14496-2/FPDAM1, January 2000

[33] G. J. Sullivan, and T. Wiegand, "Rate-Distortion Optimization for Video Compression", *IEEE Signal Processing Magazine*, Vol.15, pp.74-90, Nov. 1998.

[34] G. Cote, S. Shirani, F. Kossentini, "Optimal Mode Selection and Synchronization

for Robust Video Communications over Error-Prone Networks”, *IEEE Journal of Selected Areas of Communications*, Vol.18, No.6, pp.74-90, June 2000

[35] R. Zhang, S. L. Regunathan, K. Rose, “Video Coding with Optimal Inter/Intra-Mode Switching for Packet Loss Resilience”, *IEEE Journal of Selected Areas of Communications*, June 2000

[36] G. Cote and F. Kossentini, “Optimal Intra Coding of Blocks for Robust Video Communication over The Internet”, *Signal Processing: Image Communication*, pp. 25-34, September 1999

[37] S. Wenger and G. Cote, “Using RFC2429 and H.263+ at Low to Medium Bit-rates forLow-latency Applications”, *Proceeding Packet Video*, April 1999

[38] C. Zhao, K.N. Ngan, J. Zhang, R. Mathew and X. Zhang, “Using inter-frame dependence history to select intra-refersh blocks”, *Electronics Letters*, Vol.38, No.22, pp.1337-1338, October 2002

[39] P. Haskell and D. Messerschmitt, “Resynchronization of motion compensated video affected by ATM cell loss”, *Proceeding IEEE Conference Acoustic Speech Signal Processing*, Vol.3, pp.545-548, March 1992

[40] K. Imura, Y. Machida, "Error resilient video coding schemes for real-time an low-bitrate mobile communications", Signal Processing: Image Communication 14 (1999) 519~530

自　　傳

蘇子良: 民國 1980 年生於台南市。 2002 畢業於台灣新竹的國立交通大學電子工程學系，之後進入該校電子工程所攻讀碩士學位。以 MPEG-4 編碼器為論文研究主題。

Tzu-Liang Su was born in Tainan in 1980. He received the BS degree in Electronics Engineering, National Chiao Tung University (NCTU), HsinChu, Taiwan in 2002. His current research interest is MPEG-4 encoder.