This is to certify that the
dissertation entitled

INTERLEAVED SOURCE CODING FOR PACKET VIDEO
OVER ERASURE CHANNELS

presented by

JIN YOUNG LEE

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Electrical Engineering

Hayder Radha
Major Professor's Signature

Dec. 11, 2008
Date

INTERLEAVED SOURCE CODING FOR PACKET VIDEO OVER ERASURE CHANNELS

By

Jin Young Lee

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

ELECTRICAL ENGINEERING

2008

ABSTRACT

INTERLEAVED SOURCE CODING FOR PACKET VIDEO
OVER ERASURE CHANNELS

By

Jin Young Lee

In this dissertation, a new source coding framework, Interleaved Source Coding (ISC), is introduced and investigated. The basic idea of ISC is to code a single video sequence into multi sub-sequences using a predictive video coder while taking into consideration a given singe packet-erasure channel model and the video frames' temporal correlations to reduce the frequency and impact of the cascaded effect of packet erasures and related propagation of decoding errors resulting from the predictive nature of coded video. The ISC framework provides optimum solution for different erasure channel models such that the impact of losses are limited to a minimum number of video frames while reducing quality degradation from video frame replacements during error concealments.

Initially, we focus on generating optimum ISC streams by partitioning a predictive video sequence into two sub-sequences of coded video. The design of ISC employs Dynamic Programming based on a *reward process*. Memoryless Binary

Erasure Channel (BEC) model cases (ISC-BEC) with various erasure rates are examined to validate the initial design of the proposed framework. The initial ISC scheme is then extended and validated using a Markov Reward Process (MRP) and a Markov Decision Process (MDP) that are mapped into a packet-erasure channel with memory (ISD-MDP). Furthermore, two sub-stream canonical ISC scheme is extended to multi-stream ISC and firm benefits of interleaving has been observed with respect to channel condition and encoding characteristics. Finally, rate-distortion optimized Forward Error Correction (FEC) is adopted into ISC to maximize the performance of the proposed ISC framework. Unlike ISC-BEC and ISD-MDP, ISC-FEC employs rate-distortion optimization in the optimal interleaving-set selection process so that it can benefit from both FEC and ISC. The performance improvement using rate-optimized ISC-FEC is analyzed, evaluated, and compared with ISC-MDP.

To my father for his eternal love

And

To my mother for her endless sacrifice.

TABLE OF CONTENTS

Cha

C

Ref

LIST OF TABLES

LIST OF FIGURES

Figu

Figu

Figu

Figu

Figu

Figu

Figu

Figu

Figu
Akj

Figu
Akj

Figu
Akj

Figu
Coa

Figu
Coa

Figu
Coa

Figu

Figu

Figu

Figu

Figu

Cl

Intr

The

wor

Suc

chat

cont

prov

over

Serv

put,

exce

comr

strea

ousl,

vided

Chapter 1

Introduction

The expansion of the underlying infrastructure of the Internet Protocol (IP) network has lead to increased demand of realtime on-line streaming video services. Such services are often used in multimedia content transmission such as video chat, live news, video conferencing, video on demand (VOD), IPTV, user created content (UCC) streaming, etc. However, despite of the growth and the improvements of the Internet infrastructure, realtime streaming media transmitted over best-effort IP network often faces difficulties in guaranteeing Quality-of-Service (QoS). This is due to the network impairments, e.g., variation in throughput, delay [1] and packet losses, which are often caused by the service requests exceeding permitted limit of the networks [2]. The result is unreliable realtime communication interactivity and degraded video quality during playback since the streaming video is portioned into packets for delivery and played out simultaneously during video delivery. Detailed information on network impairments is provided in a latter section.

1.1

Mos

crea

tion

her

Du

str

for

tec

Te

ma

pl

m

in

sy

th

a

1.1 Research Problems

Most of the realtime streaming video services uses predictive video coders to increase the network utilization. The predictive video coders use motions estimation and compensation to reduce the redundancies among adjacent frames, and hence reduce the number of bits required to represent the original video content. Due to the nature of predictive video coders, the quality degradation of realtime streaming video service is mainly caused by packet losses [3-17]. Therefore, for playback quality improvement of realtime streaming video, special coding techniques resilient to packet losses are required.

Techniques such as scalable coding [3, 7, 18-20], multi-hypothesis motion estimation and compensation [21, 22], multi state video compression [23], and multiple description coding (MDC) with path diversity [24-28] are few examples of methods that are resilient to packet losses. Studies have shown that such coding schemes are resilient to packet losses, however, the coding complexities of such coders are far greater than conventional predictive video coding, and hence, their usages are limited. Detailed descriptions on the above coding schemes, along with the pros and cons of them are given in Chapter 2.

1.

In

(IS

for

sec

ing

oth

the

mul

tion

an I

sure

of th

cade

sulte

timur

cause

oryles

1.2 Contribution of Research

In this dissertation, a new source coding framework, Interleaved Source Coding (ISC), is introduced and investigated to provide packet loss resilient video-coding for predictive video sequences. The basic idea of ISC is to code a single video sequence into multiple sub-sequences using a predictive video coder while taking into consideration a given packet-erasure channel model. ISC differs from other multi-sequence coding, most notable Multiple Description Coding (MDC), in the sense that ISC sub-streams are transmitted over a single channel rather than multiple channels. Therefore, ISC eliminates channel selection, content distribution, and synchronization issues that are associated with MDC [24-28]. Therefore, an ISC channel model is based on well-known parameters, such as packet erasure rate and packet-loss correlation, of a single erasure channel. The objective of the ISC coding framework is to reduce the frequency and impact of the cascaded effect of packet erasures and related propagation of decoding errors resulted from the predictive nature of coded video. In other words, ISC provides optimum solution for different channel models such that the impact of losses caused by a given erasure channel model (with memory [4, 12, 29, 30] or memoryless [31-33]) is limited to a minimum number of video frames. In addition, the op

fra

the

this

pos

on

ten

rep

In t

par

sec

ory

mo

BE

fran

and

for t

ii) IS

optimum ISC solution also reduces quality degradation resulting from video frame replacements, which are used as low-complexity error concealments when the decoder fails to recover a video frame due to packet losses. ISC achieves this improvement by minimizing video frame replacement distances. Our proposed design of ISC coded video employs Dynamic Programming [34-37] based on a *reward process*. Furthermore, some coarse measure of video frames' temporal correlations is also taken into consideration to reflect the video-frame replacement/concealment process.

In this dissertation, ISC optimization is first examined for the canonical case of partitioning a video sequence into two sub-sequences. This canonical case (two-sequence ISC interleaving) is examined over various channel models. The memoryless Binary Erasure Channel (BEC) is known to be simplest erasure channel model [31-33] Consequently, first, the memoryless BEC-based ISC model (ISC-BEC) with various erasure rates are examined to validate the design of the ISC framework. The ISC-BEC case is analyzed and studied for both non-correlation and video frame correlation cases. In particular, the following cases are covered for the ISC-BEC scenario; i) ISC-BEC without video frame correlation measure, ii) ISC-BEC with sequence specific correlation measure, iii) ISC-BEC with gene

cha

coc

inte

Bu

is e

Pro

cha

mo

wit

mo

als

fom

ing

Upo

ten

sub

in te

neric frame correlation measure. Each scenario is examined with various BEC channel capacities. MPEG-4 video coder [38] is the choice of predictive video coder and results from each evaluation scenario are compared to non-interleaving single layer coding to validate the performance improvement of ISC. Building upon the initial ISC design over a BEC channel model, the ISC scheme is extended using a Markov Reward Process (MRP) and a Markov Decision Process (MDP)) [34, 35, 37, 39, 40] that are mapped into a packet-erasure channel with memory. We focus on the two-state Markov Model, *a.k.a. Gilbert model* [15, 29, 30, 34, 38, 41-43], which is known to model network channel with packet losses more realistically when compared to the memoryless BEC model. Similar to the BEC case, ISC over channels with memory (ISC-MDP) is also evaluated for both non-correlation and correlation video cases. The performance improvement using ISC over non-interleaving traditional predictive coding method is validated.

Upon the validation of the ISC design over channel with memory, ISC-MDP is extended to multi-stream ISC to find relationships among the number of interleaving sub-streams or the GOV size of an interleaving sub-stream and the performance in terms video quality. Three and four sub-stream ISC cases are evaluated. The ev

ha

ha

tic

Ba

Fo

pe

BE

int

Th

eva

1.3

The

reg

coo

ter

era

evaluations have shown that increasing the number of interleaving sub-streams have no noticeable improvement on quality; however, firm benefits of interleaving has been observed with respect to channel condition and encoding characteristics.

Based on the observations from the previous parts, rate-distortion optimized Forward Error Correction (FEC) is adopted into ISC (ISC-FEC) to maximize the performance of the proposed Interleaving source coding framework. Unlike ISC-BEC and ISD-MDP, ISC-FEC takes a slightly different approach in the optimal interleaving-set selection process such that it can benefit from both FEC and ISC. The performance improvement using rate-optimized ISC-FEC is analyzed, evaluated, and compared with ISC-MDP.

1.3 Organization

The remainder of this dissertation is organized as follows: Background material regarding network impairments, predictive coding, various packet loss resilient coding methods, and Forward Error Correction (FEC) coding are given in Chapter 2. In Chapter 3, the proposed ISC coding method is introduced with a general description on interleaving, an analytical approach for identifying the optimum interleaving set over the memoryless Binary Erasure Channel (BEC) using Reward based Decision Process (RDP). Chapter 4 extends the ISC scheme for channel with memory using a Dynamic Programming algorithm in conjunction with a Markov Reward Process (MRP) and a Markov Decision Process (MDP). In Chapter 5, previously evaluated and validated the ISC design is extended to multiple sub-stream ISC to validate benefits of ISC as a function of the channel model and encoding characteristics variation. Chapter 6 extends the ISC framework to adopt rate-distortion optimized Forward Error Correction (FEC), ISC (ISC-FEC), which further improves the performance of the proposed ISC framework. Finally, Chapter 7 concludes this dissertation and identifies future directions.

C

B

Th

Fo

re

A r

Firs

pre

sio

ada

pas

dec

Chapter 2

Background Information

This chapter gives an overview of the network impairments, predictive coding, Forward Error Correction (FEC) coding, and past and current work in packet loss resilient coding method for realtime streaming video applications.

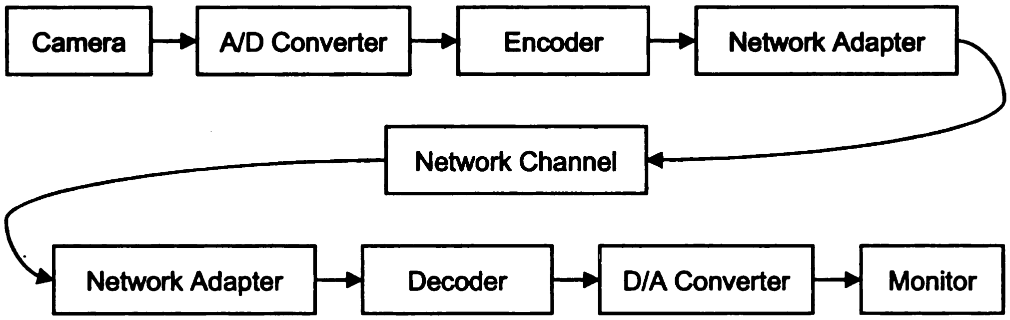


Figure 1 Realtime Streaming Video Service System

A realtime streaming video service system can be represented as in Figure 1. First, video signal is digitized and fed into the encoder. Then the encoder compresses the digitized video stream to reduce the bandwidth needed for transmission. Once compressed, the encoded signal is passed on to the network adapter where it is broken into packets to be transmitted over the network. After passing through various links and routers, packets reach the destination where decoding and error handling of the transmitted packets are performed for presenta

W

to

co

ue

de

pr

Ne

bu

is

tha

los

clie

At a

dicta

for r

tation.

When measuring the Quality-of-Service (QoS) of the delivered media, degradation can be caused by any component in Figure 1. Coding perspective, signal conversions, A/D and D/A, affect the quality depending on the quantization values taken. In addition, compression scheme [38, 41] plays major role in quality degradation, however, improvement of coding schemes can overcome such problem by providing high perceptual quality with high compression ratios. Network perspective, the main cause is packet loss, and this is usually caused by buffer overflow of packet transmitting network components. The buffer overflow is usually observed when data arrives to the network components at a higher rate than the buffer handling capacity of component. Most cases, such bottleneck loss occurs at the packet switching routers in the network between servers and clients.

At any rate, while coding losses are controllable, network losses are often unpredictable uncontrollable, method to reduce the impact of network losses is in need for realtime streaming media service system.

2.

D

te

m

na

th

of

eff

lin

2.2

Cu

en

du

Int

cee

wh

2.1 Network Impairments

Due to the realtime delivery constraints of realtime streaming video services often require reliable network that meets their quality of service (QoS) requirements, e.g., throughput, delay, and error resiliencies. However, the unpredictable nature of internet traffic results in network impairments such as variation in throughput, delay, and packet losses, which in turn severely degrades the quality of delivered video[1, 3-17, 32, 44-48]. The network impairments of the best-effort IP network are associated with the main components, the routers and the links interconnecting the routers, and often related to excess service requests.

2.2 Variation in throughput

Currently, the best-effort IP network does not provide end-to-end bandwidth reservation mechanism. It is well known that the available bandwidth is unknown due to fluctuations of the Internet traffic. As a result of a sudden increment of Internet data transmission requests or a transmission rate of a media stream exceeding the available bandwidth of the connection, network congestion occurs which in turn causes bursty packet losses or delays in media delivery.

2.3

Ba

are

net

out

cor

Wh

size

low

her

It is

bes

incr

the

con

It is

paci

2.3 Packet Loss

Basic operations of the packet switching routers are as follows: Incoming packets are stored or queued in the input buffers of the router to be processed by the network processors. When processed, packets is forwarded to an appropriate output port and stored in the output buffer to be released through the output link connecting intermediate router.

When the incoming data rate is higher than the forwarding rate of the routers, the size of processing queue increases, and unless the incoming rate decreases below the forwarding rate, queue reaches the size of the buffer, *bottleneck*, and hence, any new incoming packet is dropped automatically, *buffer overflow*.

It is well known that buffer overflow is the main cause of the packet losses in best-effort packet switching network. It is possible to prevent buffer overflow by increasing buffer size, however, it does not resolve packet loss problem due to the expiration of the time-to-live (TTL) in the header of the IP packets, another congestion control mechanism for the network.

It is seldom, however, link failures or unstable lossy links are another cause of packet losses. In most cases, link related losses are observed in the wireless net

los

link

con

pac

2.4

The

wa

The

Era

sio

orig

pas

pac

The

dis

are

network environment. In the wireless environment, packets are lost when the lossy link condition is observed due to channel noise or signal strength fading, or link failure from handoffs or intermittent loss of connection. In case of lossy link condition, bit error rate (BER) increases and if the errors cannot be corrected, the packet is dropped by network adapter.

2.4 Channel Modeling

The packet transmitting network channel is usually modeled in two different ways; memoryless channel and channel with memory.

The memoryless channel, a very simplistic channel model, is often called Binary Erasure Channel (BEC) [31-33]. In this channel model, the packet's transmission loss or success at time t is determined only by the packet transmission originated at time t and does not have any relationship with prior or posterior packet transmissions. From the statistical perspective, it can be said that packet is lost with a certain probability p without any conditional probability. Therefore, the memoryless channel shows independent and identical packet loss distribution and hence, the packet losses observed in memoryless channel model are mostly single isolated losses and the burst loss occurrences are very rare.

Ho

te

in

qu

lac

ch

los

38

Ma

ch

for

for

to

2.5

2.5

Du

stre

However, studies on Internet traffic have shown that majority of packet loss patterns observed are bursty [4, 5, 11-15, 29, 49, 50] due to the reasons described in the previous section, hence the simple memoryless channel model is not adequate to model the Internet channel realistically. Therefore, to supplement the lack of adequacy, memory is added to the BEC model. By adding memory, the channel with memory model allows to define a conditional probability of a packet lost based on the previous packet(s) transmission state, success or lost [29, 30, 38, 41]. The most widely used model is the Gilbert model which is a two state Markov model. The model consists of two states: the good (G) and bad (B) channel state with two transition probabilities, p_{01} and p_{10}. The parameters for the model can be easily obtained from a packet loss trace and the parameters for the model replication are set such that the model generates loss bursts close to the realistic packet trace.

2.5 Video Compression and Predictive Video Coding

2.5.1 Video Compression

Due to the bandwidth limitation of the underlying internet infrastructure, video streams are often compressed before transmission to meet the realtime delivery cons

ofter

are t

dia is

spac

capa

fits, fo

coder

comp

Storag

desire

coder

entatio

pressio

constra

video fo

quality

constrains of the realtime streaming video service. Video compression methods, often called video source coders, such as MPEG-2, MPEG-4, H.263, or H.264 are the ones that are widely used in today's multimedia industry. When the media is encoded using the source coder, it not only helps to reduce the storage space, but also helps to maximize the utilization of bandwidth, limited by control capability of underlying network infrastructure. However, aside from such benefits, followings must be taken into the consideration before employing source coders; computation complexities of encoder and decoder, and distortions from compression.

Storage and bandwidth utilization perspective, high compression ratio is always desired, however, the trade-offs are increased computation complexities of encoder and decoder of which the effect is severe enough to exceed realtime presentation constraint, coding delays. In addition, distortions from high ratio compression can also degrade presentation quality that will not meet desired quality constraint set by the service or the viewer. Therefore, source coding of the video for realtime application must be performed deliberately to meet delay and quality constraints of the service.

2.5.

Vid

algo

mos

cos

is u

dan

tion

repr

ing

The

still-

pen

caus

fines

codi

curre

crea

2.5.2 Predictive Video Coding

Video compression standards [38, 41] use some form of redundancy reduction algorithm to improve the compression ratio with minimal quality degradation. In most cases, for spatial redundancy reduction, intra-frame coding, the discrete cosine transformation (DCT) in conjunction with the variable length coding (VLC) is used where motion estimation and compensation are used for temporal redundancy reduction, inter-frame coding. Since motion estimation and compensation depend on previously encoded frame to determine motion vectors, which represent the transformation of current frame from previous one, such video coding schemes are also called predictive video coding.

The compression efficiency of the predictive video coding is far greater than that still-image based non-predictive video coding, however, due to the temporal dependent nature, the predictive video coding is more prone to error propagation caused by packet losses than the other. While non-predictive video coding confines reconstruction error from a single packet loss to one frame, predictive video coding propagates reconstruction error to all future frames that depend on the current frame from the motion estimation and compensation perspective, hence creates similar reconstruction error patterns as with bursty packet losses.

Th

pa

2.

Fo

wh

co

co

Sc

Mu

ing

2.6

Sca

com

the

a m

scai

Therefore, it is important to reduce or eliminate the error propagation from the packet loss in predictive video coding.

2.6 Error Resilient Video Coding

For realtime streaming video services over best-effort IP based packet network where retransmission of lost packets are inadequate due to realtime delivery constraint, to adequately minimize the error propagation effect in predictive video coding, various coding techniques are adopted that are resilient to packet losses. Scalable Video Coding (SVC) [3, 7, 18-20], Multiple Description Coding, and Multi-hypothesis Coding schemes are few examples of error resilient video coding scheme that are available today.

2.6.1 Scalable Video Coding (SVC)

Scalable Video Coding (SVC) is the name of the H.264/MPEG-4 AVC video compression standard extension [41]. However, before the name was fixed to the current video compression standard, scalable video coding in general meant a multi-layered coding scheme that provides spatial, temporal, and SNR/quality scalabilities[3, 7, 18-20].

Th

pie

dep

vid

bas

har

of t

Hov

sub

bas

data

shov

pack

cod

for th

2.6.2

Multi

[1] H

The objective of the SVC[1], for both old and new terminology, is to provide multiple scalabilities in one coded stream which can be transmitted at different bitrate depending on the network condition. In order to achieve the goal, SVC encodes video stream into one, *base layer*, or more layers, *enhancement layers*. The base layer is necessary for the media stream to be decoded, whereas the enhancement layers are applied to improve stream quality, and yet the transmission of the enhancement layers are optional depending on the channel condition. However, since each enhancement layer depends on either the base layer or its subordinate layer, decoding of enhancement layers are interrupted whenever the base layer and/or the subordinate layers are missing and, as a consequence, the data of the respective enhancement layers is rendered useless. Studies have shown that video streams coded with multiple scalable layers are resilient to packet losses, however, despite of the improvement, due to the increased the coding complexity compared to single layer coding, it is more feasible to be used for the pre-encoded, stored media services.

2.6.2 Multiple Description Coding (MDC)

Multiple Description Coding (MDC) [24-28] is a coding scheme which encodes a

[1] Here, the acronym SVC is used for both old and new terminology.

s

s

te

q

H

tic

pa

em

cor

cor

2.6.

Mu

esti

dard

the r

ory c

such

single media stream into n independent sub streams (n >= 2), referred to as descriptions, and transmits each sub streams over multiple paths. When transmitted, any description can be used for decoding and presentation, however, the quality can be improved with the number of descriptions received in parallel. Hence MDC resilient to the packet losses since an arbitrary subset of descriptions can be used to decode the original stream, therefore, interruptions from packet losses are minimal, they only degrades the quality of video. Despite the error resilient property of MDC, the use of MDC is minimal due to the high coding complexity, path selection problem for each description, and synchronization complexity of received descriptions.

2.6.3 Multi-hypothesis Coding

Multi-hypothesis (MHC) coding [21, 22] uses multiple reference frames for motion estimation and compensation. It differs from B-frame coding concept of standard predictive video coders since MHC uses only past frames for reference and the number of references can be set arbitrary depending on the long-term memory capacity of the encoder and decoder. The error resilient property of MHC is such that a frame can be decoded with marginal quality as long as there is a refere

pict

and

qua

mini

2.7

Forw

rects

corre

tage

serve

adeq

retran

mum

ness a

erence frame presents in the memory of the decoder. However, total group of picture decoding failure problem still presents if the first frame in the group is lost, and yet, the packet loss related error propagation of MHC is almost limited to quality degradation, not total frame losses. Similar to MDC, the use of MHC is minimal due to increased coding complexity.

2.7 Forward Error Correction (FEC)

Forward error correction (FEC) [6, 49-63] is an error correction system which corrects errors at the receiving end of data transmission with the redundant error correction codes that are added by the sender before transmission. The advantage of FEC is that the receiver can avoid retransmission of data if error is observed in the data. Therefore FEC is usually used when retransmission is not adequate due to increase cost from retransmission or delay constraint of the data retransmission. There are many types of FEC, but the most notable is Maximum Distance Separable (MDS) coding [49-51, 53-59] because of its compactness and adaptation simplicity compared to other FEC schemes.

C

Int

The

coo

1)

con

cha

dei

In

ret

To

pos

into

ove

que

tion

Chapter 3

Interleaved Source Coding over Binary Erasure Channel

The purposed of this research is to propose a new packet loss resilient video-coding approach for predictive video sequences with the following constraints: 1) It must take channel condition, in terms of loss probability. 2) The coding complexity cannot be greater than any of the methods described in the previous chapter. 3) Network transmission overhead must be minimized such that the delay at the receiving end is confined by the minimum network transmission only. In other words, neither frame synchronization delay for multiple path delivery, nor retransmission delay is allowed.

To achieve the objectives stated above, Interleaved Source Coding (ISC) is proposed and introduced in this deliverable. ISC codes a single video sequence into multiple sub-sequences based on the network condition and transmits them over a single erasure channel. The objective of ISC is to minimize the frequency and impact of the cascaded effect of packet losses and related propagation of errors resulted from the predictive nature of predictive video coders. Particu

imp

mer

cas

vide

The

(MD

sign

elim

know

leve

less

miss

In th

eras

from

terle

terlea

ticularly, the target is to design and optimum interleaving method such that the impact of losses caused by a given erasure channel model (with memory or memoryless) is limited to a minimum number of video frames. In addition, in case of decoder failed frame replacement, frozen frames, ISC presents smoother video compared to the non-interleaving method.

The proposed ISC video coding differs from previous Multiple-Description-Coding (MDC) based methods (e.g., ones proposed in [24-28]) since ISC is primarily designed for transmission of encoded sequences over a single channel. This eliminates channel selection, content distribution, and synchronization issues known to present with MDC [24-28] . In addition, interleaving could reduce the level of coding inefficiency that normally characterizes MDC coding. Nevertheless, the proposed interleaved coding framework can be generalized for transmission over multiple channels, and hence, it could include some form of MDC. In this research, however, the main focus is on interleaved coding for the single erasure-channel case. Furthermore, the proposed ISC framework is different from other interleaving frameworks [16, 46, 52] since ISC is based on *Frame* interleaving where others are based on *packet* or *cell* interleaving. To find an interleaving set, a Markov Decision Process (MDP) [34, 35, 37, 39, 40] and a Dynar

moc

com

side

to e

amo

dela

pose

3.1 (

Trac

num

that

GOV

The

dictiv

ers in

howe

namic Programming algorithm [34-37] in association with a realistic packet loss model are employed [4, 5]. In addition, some coarse measure of the temporal correlation among pictures within a given video sequence is also taken into consideration. This temporal correlation results in interleaving sets that are unique to each video sequence. However, since measuring the temporal correlation among video frames may not be always feasible for realtime applications due to delay, complexity, and memory constraints, a generic correlation model is proposed as well in case where the actual correlation cannot be computed.

3.1 General Interleaving

Traditional predictive video coding partitions a single lengthy sequence into a number of shorter length Group Of Video object planes (GOVs). It is well known that this partitioning limits the impact of possible errors or losses into individual GOVs.

The proposed *interleaved source coding* (ISC) is a pre- and post-process of predictive source coders (Figure 2). It is possible to integrate *Interleavers and Mergers* into the predictive source coders and use a single encoder and decoder; however, to simplify ISC adaptation, we employ ISC as a pre- and post-process c

k

c

Br

a

te

en

str

IS(

the

tra

of the coders and leave the coders untouched. ISC reduces the impact of losses within a given GOV and improves the overall quality of predictive video over lossy packet networks.

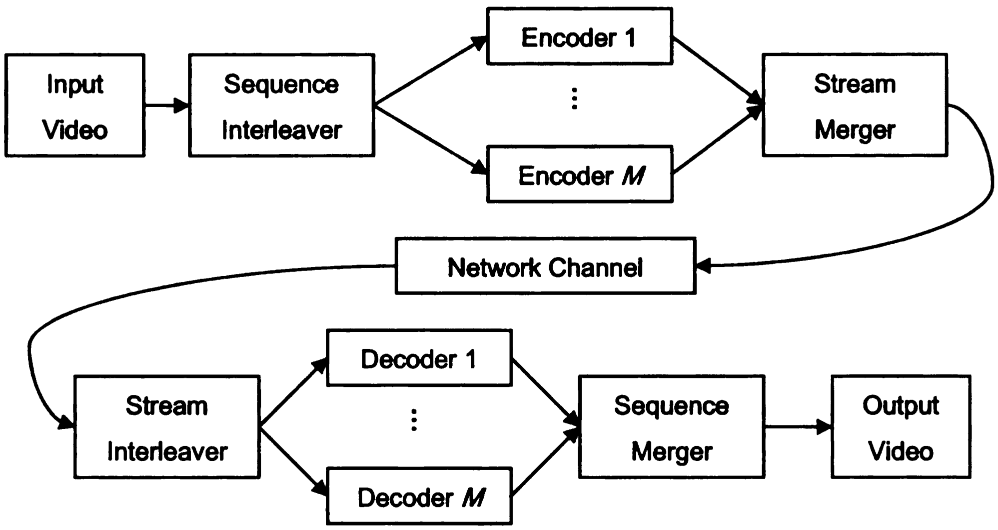


Figure 2 Interleaving of Predictive Video Coding
Illustration is given for two substream interleaving.

Brief description of the overall ISC process is the following: First, ISC separates a single video sequence into M multiple sub-sequences using a *Sequence Interleaver*, and the resulting sub-sequences are encoded using separate video encoders. Then, a *Stream Merger* merges the encoded frames into a single stream in the original-sequence frame order for transmission. In addition to the ISC merged-stream, information regarding the interleaving pattern employed by the encoder must be transmitted to the decoder prior to the ISC merged-stream transmission. At the decoder side, the interleaving pattern is used by a corresponding pair of *Stream Interleaver* and *Sequence Merger*. Hence, the decoder side's *Stream Interleaver* separates the incoming frames or associated packets into M sub-streams according to the transmitted interleaving pattern information. The separated streams are decoded independent to each other and the *Sequence Merger* finalizes the process by merging the sub-sequences' frames into the proper order for playback.

When separating a single sequence into M sub-sequences, $s^{(j)}$, represented by an index set, $j = \{1, 2, \ldots M-1, M\}$, interleaving set can be found with;

$$s = \left\{0 \quad 1 \quad \cdots \quad M \times N - 1\right\} = \bigcup_{j=1}^{M} s^{(j)}, \qquad \bigcap_{j=1}^{M} s^{(j)} = \varnothing, \qquad \forall j, size\left(s^{(j)}\right) = N \quad (1)$$

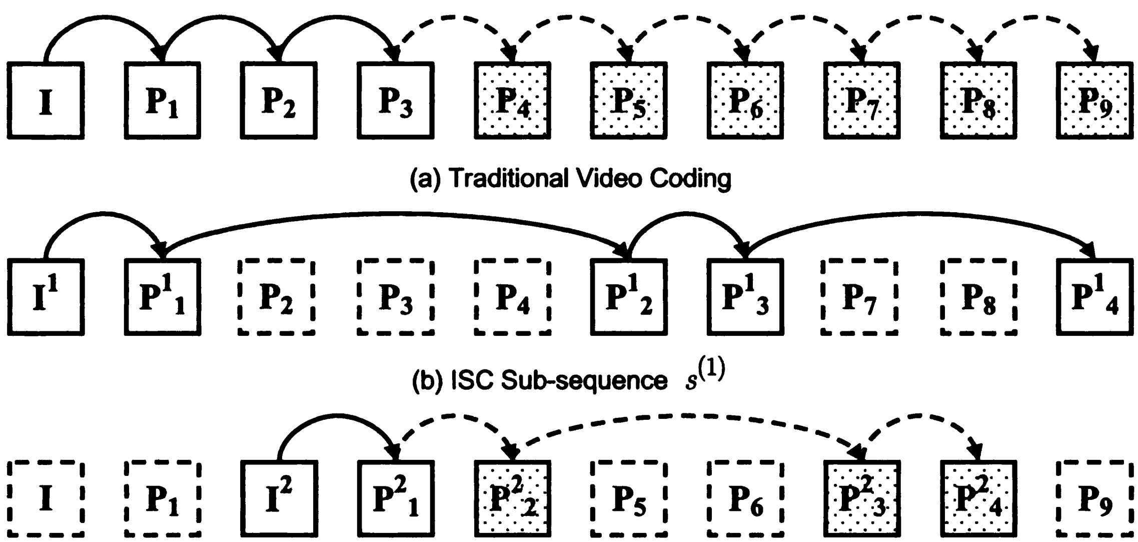
In addition, for true interleaving, we adopt the following ISC interleaving constraints;

$$\sum \left\{s^{(j)}(2) - s^{(j)}(1), \cdots, s^{(j)}(N) - s^{(j)}(N-1)\right\} > N - 1 \qquad (2)$$

where $(M \times N - 1)$ is the number of frames in the original non-interleaved sequence s. In practice, $(M \times N - 1)$ could be the number of frames in a GOV, and hence, the same interleaving is applied to all GOVs in the sequence or a scene.

For example, in two sub-stream case, for a non-interleaved sequence with a GOV size of 10, let $\mathbb{S} = \left\{s^{(1)}, s^{(2)}\right\}$ be an interleaving sub-sequence set with

$s^{(1)} = \{0 \quad 1 \quad 5 \quad 6 \quad 9\}$ and $s^{(2)} = \{2 \quad 3 \quad 4 \quad 7 \quad 8\}$ (Figure 3).



(a) Traditional Video Coding

(b) ISC Sub-sequence $s^{(1)}$

(c) ISC Sub-sequence $s^{(2)}$

Figure 3 Traditional vs. ISC Video Coding

Packet loss in the frame location of P_4 in (a). The arrowed lines represent the coded frames temporal dependencies in the predictive video coding. The dotted frames are the decoder failed frames due to the loss. The shaded frames are belonged to the other sub-sequence in (b) and (c).

Here, the numbers in $s^{(j)}$ represent the frame locations in the non-interleaved sequence and the coded stream's frame transmission order. This interleaving information is required to be transmitted (e.g., as metadata) with the coded substreams as stated previously. Once separated, the sub-sequences are encoded as $I^1P_1^1P_2^1P_3^1P_4^1$ and $I^2P_1^2P_2^2P_3^2P_4^2$ for $s^{(1)}$ and $s^{(2)}$, respectively, and they are transmitted in the following order: $I^1P_1^1I^2P_1^2P_2^2P_2^1P_3^1P_3^2P_4^2P_4^1$; in other words, the merged coded sequence is transmitted in the same frame transmission order of the non-interleaved traditional video coder.

During transmission, if a packet is lost that, for example, is a part of the 5th frame (P_4, in Figure 3-(a)), all 6 frames from P_4 to P_9 of the non-interleaved coding are impacted severely and would not be decoded correctly. However, with interleaving, all the frames in sub-sequence $s^{(1)}$ are decoded successfully and only three frames, P_2^2, P_3^2, and P_4^2, from the sub-sequence $s^{(2)}$ are not decoded. Hence interleaving improves overall playback quality by limiting errors (due to packet losses) to $s^{(2)}$.

Since the formation of the optimal interleaving set could vary depending on the channel model and the transmitting sequence, a problem rises here in choosing the optimal set from the set of all possible interleaved sequences. Let K be the set of all possible interleaving sets for a given GOV size. The size of the set K can be expressed as follows:

$$size(K) = \prod_{i=0}^{M-2} \frac{(M \times N - i \times N)!}{(M-i)(M \times N - (i+1) \times N)!N!} \tag{3}$$

Table 1 Number of Possible Interleaving Set, K, from (1) for $M = 2$

GOV SIZE	10	12	14	16	18	20
SIZE K	126	462	1716	6435	24310	92378

As shown in Table 1 , the size of the set K could be quite large for any reasonable GOV size $(2N-1)$. Hence, identifying the optimum interleaving set that produces the best quality decoded video transmitted over a lossy network channel could be very computationally expensive task due to the vast size of K (Table 1). Therefore, an efficient decision-based search algorithm is required to choose and optimal interleaving set that gives the best quality video for a given erasure-channel model and a video sequence.

3.2 Interleaved Source Coding over Binary Erasure Channel (ISC-BEC)

3.2.1 Binary Erasure Channel

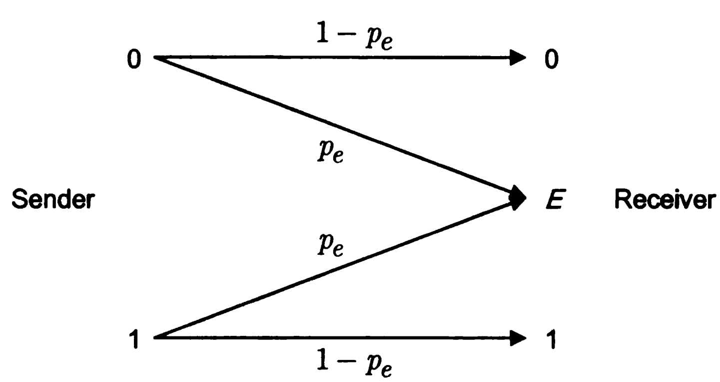


Figure 4 Binary Erasure Channel's Channel Model

Binary Erasure Channel (BEC) model is one of the simplest communication channel modeling methods for data erasure channel (Figure 4). This model characterizes the bit erasure phenomenon caused by channel noise or other data erasure factors, with the erasure probability p_e. The distribution of the binary erasure channel is known to be Independent and Identically Distributed (i.i.d.), hence data arrive in memoryless fashion and the channel capacity is known be $1-p_e$.

Even though the modeling of BEC is based on the binary data bit transmission, this can simply expanded to model packet erasure channel as well. For packet erasure channel, the bit erasure probability can be translated as the packet erasure probability, whereas the channel capacity is same as that of the BEC.

3.2.2 Optimal Interleaving with Reward based Decision Process (RDP)

For the transmission of a predictive coded (and packetized) sequence over a packet erasure channel, the aggregated reward $v(n-1)$ can be defined as a function of the number of transmitted packets. For any final time $n-1$, in other words, after n packet transmissions, define *stage* m as m time units before final time, i.e., as time $n-1-m$ in Figure 5.

0	1	2	$\cdots$	$n-3$	$n-2$	$n-1$	*Time*
$n-1$	$n-2$	$n-3$	$\cdots$	2	1	0	*Stage*

$$\Rightarrow$$

0	1	2	$\cdots$	$n-3$	$n-2$	$n-1$	*Time*
0	1	2	$\cdots$	$n-3$	$n-2$	$n-1$	*Stage*

Figure 5 View Alternation of Stages

Hence, if the final time is time $n-1$, the stage m corresponds to time $n-1$.

For the aggregated reward $v(n-1)$ represents the performance of predictive sequence transmission over a erasure channel with a channel's packet erasure rate p_e.

$$p = \begin{bmatrix} Good & Erasure \end{bmatrix}^T = \begin{bmatrix} 1-p_e & p_e \end{bmatrix}^T \tag{4}$$

To ease the matrix computation, p is transformed into the diagonal matrix P.

$$P = diag(p) = \begin{bmatrix} 1-p_e & 0 \\ 0 & p_e \end{bmatrix} \tag{5}$$

$$v(0) = r = \begin{bmatrix} r_{Good} & r_{Erasure} \end{bmatrix}^T \tag{6}$$

$$v(1) = r + Pv(0) \tag{6}$$

$$\begin{aligned} v(n-1) &= \begin{bmatrix} v_{Good}(n-1) & v_{Erasure}(n-1) \end{bmatrix}^T \\ &= r + Pv(n-2) \\ &= r + Pr + P^2 r + \cdots + P^{n-2} r + P^{n-1} r \\ &= \left(1 + \sum_{i=1}^{n-1} P^i\right) r \end{aligned} \tag{8}$$

For example, in BEC, if the instant rewards are $\{r_{Good}, r_{Erasure}\} = \{1, 0\}$, the reward process is awarded with *1* for a successful packet and *0* for a lost packet during the transmission. In this case, after n packet transmissions, the aggregated rewards, $v_i(n-1)$, represent the expected number of good packet transmissions with the initial packet transmission at state i, *Good or Erasure.*

To reflect ISC scheme into the above reward process, a decision process is employed to find an interleaving set that is most suitable for a given channel condition. In general, the objective is to maximize the number of frames (or associated packets) that can be decoded *correctly*. Hence, the following decision process could guide us toward an "optimal" interleaving for a given erasure channel model that achieves our objective; the interleaving set that provides the highest sum of aggregated reward. In this decision process, a set of *discount factors*, γ_a is applied. The discount factors decide the amount of aggregated reward to be propagated to the next stage. Incorporating equation (8) with the discount factors gives an aggregated reward equation (9), where each coded frame in a GOV is considered to be a single packet transmission iteration in BEC.

$$
\begin{aligned}
v(n-1) &= r_{a(n-1)} + diag(\gamma_{a(n-1)})Pv(n-2) \\
&= r_{a(n-1)} + diag(\gamma_{a(n-1)})Pr_{a(n-2)} + \\
&\cdots + \left(\prod_{i=2}^{n-1} diag(\gamma_{a(i)})\right)P^{n-2}r_{a(1)} \cdots + \left(\prod_{i=1}^{n-1} diag(\gamma_{a(i)})\right)P^{n-1}r_{a(0)}
\end{aligned}
\tag{9}
$$

In ISC, one of the two actions, Coding (C), or Skip (S), is taken for each iteration where $a(n)$ denotes an action taken for the n^{th} frame in a GOV. Let the set of ISC sub-sequences in Figure 3 be the interleaving set k, where $k \in K$. With respect to k, an ISC set is written as $\mathbb{S}^{(k)} = \left\{s^{(k,1)}, s^{(k,2)}\right\}$. In ISC model, each sub-sequence has its own Intra-coded I frame. It is possible to have a single I framed shared among the interleaved sub-sequences, the main design issue will be the interleaving of the predictive frames within the sequence GOVs. Consequently, the frame numbers are rewritten so that each sub-sequence's reward computation starts from the time instance *0* and backward.

$$s^{*(k,j)}(n) = s^{(k,j)}(n) - s^{(k,j)}(0), \quad \text{for } 0 \le n \le N-1 \tag{10}$$

Hence, $\mathbb{S}^{(k)}$ from Figure 3 are $\left\{s^{*(k,1)}(0) \ \cdots \ s^{*(k,1)}(N-1)\right\} = \{0\ 3\ 4\ 8\ 9\}$ and $\left\{s^{*(k,2)}(0) \ \cdots \ s^{*(k,2)}(N-1)\right\} = \{0\ 1\ 4\ 5\ 6\}$. For each sub-sequence, frames are coded, or in other words, action C is performed at frame locations specified in $s^{*(k,j)}$. When the difference between two adjacent numbers in $s^{*(k,j)}$ exceeds 1, which indicates the presence of skipped frames, action S is performed for the frames in location l.

$$\begin{gathered} l = \bigcup_{n=0}^{N-1} \left\{s^{*(k,j)}(n) + 1, \cdots, s^{*(k,j)}(n+1) - 1\right\} \\ \forall n \quad | \quad s^{*(k,j)}(n+1) - s^{*(k,j)}(n) > 1 \end{gathered} \tag{11}$$

This gives the action sets $a^{(k,j)}$ for the interleaving set $\mathbb{S}^{(k)}$ from Figure 3 as $a^{(k,1)} = [C\ S\ S\ C\ C\ S\ S\ S\ C\ C]$ and $a^{(k,2)} = [C\ C\ S\ S\ C\ C\ C]$.

The instant reward For ISC over BEC, the discount factors for the coded frames, action C, are $\gamma_c = \{\gamma_{Good}, \gamma_{Erasure}\} = \{1,0\}$, since packet *Erasure* forces the decoder to stop and no further decoding is possible, hence aggregated reward is not propagated unless the decoder is restarted.

Therefore, the proposed aggregated reward equations for single-packet-per-frame are:

$$v^{(k,j)}\left(s^{*(k,j)}(0)\right) = r_C \tag{12}$$

$$v^{(k,j)}\left(s^{*(k,j)}(n-1)\right) = r_C + diag(\gamma_C)Pv^{(k,j)}\left(s^{*(k,j)}(n-2)\right), \forall n \in s^{*(k,j)} \tag{13}$$

This is valid since the reward aggregation for skipped frame sections can be ignored due to memoryless nature of BEC, each frame arrived independently. When coded sequences are packetized, the number of packets per frame varies with the bitrate and frame rate of the encoder, and the packet size. In addition, within a sequence, the number of packets per frame varies depending on the coding type, (e.g., Intra-frame coding (*I*-frame) and Inter-frame coding (*P*-frame)), and the motion of the sequence. Therefore, due to the unpredictability of the variation of the number of packets per each coded frame, an average number of packets per frame η is used and the aggregated reward equations are as follows.

$$\eta = \left\lceil \frac{bitrate}{framerate \times packetsize} \right\rceil \tag{14}$$

$$v^{(k,j)}\left(s^{*(k,j)}(n)\right) = r_C + diag(\gamma_C)P^{\eta}v^{(k,j)}\left(s^{*(k,j)}(n-2)\right), \quad \text{for } 1 \le n \le N-1 \tag{15}$$

The term P^{η} is multiplied to the aggregated reward since a frame is decoded if and only if all the packets in the coded frames are successfully transmitted. For each interleaving set k, the sum of aggregated rewards gives corresponding expected number of successfully decoded frames.

$$v^{(k)} = \sum_{j=1}^{M} \sum_{n=0}^{N-1} v^{(k,j)} \left(s^{*(k,j)}(n) \right) \tag{16}$$

Hence, the set of aggregated rewards is expressed as:

$$\begin{aligned} V^{(k)} &= \bigcup_{j} V^{(k)}(s^{(k,j)}) \\ \text{where } V^{(k)}(s^{(k,j)}(n)) &= v^{(k,j)}\left(s^{*(k,j)}(n)\right), \\ \text{for } 0 \le n \le N-1 \end{aligned} \tag{17}$$

With the following equation, an interleaving set k is found such that our decision criteria is satisfied, a set with the highest aggregated reward.

$$\arg\max_{k} \left[v^{(k)} \right] \tag{18}$$

3.2.3 ISC over BEC with Frame Correlation

In predictive video coding, when the decoder encounters a packet loss (or errors in a transmitted packet), to continue the smooth video presentation (without blank screen or distorted frames), a playback application often replaces the decoder failed frames with the last successfully decoded frame until a successfully decoded frame arrives to restart the decoding process. Here, we refer to this last

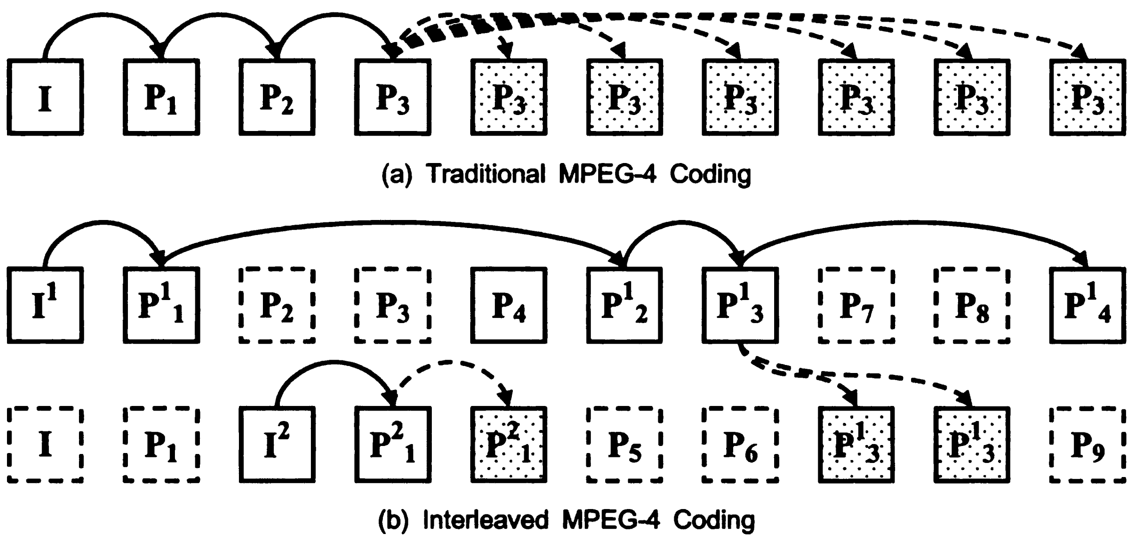


Figure 6 Frame Replacement Illustrations

Packet loss in the frame location of P_4 in (a). The dotted arrowed lines represent the frame replacement relationship for the decoder failed frames (dotted frames).

successfully decoded frame as the *replacement* frame. When the decoder failed frames are replaced, the *distances* (in terms of number of pictures) between the replacement frame and the replaced frames have effects on the smoothness of the sequence flow and the overall quality of the playback sequence. This is due to the fact that the shorter distance between the replacing frames indicates highly correlated frame replacement in place of decoder failed frames. Figure 6 illustrates the frame replacement actions in case of decoder failure.

Table 2 Average Frame Replacement Distance with a Single Lost Packet in a GOV

GOV SIZE	10	12	14	16	18	20
NON-ISC	4.0000	4.6667	5.3333	6.0000	6.6667	7.3333
ISC	2.8265	2.9686	3.0740	3.1561	3.2230	3.2793

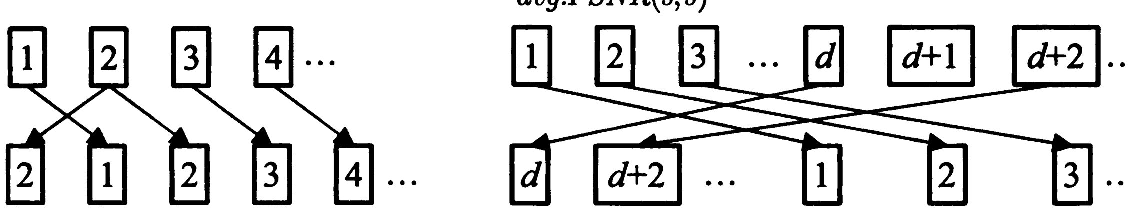
As shown in Table 2 , the average frame replacement distances due to a single lost packet in a GOV is shorter for ISC than the traditional transmission method. Hence it is expected that ISC produces smoother and higher quality video over erasure channels with decoder failed-frame replacements.

To incorporate the quality improvement from frame replacements, correlation gain $g^{(k)}$ is added to equation (16) and a Dynamic Programming is used to find an interleaving set that produces the highest RDP sum of the aggregated reward with the correlation gain $g^{(k)}$.

$$\arg\max_{k}\left[v^{(k)} + g^{(k)}\right] \tag{19}$$

The correlation gain $g^{(k)}$ is computed with the following steps. First, temporal correlations are computed with average PSNR between original sequence and temporally shifted sequences.

$$\rho(d) = \frac{avg.PSNR(s, s+d)}{avg.PSNR(s,s)} \tag{20}$$



(a) Shifted by 1 (b) Shifted by d

Figure 7 Sequence Shifting Illustration for Temporal Correlation Measurement

Figure 7 shows illustration on sequence shifting for the temporal correlation measurement and the correlations are computed with equation(20) .

Second, a curve fitting method with the Minimum Mean Square Estimator (MMSE) is used to obtain a function that represents temporal correlation of a given sequence.

$$\underset{\{a,b,c\}}{\arg\min}\left[MSE\left\{\rho(d), a\times\exp\left(-d^{b}\right)+c, \forall d\right\}\right] \tag{21}$$

Table 3 Distance Matrix $D^{(k)}$ for $\mathbb{S}^{(k)}$ shown in Figure 7 -(b)

	1	2	3	4	5	6	7	8	9	10
1	1	2	0	0	0	1	2	0	0	1
2	0	1	0	0	0	1	2	0	0	1
3	0	0	1	2	3	0	0	1	2	0
4	0	0	0	1	2	0	0	1	2	0
5	0	0	0	0	1	0	0	1	2	0
6	0	0	0	0	0	1	2	0	0	1
7	0	0	0	0	0	0	1	0	0	1
8	0	0	0	0	0	0	0	1	2	0
9	0	0	0	0	0	0	0	0	1	0
10	0	0	0	0	0	0	0	0	0	1

Third, a $2N$ by $2N$ upper triangular distance matrix $D^{(k)}$ (Table 3) is generated for each ISC set k for single-packet-loss per GOV cases, since the main purpose of interleaving method is to isolate decode failure to one sub-sequence. The distance matrices' diagonal indices indicate the first frame location in a GOV impacted by a single packet loss. Hence, the non-zeros entries of the distance matrix represent the distances from replacement frames to the replaced ones. Finally, the correlation gain is computed with the following equations. $W^{(k)}$ is the correlation weight matrix with respect to the distances from replacement frames to the replaced ones. In case of replacements, the weight is multiplied by the aggregated reward of the replacement frame and the discounted reward is given to the replaced frame. $G^{(k)}$ is the correlation computed aggregated reward gain matrix.

$$W_{x,y}^{(k)} = \begin{Bmatrix} a \times \exp\left[-\left(D_{x,y}^{(k)}\right)^{b}\right] + c, & \forall D_{x,y}^{(k)} \neq 0 \\ 0, & \text{otherwise} \end{Bmatrix} \tag{22}$$

$$V^{(k)*} = \left[V^{(k)}(M \times N - 1) \quad V^{(k)}(0) \quad \cdots \quad V^{(k)}(M \times N - 2)\right] \tag{23}$$

$$G_{x,y}^{(k)} = \begin{Bmatrix} W_{x,y}^{(k)} \times V^{(k)*}\left(y - D_{x,y}^{(k)}\right), & \forall D_{x,y}^{(k)} \neq 0 \\ V^{(k)}(y), & \text{otherwise} \end{Bmatrix} \tag{24}$$

$$g^{(k)} = \sum_{x,y=1}^{\text{GOV SIZE}} G_{x,y}^{(k)}, \qquad \forall x, y \leq \text{GOV SIZE} \tag{25}$$

3.2.4 Generalization of the Temporal Correlation Measurement

Measuring the temporal correlation among video frames within a complete GOV may not be always feasible for realtime applications due to delay, complexity, and memory constraints. Therefore, a more generic correlation model may be required for the cases when the actual correlation cannot be computed. Below, we present such a generic model. $VI^{(k)}$ is the set of the reward increments at each sub-sequences' reward calculation iteration.

$$\begin{gathered} VI^{(k)} = \bigcup_{j} VI^{(k)}(s^{(k,j)}) \\ \text{where } VI^{(k)}(s^{(k,j)}(0)) = v^{(k,j)}\left(s^{*(k,j)}(0)\right), \\ VI^{(k)}(s^{(k,j)}(n)) = v^{(k,j)}\left(s^{*(k,j)}(n)\right) - v^{(k,j)}\left(s^{*(k,j)}(n-1)\right), \\ \text{for } 1 \le n \le N-1 \end{gathered} \tag{26}$$

With respect to $D^{(k)}$ and $VI^{(k)}$, the weight matrix $W^{(k)*}$ is calculated with the following equation. Here, $\left(\left(D^{(k)*} \times VI^{(k)^T}\right) \div \sum_{y=1}^{\text{GOV SIZE}} D_{x,y}^{(k)*}, \forall x\right)$ is the average reward increment of the successfully decoded frames in case of a single packet loss in a GOV. Since the decoder failed frames are copied by the last successfully decoded frames, multiplying this value by the replacement frame's aggregated reward estimates the correlation-based aggregated reward of the replaced frame. Hence, the decrement is assumed to be exponential with respect to temporal distances from the replacement frames to the replaced ones.

$$W^{(k)*} = \left(\left(D^{(k)*} \times VI^{(k)^T} \right) \div \sum_{y=1}^{\text{GOV SIZE}} D_{x,y}^{(k)*}, \forall x \right)^{\wedge} D^{(k)}$$
$$\text{where } D^{(k)*} = \begin{Bmatrix} 0, & \forall D_{x,y}^{(k)} \neq 0 \\ 1, & \text{otherwise} \end{Bmatrix} \tag{27}$$

$$G_{x,y}^{(k)*} = \begin{Bmatrix} W_{x,y}^{(k)*} \times V^{(k)*}\left(y - D_{x,y}^{(k)}\right), & \forall D_{x,y}^{(k)} \neq 0 \\ V^{(k)}(y), & \text{otherwise} \end{Bmatrix} \tag{28}$$

$$g^{(k)*} = \sum_{x,y=1}^{\text{GOV SIZE}} G_{x,y}^{(k)*}, \qquad \forall x, y \leq \text{GOV SIZE} \tag{29}$$

The optimal interleaving set using the above generic correlation model can be found using the following equation

$$\arg\max_{k} \left[v^{(k)} + g^{(k)*} \right] \tag{30}$$

3.3 Evaluation and Analysis

3.3.1 Simulation Setup

For evaluation, CIF sequences of *Akiyo*, *Foreman*, *Coastguard*, and *Mobile* were coded into an *IPPP...* GOV structure using an MPEG-4 encoder. GOV sizes (un-interleaved size) of 10, 12, 14, 16, 18, and 20 were used to partition the evaluation sequences. Frame rate of 15 frames per second, bitrate of *100kbps*, *500kbps*, and *1Mbps*, and packet size of *512 Byte* are used to represent emerging Internet-access technologies (e.g., Modem, DSL/Cable and LAN connections). For Simulation, CIF sequences, *Akiyo*, *Foreman*, *Coastguard*, and *Mobile*, are encoded using MPEG-4 encoder. Only Intra- (I) and Inter- (P) coded frames are used to form GOV.

When the coded sequences are packetized, to limit the impact of a single packet loss to a single frame, no packets are shared among two consecutive coded frames. (In other words, each packet contains data that belongs to only one video frame.) In addition, partial decoding is not employed for the frames with losses and frozen frames for both ISC and traditional (non-ISC) cases. Three ISC scenarios are simulated: (a) the non-correlation computation model (equation (18)), (b) sequence specific correlation gain computation model (equation (19)), and (b) generic correlation gain computation model (equation (30)). We refer to these scenarios as ISC-BEC-NC (non-correlation model), ISC-BEC-SC (sequence specific correlation model), and ISC-BEC-GC (generic correlation model), respectively. The ISC-BEC-NC scenario generates an optimal interleaved pattern that is independent of the video sequence, and hence, it generates ISC pattern depending on the erasure-channel model only. It is important to note that the ISC-BEC-GC case captures the correlation among frames in a generic sense, and it does not measure correlation based on actual computation of the correlation among the video frames. Hence, the ISC-BEC-GC scenario is mainly dependent on the original GOV size of the video sequence being coded. To simulate a statistically viable experiments and to capture a realistic network loss patterns, ten packet loss traces were generated for each packet erasure rate of 2%, 4%, 6%, ... 18%, and 20%. Each evaluation case is fitted into these packet loss traces and the PSNR values are averaged to provide statistically satisfying results for analysis.

3.3.2 Simulation Results

The simulation results are given in the following order: a) ISC-BEC-NC, b) ISC-BEC-SC, and c) ISC-BEC-GC. The figures show the average PSNR difference between ISC-BEC-XX and non-interleaving single layer MPEG-4 coding with various BEC erasure rates and GOV sizes.

The sequence specific temporal correlations are computed following the illustration Figure 7 and the MMSE coefficients computed for the sequences in Figure 8 with equation (21) are given in Table 4 .

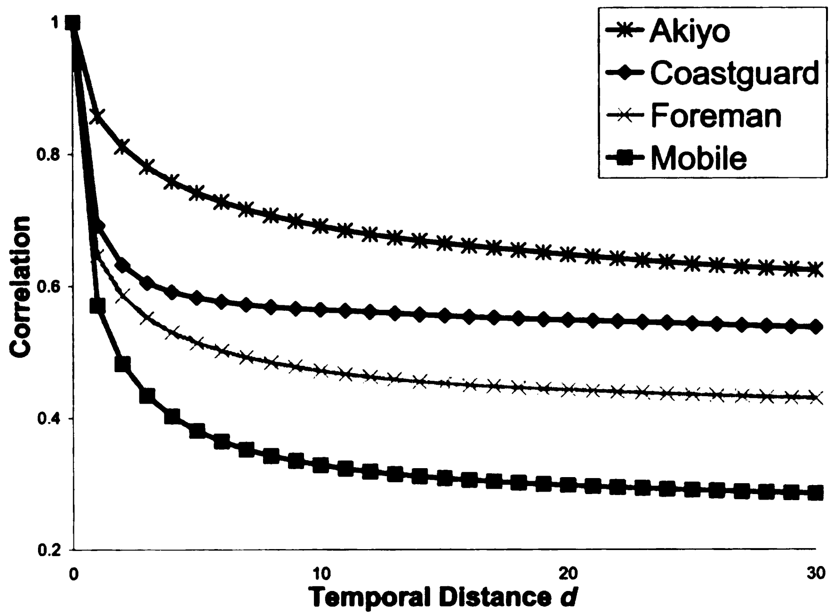
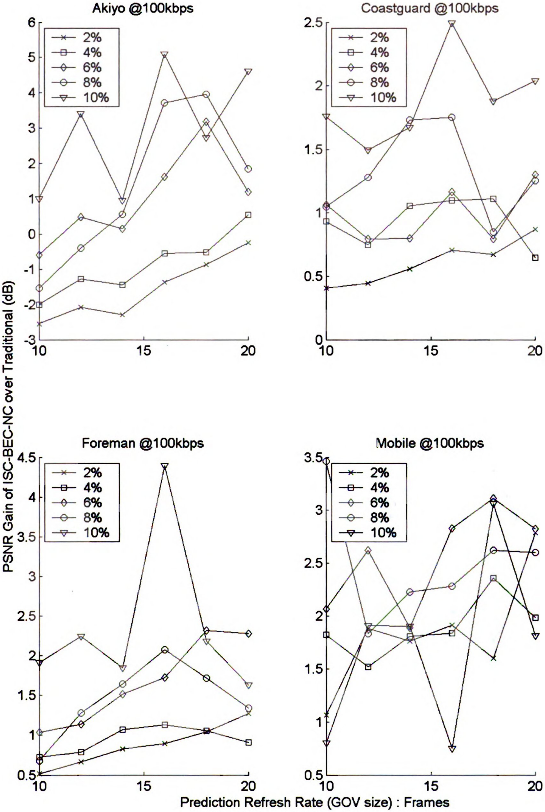


Figure 8 Temporal Correlation of the Evaluation Sequences

Table 4 MMSE Coefficients, $\{a, b, c\}$ for given test sequences

	a	b	c
Akiyo	0.5148	0.2740	0.5633
Coastguard	0.6279	0.2567	0.3609
Foreman	0.6217	0.1251	0.3769
Mobile	0.7262	0.3718	0.2456


Akiyo @100kbps
Coastguard @100kbps
Foreman @100kbps
Mobile @100kbps
2%
4%
6%
8%
10%
PSNR Gain of ISC-BEC-NC over Traditional (dB)
Prediction Refresh Rate (GOV size) : Frames

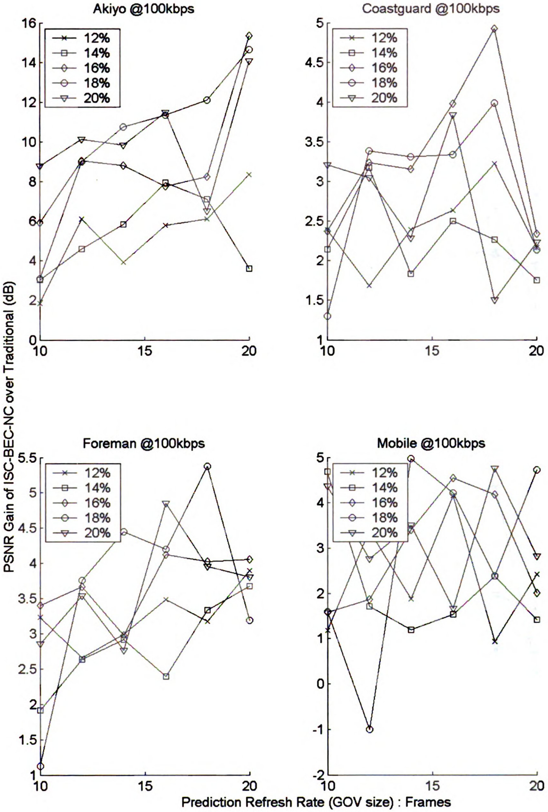
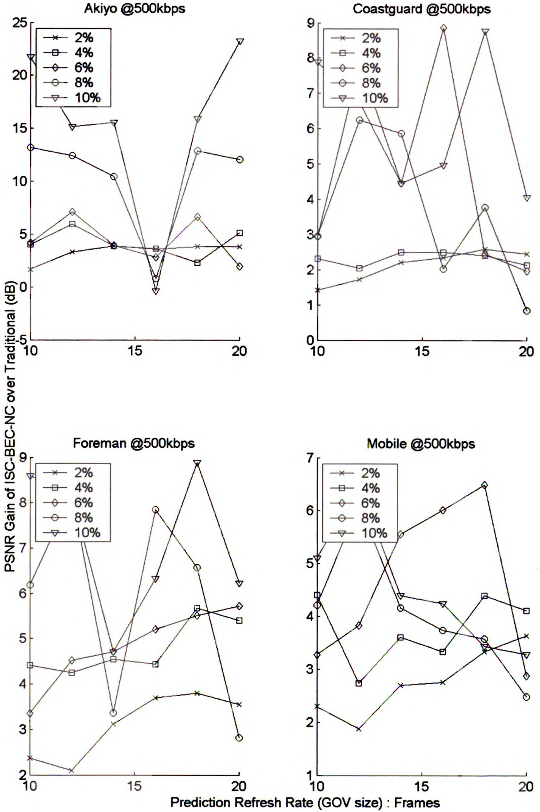


Figure 9 PSNR Differences between ISC-BEC-NC and Traditional *@ 100KBPS*



Akiyo @500kbps
Coastguard @500kbps
Foreman @500kbps
Mobile @500kbps
2%
4%
6%
8%
10%
PSNR Gain of ISC-BEC-NC over Traditional (dB)
Prediction Refresh Rate (GOV size) : Frames

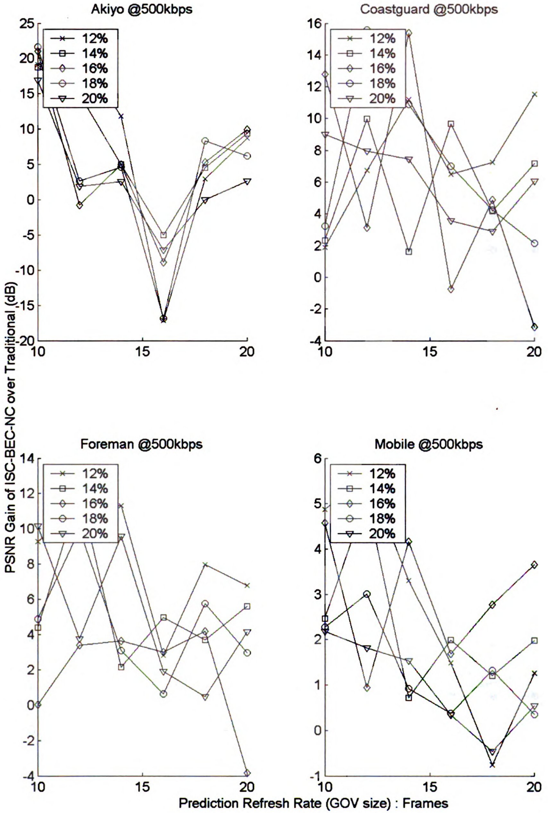
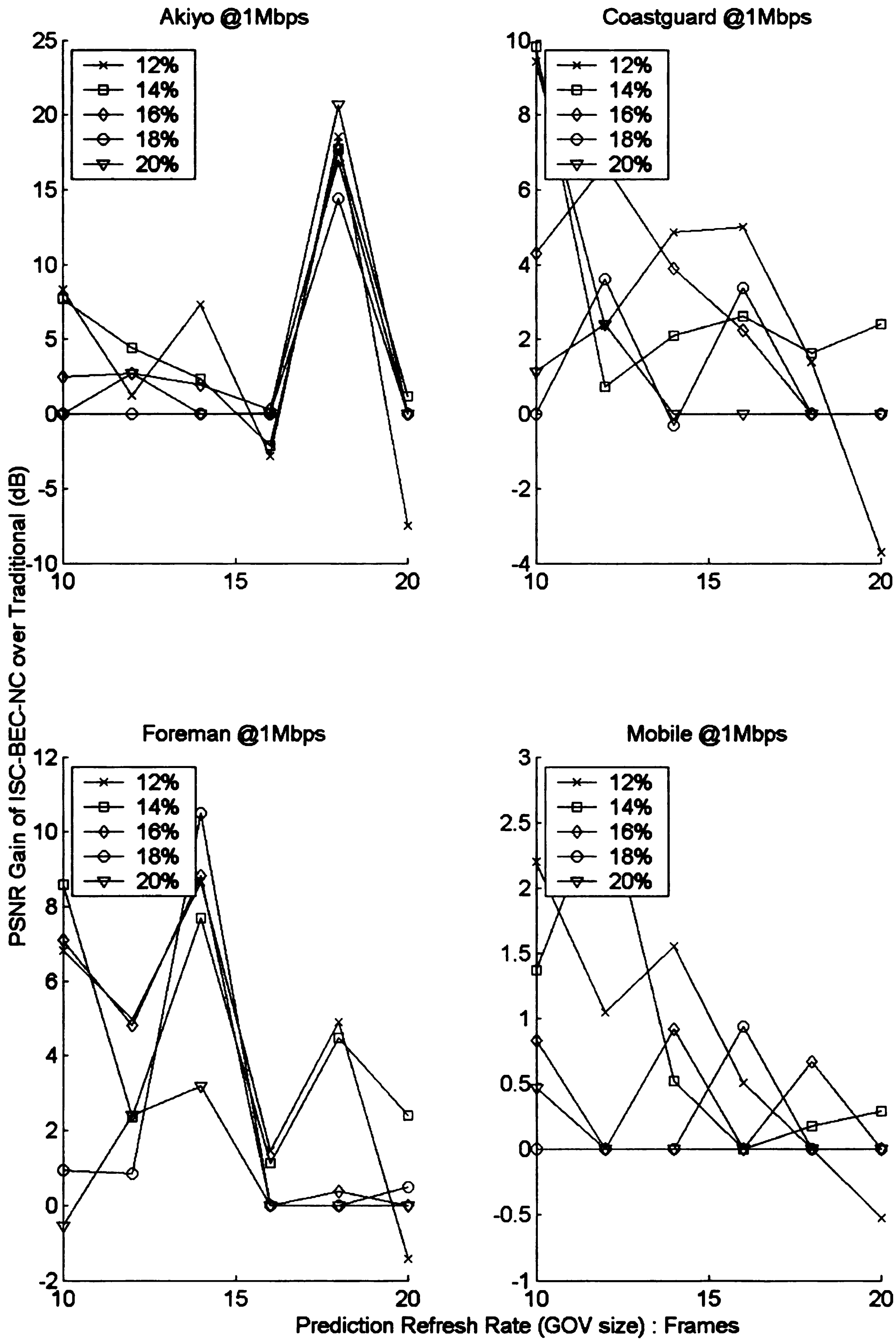


Figure 10 PSNR Differences between ISC-BEC-NC and Traditional *@ 500KBPS*


Akiyo @1Mbps
Coastguard @1Mbps
Foreman @1Mbps
Mobile @1Mbps
12%
14%
16%
18%
20%
PSNR Gain of ISC-BEC-NC over Traditional (dB)
Prediction Refresh Rate (GOV size) : Frames

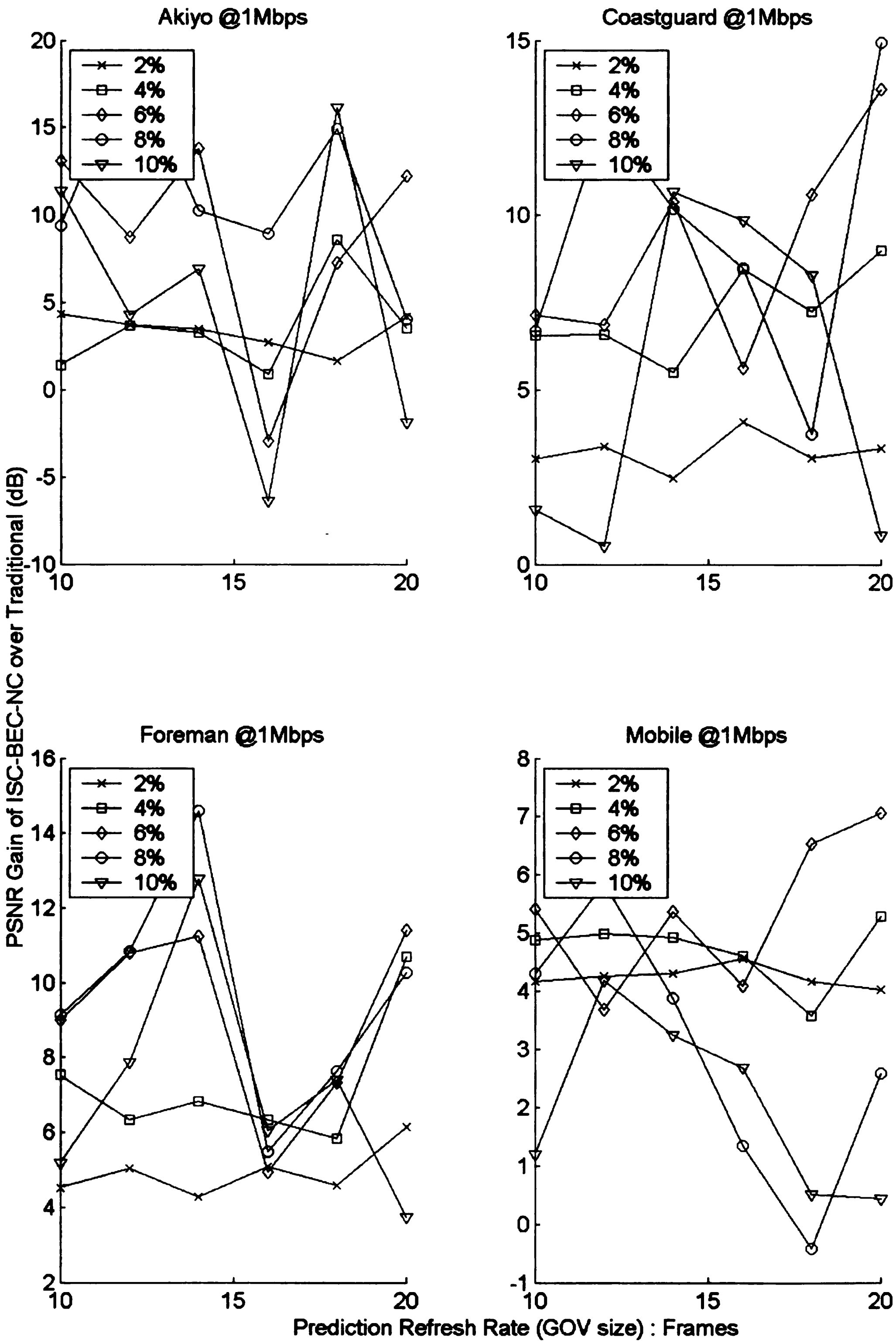
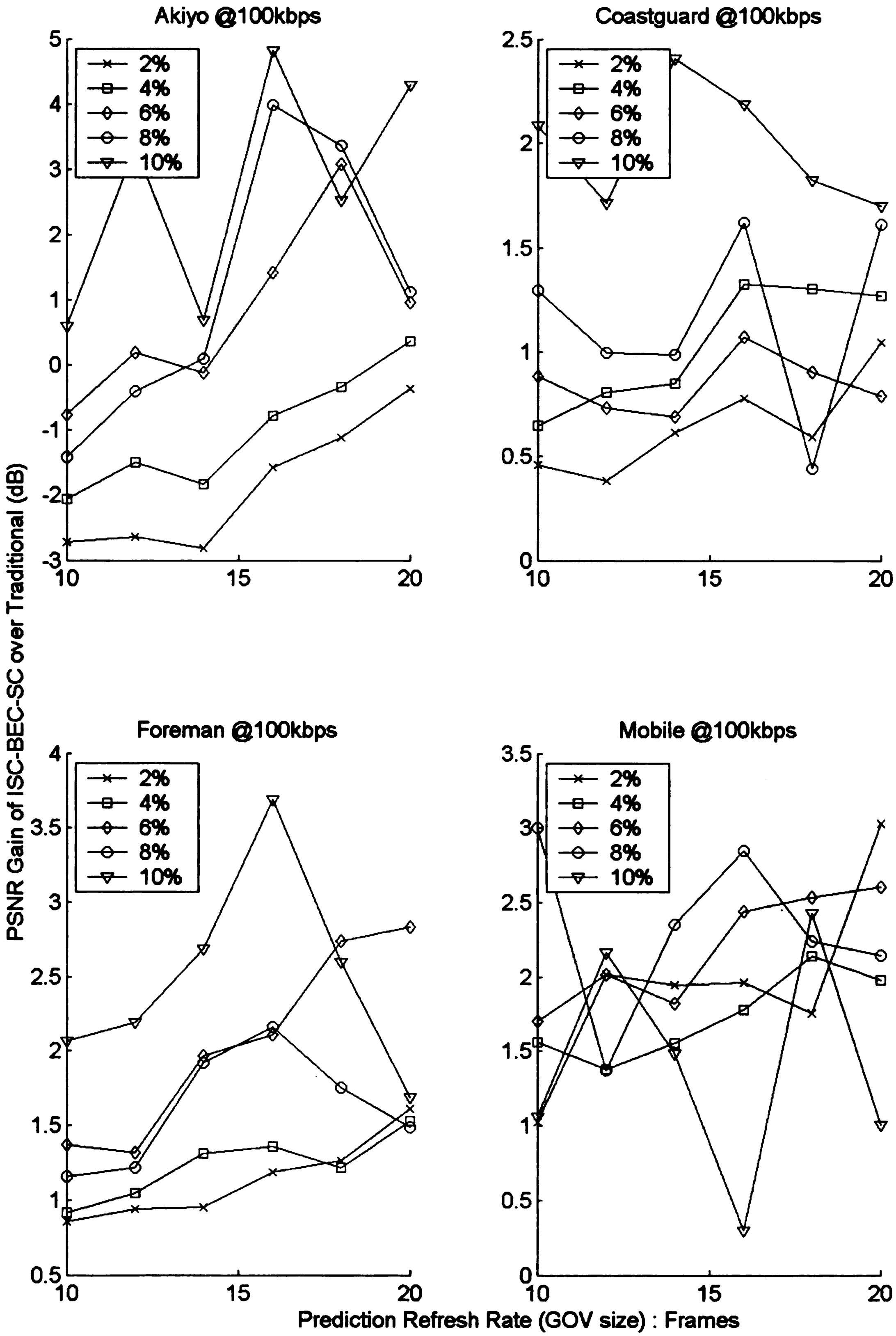


Figure 11 PSNR Differences between ISC-BEC-NC and Traditional *@ 1MBPS*


Akiyo @100kbps
Coastguard @100kbps
Foreman @100kbps
Mobile @100kbps
2%
4%
6%
8%
10%
PSNR Gain of ISC-BEC-SC over Traditional (dB)
Prediction Refresh Rate (GOV size) : Frames

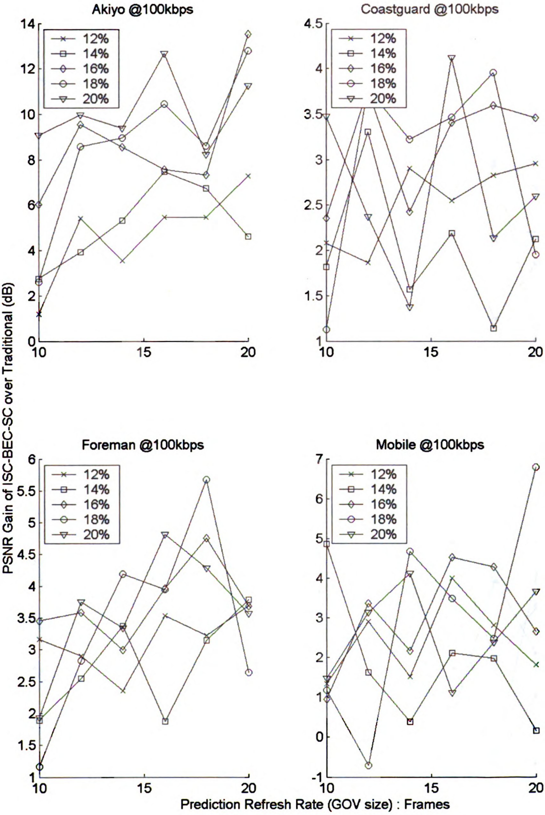
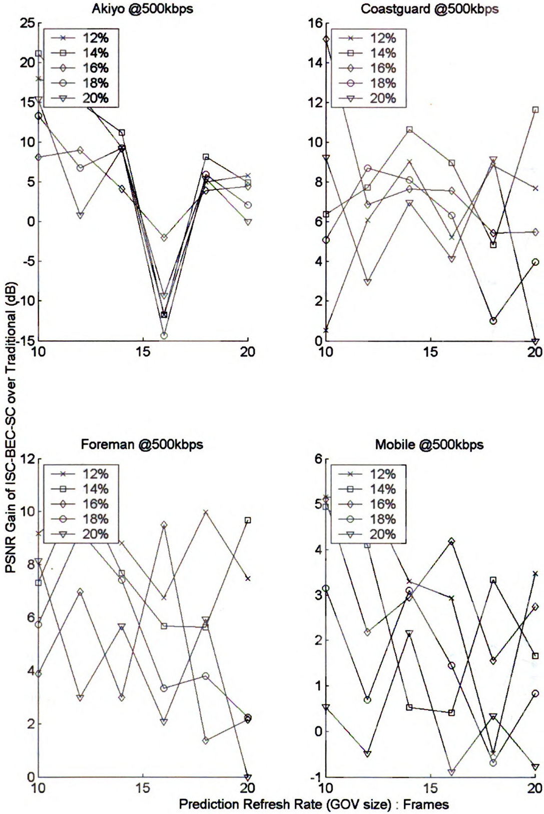


Figure 12 PSNR Differences between ISC-BEC-SC and Traditional *@ 100KBPS*



Akiyo @500kbps
Coastguard @500kbps
Foreman @500kbps
Mobile @500kbps
12%
14%
16%
18%
20%
PSNR Gain of ISC-BEC-SC over Traditional (dB)
Prediction Refresh Rate (GOV size) : Frames

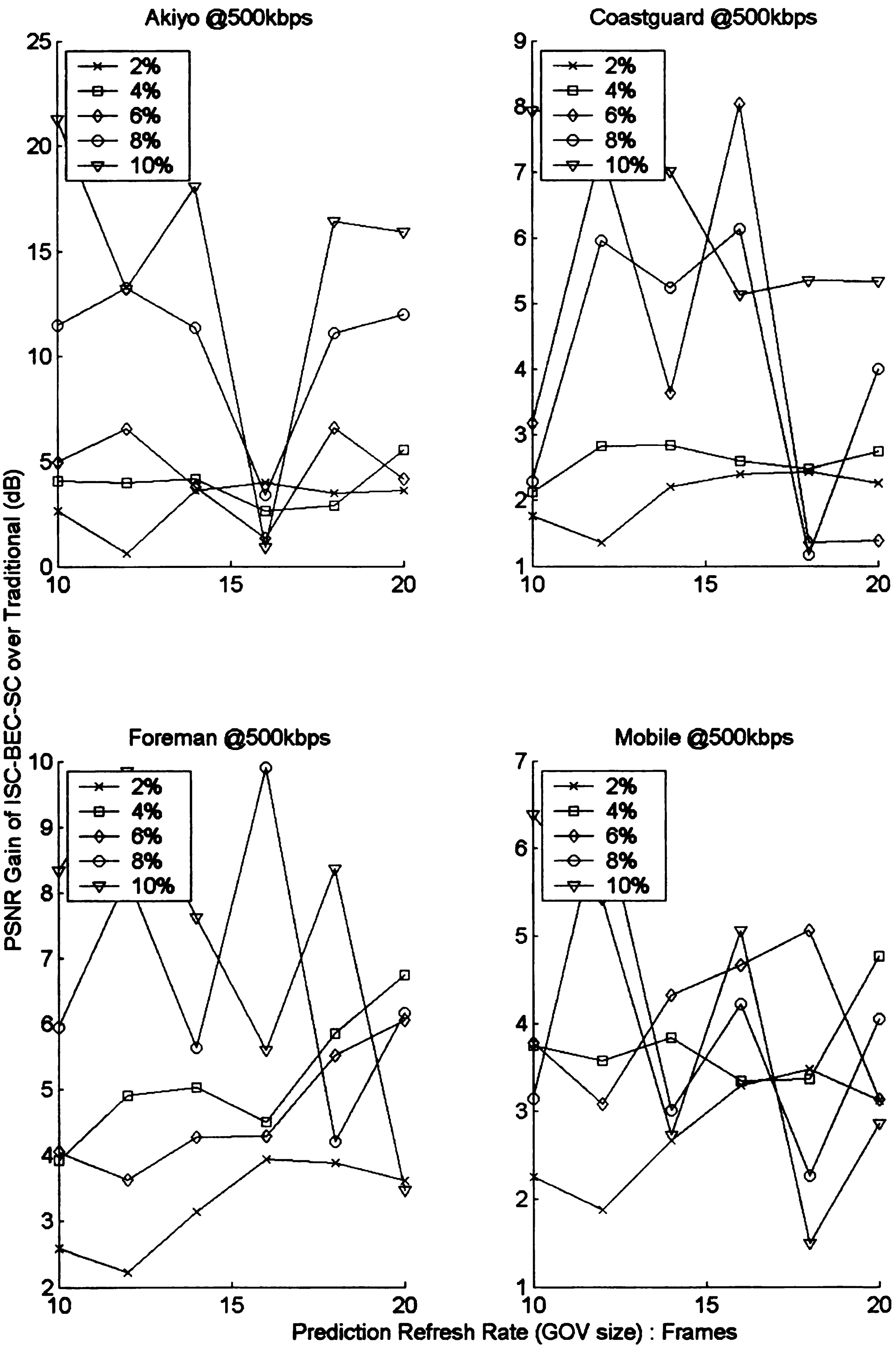
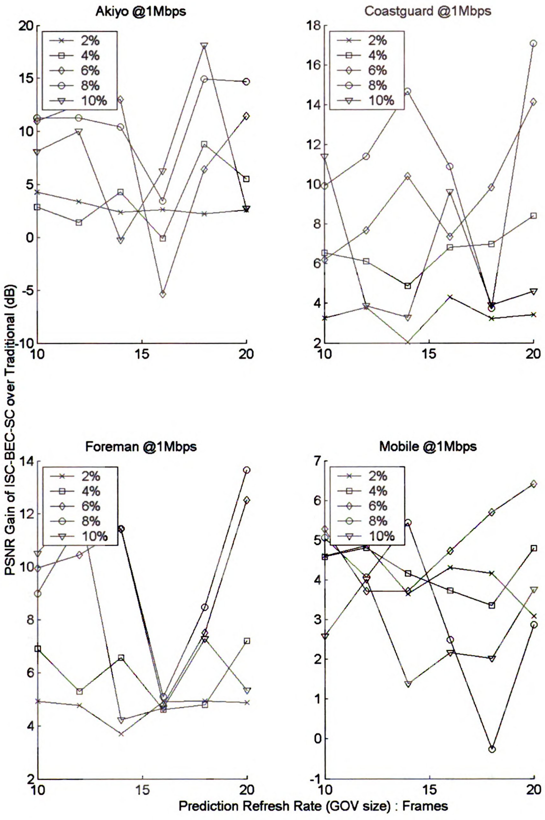


Figure 13 PSNR Differences between ISC-BEC-SC and Traditional *@ 500KBPS*


Akiyo @1Mbps
Coastguard @1Mbps
Foreman @1Mbps
Mobile @1Mbps
2%
4%
6%
8%
10%
PSNR Gain of ISC-BEC-SC over Traditional (dB)
Prediction Refresh Rate (GOV size) : Frames

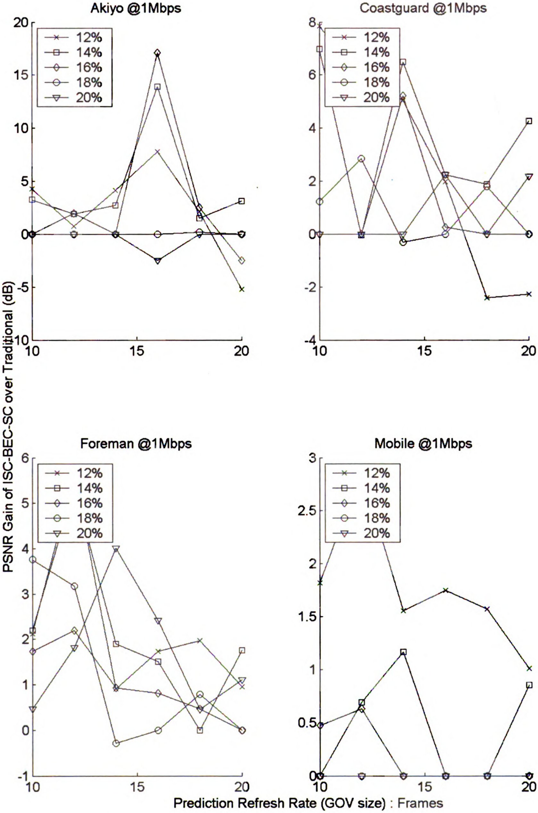
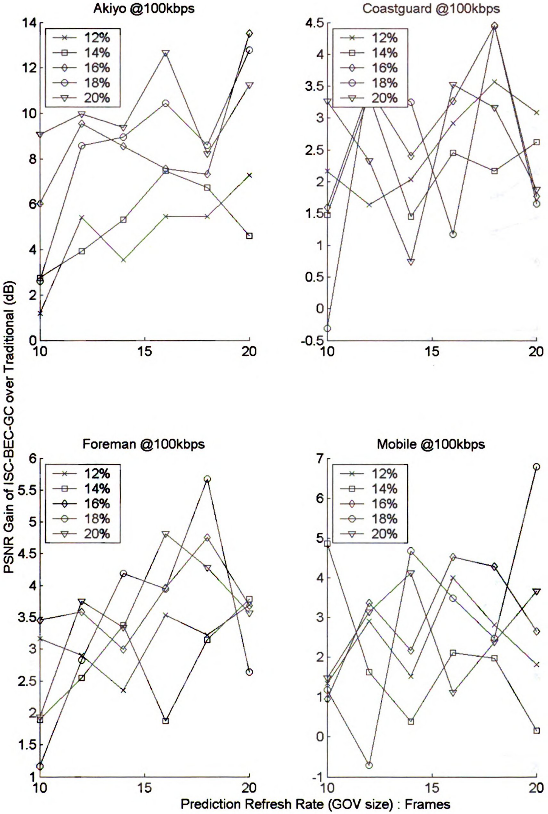


Figure 14 PSNR Differences between ISC-BEC-SC and Traditional *@ 1MBPS*


Akiyo @100kbps
Coastguard @100kbps
Foreman @100kbps
Mobile @100kbps
12%
14%
16%
18%
20%
PSNR Gain of ISC-BEC-GC over Traditional (dB)
Prediction Refresh Rate (GOV size) : Frames

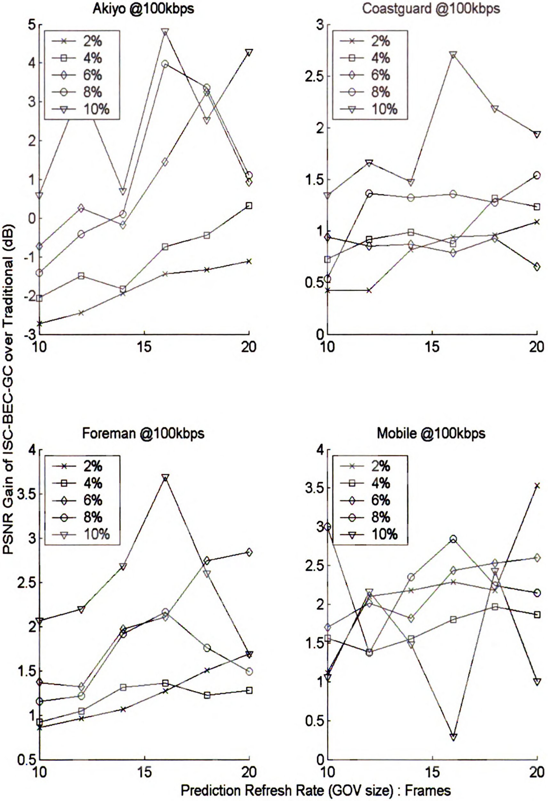
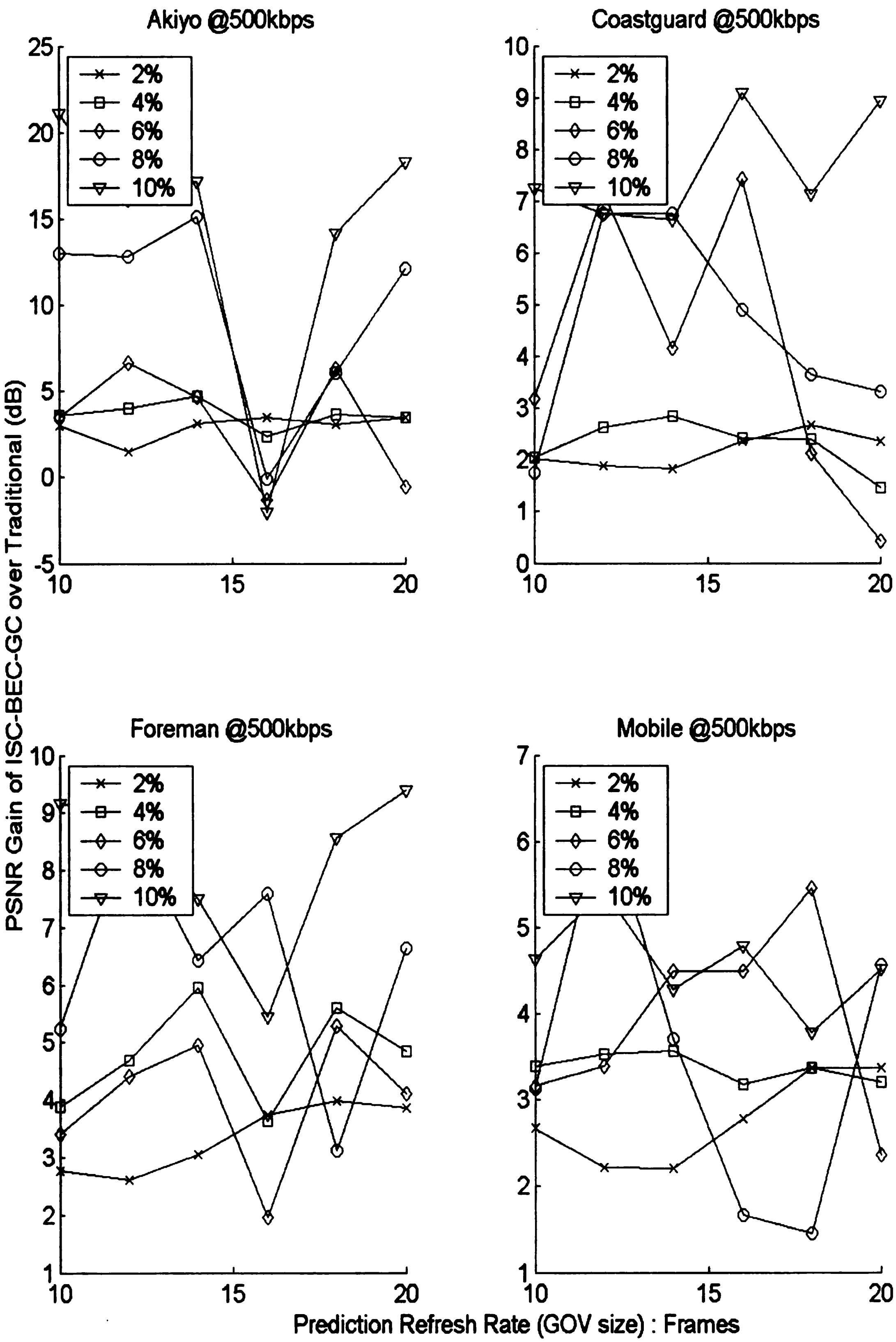


Figure 15 PSNR Differences between ISC-BEC-GC and Traditional *@ 100KBPS*


Akiyo @500kbps
Coastguard @500kbps
Foreman @500kbps
Mobile @500kbps
2%
4%
6%
8%
10%
PSNR Gain of ISC-BEC-GC over Traditional (dB)
Prediction Refresh Rate (GOV size) : Frames

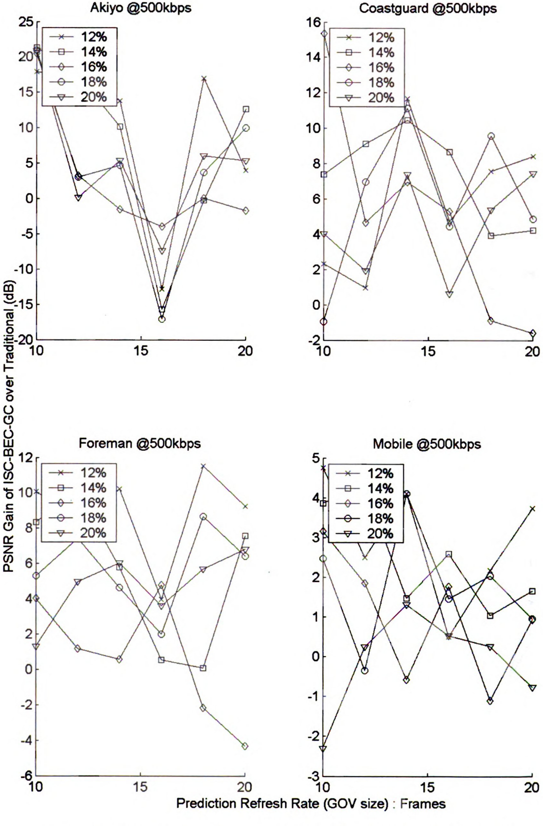
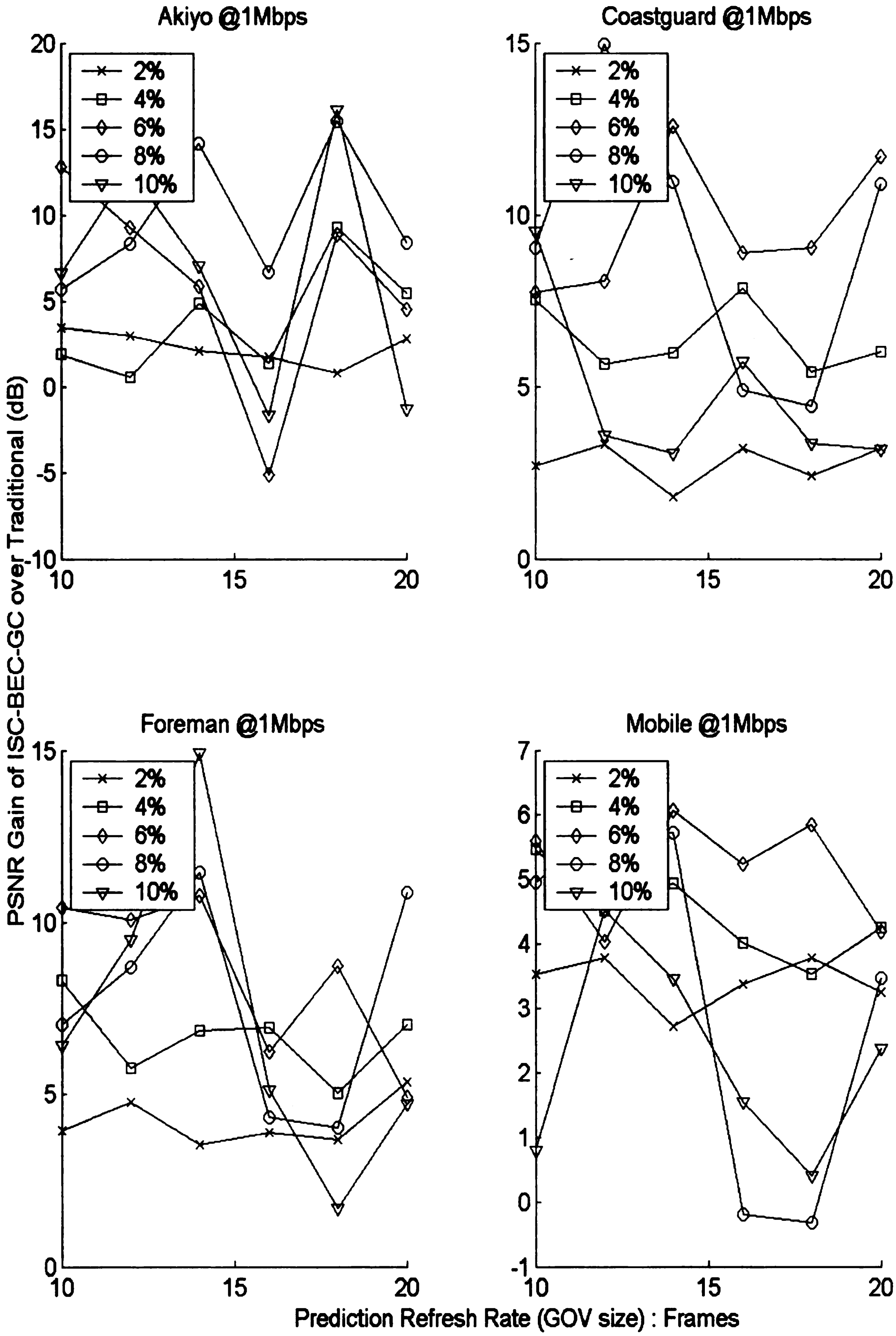


Figure 16 PSNR Differences between ISC-BEC-GC and Traditional *@ 500KBPS*



Akiyo @1Mbps
Coastguard @1Mbps
Foreman @1Mbps
Mobile @1Mbps
2%
4%
6%
8%
10%
PSNR Gain of ISC-BEC-GC over Traditional (dB)
Prediction Refresh Rate (GOV size) : Frames

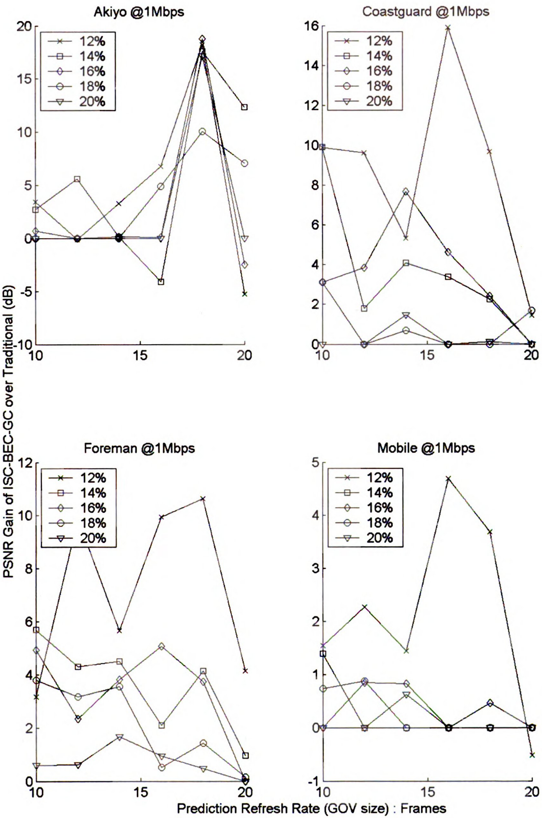


Figure 17 PSNR Differences between ISC-BEC-GC and Traditional *@ 1MBPS*

3.3.3 Analysis

Due to the nature of Binary Erasure Channel, memoryless, hence the packet erasures are mostly isolated single erasures, it is difficult to generalize the performance variance depending on channel erasure rate, GOV size, and correlation implications. However, in most cases, the Interleaved Source Coding framework has shown improvement over traditional non-interleaving single layer coding. Noticeable observation is that when the packet erasure rate is low, i.e., 2% ~ 10%, the ISC-BEC finds the optimum interleaving pattern that is very close to dividing non-interleaving coding scheme's GOV size in half. In fact, if there was no ISC constraint, at least one skipped frame within a substream, ISC would have found the half GOV size of the non-interleaving coding as the optimum interleaving set. Overall observation indicates that under the given channel model, even though the intra-frame coded I frames consumes more bandwidth than the inter-frame coded P frames, the frequent appearance of I frame improves the overall quality of vided when channel is prone to packet losses. In addition, when the different types of correlation models are applied to ISC-BEC, even though the superiority is hard to distinguish, the generic correlation model shows competitive results and yet the use of generic correlation model is acceptable when actual correlation computation is not feasible.

In summary, the Interleaved Source Coding has shown its superiority compared to non-interleaving coding, and ISC adaptation to the realistic channel model, channel with memory, is permissible and feasible.

Chapter 4

Interleaved Source Coding over Channel with Memory

4.1 Interleaved Source Coding over Channel with Memory

4.1.1 Gilbert Channel Model

Previous efforts for the analysis and modeling of packet losses over the Internet (e.g.,[4, 5, 10-14, 16]) and wireless networks (e.g.,[29, 30]) have shown that these losses exhibit Markovian properties.

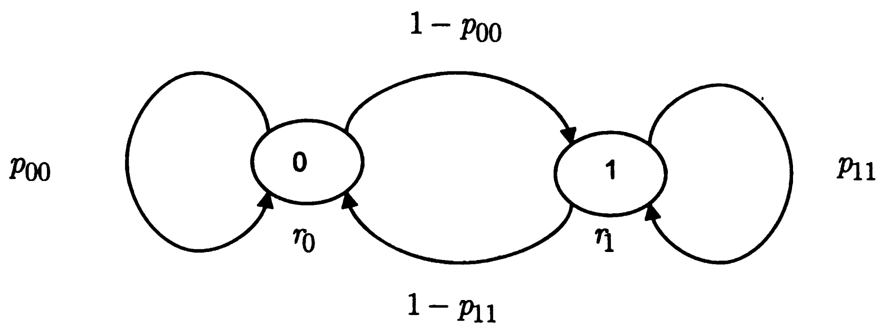


Figure 18 Two state Markov model with rewards r_i

The two state Markov Model, *a.k.a. Gilbert Model,* is proven to replicate an acceptable erasure-channel model (Figure 18). It is possible to use higher order Markov models; however, to reduce computational complexity, two state Markov model is used throughout this research.

In the Gilbert channel model, the steady state probabilities in good state and bad state is represented as following.

$$\pi(0) = \frac{p_{10}}{p_{01} + p_{10}} \text{ and } \pi(1) = \frac{p_{01}}{p_{01} + p_{10}} \tag{31}$$

These values give coarse measure of a given channel's packet transmission behavior. However, for statistical channel modeling, instead of the above probabilities, the transition probabilities p_{01} and p_{10} (or $p_{00} = 1 - p_{01}$ and $p_{11} = 1 - p_{10}$) could be used to characterize Gilbert channel model. Since it is difficult to properly model a Gilbert channel with arbitrary transition probabilities, p_{01} and p_{10}, a more meaningful pair of parameters are the average loss rate, p_1, and the *packet correlation*, ρ; this pair can provide a practical and useful insight of the channel while representing the state transition probabilities.

The average loss rate and the correlation between two consecutive packets can be defined as follows:

$$p_1 = \frac{p_{01}}{p_{01} + p_{10}}, \qquad \rho = p_{01} + p_{10} - 1 \tag{32}$$

Hence, the transition probabilities represented by p_1 and ρ are:

$$\begin{aligned} p_{00} &= 1 - p_1(1-\rho), & p_{01} &= p_1(1-\rho) \\ p_{10} &= (1-p_1)(1-\rho), & p_{11} &= 1-(1-p_1)(1-\rho) \end{aligned} \tag{33}$$

In addition, the steady state probabilities are directly related to the loss rate p_1; $\pi(0) = 1 - p_1$ and $\pi(1) = p_1$. Furthermore, the packet erasure correlation ρ provides an average measure of the correlation of two consecutive packets. In particular, when $\rho = 0$, $p_{01} + p_{10} = 1$, the loss process is memory-less, and the above probability measures reduce to the special case of a memory-less Binary Erasure Channel (BEC). In the sequel, we analyze the impact of the level of correlation among consecutive packets, as represented by ρ, on ISC-based packet video over a wide range of loss rate p_1 values.

4.1.2 Interleaved Source Coding with Markov Decision Process (ISC-MDP)

For a Markov channel model, a *Markov Reward Process* (MRP) (e.g.,[34, 35]) can estimate the system's performance using (a) the Markov channel's transition probabilities based on the packet transmission and (b) some model for the *rewards* that are associated with each system state. This reward-based MRP could be used to measure the system's performance after n packet transmissions, and this, in turn, could guide the design of our ISC coding system (as explained further below).

For the transmission of a predictive coded (and packetized) sequence over a lossy Markov channel with a channel's state transition matrix P, we define the aggregated reward $v(n-1)$ as a function of the number of transmitted packets.

Table 5 Two state Markov transition matrix, P

Current \ Future	0	1
0	p_{00}	$1-p_{00}$
1	$1-p_{11}$	p_{11}

After n packet transmissions, the aggregated reward $v(n-1)$ represents the performance of predictive sequence transmission over a lossy channel with a channel's state transition matrix P. The reward equations (6)-(8) for BEC are still valid for Gilbert Channel model, hence the variables are defined as following: In a two state Markov channel model, if the instant rewards are $\{r_0, r_1\} = \{1, 0\}$, the reward process is awarded with *1* for a successful packet and *0* for a lost packet during the transmission. In this case, after n packet transmissions, the aggregated rewards, $v_i(n-1)$, represent the expected number of good packet transmissions with the initial packet transmission at state i. To establish a MRP for an erasure-channel model, let $\{0,1\}$ be the corresponding state space to good (0) and bad (1) packet transmissions. The instant rewards r_i are assigned for each state and they are awarded to the process whenever it reaches state i (Figure 18).

A Markov Decision Process (MDP) associates a Markov reward process with a series of actions and decision criteria to find an "optimal" interleaving for a given erasure channel model properties; packet loss rate and packet correlation value. Similar to the decision process of BEC model, MDP finds the interleaving set that provides the highest sum of MRP aggregated reward.

Different from BEC case, the reward aggregation for skipped frames cannot be ignored any more due to the memory constraint, therefore, in MDP, a set of *policies*, mappings from states to actions, are associated with a new set of instant rewards, a new set of discount factors and a modified set of state transition probabilities (Table 6).

Table 6 Properties of MDP for Multimedia Stream Interleaving

Policies $\{\text{Action,Current State}\}$	***Instant Reward*** r_a		***Discount Factor*** γ_a		***Transition Probabilities***		
						0	1
$\{C,0\}$	r_C	1	γ_C	1	P_C	p_{00}	p_{01}
$\{C,1\}$		0		0		0	1
$\{S,0\}$	r_S	0	γ_S	1	P_S	p_{00}	p_{01}
$\{S,1\}$		0		1		p_{10}	p_{11}

In the instant reward perspective, instant reward of *1* is awarded for successful transmission and decoding for the policy $\{C,0\}$. For all other policies, the instant reward is set to *0,* since no frames can be decoded, due to loss, or need to be decoded, for skipped frames.

For the discount factor, the aggregated reward from the previous stage is fully propagated to the current stage, hence the discount factor is set to *1* for all policies except policy $\{C,1\}$. For the policy $\{C,1\}$, since the decoder of predictive coding is forced to stop when a lost packet is detected, the state *1* is considered as a trapping state for action C, hence the discount factor is set to *0* for the policy. Furthermore, in ISC MDP model, once the decoder is stopped due to a lost packet, it uses the last successfully decoded picture to replace the missing and effected frames, and then it restarts when a successfully transmitted I frame of a new GOV arrives to the decoder. Therefore, to reflect the trapping state, the transition probabilities for the policy $\{C,1\}$ are set to $\{p_{10}, p_{11}\} = \{0,1\}$. For all other policies, the channel's transition probabilities are used since the frame with successfully transmitted packets or lost packets in skipped frames do not affect the decoder. Therefore, the proposed MDP model's aggregated reward equations for single-packet-per-frame are:

$$v^{(k,j)}\left(s^{*(k,j)}(0)\right) = r_C \tag{34}$$

$$v^{(k,j)}\left(s^{*(k,j)}(n-1)\right)$$

$$= r_C + diag(\gamma_C) P_C P_S^{\left(s^{*(k,j)}(n) - s^{*(k,j)}(n-1) - 1\right)} v^{(k,j)}\left(s^{*(k,j)}(n-2)\right) \quad (35)$$

$$, \forall n \in s^{*(k,j)}$$

This is valid since the aggregated reward for a skipped frame is:

$$\begin{aligned} v^{(k,j)}(l) &= r_S + diag(\gamma_S) P_S v^{(k,j)}(l-1) \\ &= \begin{bmatrix} 0 & 0 \end{bmatrix}^T + diag(\begin{bmatrix} 1 & 1 \end{bmatrix}^T) P_S v^{(k,j)}(l-1) \\ &= P_S v^{(k,j)}(l-1) \end{aligned} \quad (36)$$

For multiple packets per frame case, the reward equations are following:

$$v^{(k,j)}\left(s^{*(k,j)}(0)\right) = r_C \quad (37)$$

$$v^{(k,j)}\left(s^{*(k,j)}(n)\right)$$

$$= r_C + diag(\gamma_C) P_C{}^{\eta} P_S^{\eta \times \left(s^{*(k,j)}(n) - s^{*(k,j)}(n-1) - 1\right)} v^{(k,j)}\left(s^{*(k,j)}(n-2)\right) \quad (38)$$

$$\text{for } 1 \le n \le N-1$$

The term $P_C{}^{\eta}$ is multiplied to the aggregated reward since a frame is decoded if and only if all the packets in the coded frames are successfully transmitted. Similarly, $P_S^{\eta \times \left(s^{*(k,j)}(n) - s^{*(k,j)}(n-1) - 1\right)}$ is multiplied to represent packets associated with skipped frames. As in BEC case, for each interleaving set k, the sum of aggregated rewards gives corresponding expected number of successfully decoded frames. Hence the equations (16)-(18) are still valid in finding an optimal interleaving set k that satisfies our decision criteria, a set with the highest MRP aggregated reward.

4.1.3 ISC-MDP with Frame Correlation

As described in section 3.3, frame correlation is also taken into consideration for ISC-MDP to reflect the frame dependent nature of predictive video coding and frame replacement process for decoder failed frames. Detailed description on frame correlation calculations and associated reward computation equations are provided in section 3.3. When incorporated with the ISC-MDP equations, (31)-(38), the decision criteria and equations, (19)-(30), of ISC-BEC model are still valid for ISC-MDP.

4.2 ISC-MDP Evaluation and Analysis

4.2.1 Simulation Setup

For evaluation, CIF sequences of *Akiyo*, *Foreman*, *Coastguard*, and *Mobile* were coded into an *IPPP...* GOV structure using an MPEG-4 encoder. GOV sizes (un-interleaved size) of 10, 12, 14, 16, 18, and 20 were used to partition the evaluation sequences. Frame rate of 15 frames per second, bitrate of *250kbps* and *500kbps*, and packet size of *512 Byte* are used to represent emerging Internet-access technologies (e.g., DSL/Cable and LAN connections). Only Intra- (*I*) and

Inter- (*P*) coded frames are used to form GOV.

The evaluation scenarios are as same as the ones from ISC-BEC evaluation: (a) no correlation (ISC-NC), sequence specific correlation (ISC-C), and generic correlation (ISC-GC)

For the realistic network channel model simulation, packet-loss Markov transition probabilities from [4, 5], $p_{00} = 0.9734, p_{11} = 0.7052$, are used to capture realistic network loss patterns. In addition, 5%, 10%, and 15% packet loss rates were used and packet correlation value of 0.3, 0.6, and 0.9 were used to represent low, medium, and high correlation between the transmitted packets. For each network loss probabilities, from [4, 5] or $p_1 - \rho$ pair, ten packet loss traces were generated. Each evaluation case is fitted into these packet loss traces and the PSNR values are averaged to provide statistically satisfying results for analysis.

The simulation results are given in the following order: (a) real network model simulation, (b) variable $p_1 - \rho$ pair simulation.

4.2.2

41.5

40.5

39.5

38.5

33.5

32.5

31.5

30.5

4.2.2 Simulation Results

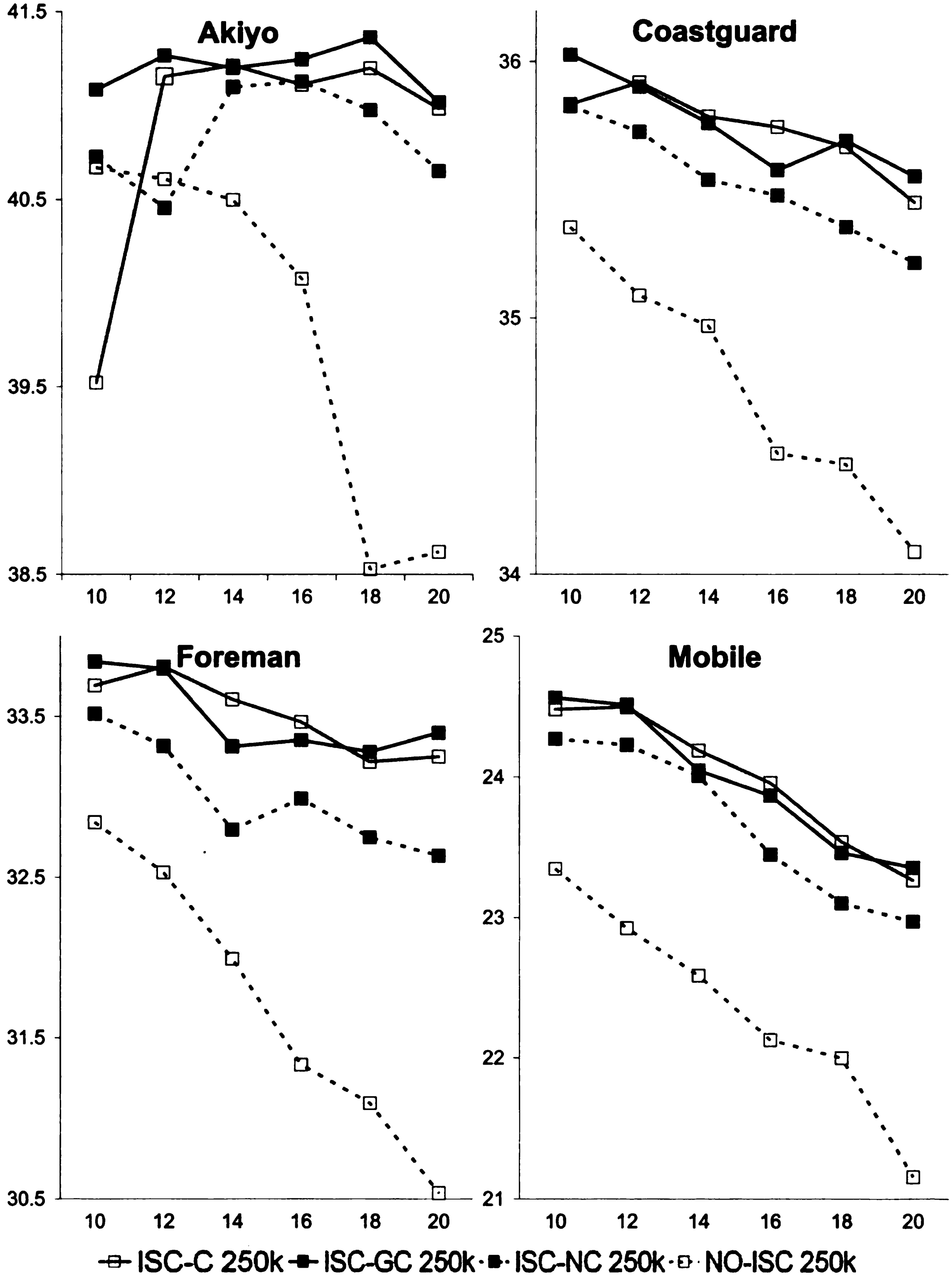


Figure 19 Average PSNR (GOV Size vs. PSNR(dB)) *@ 250kbps*

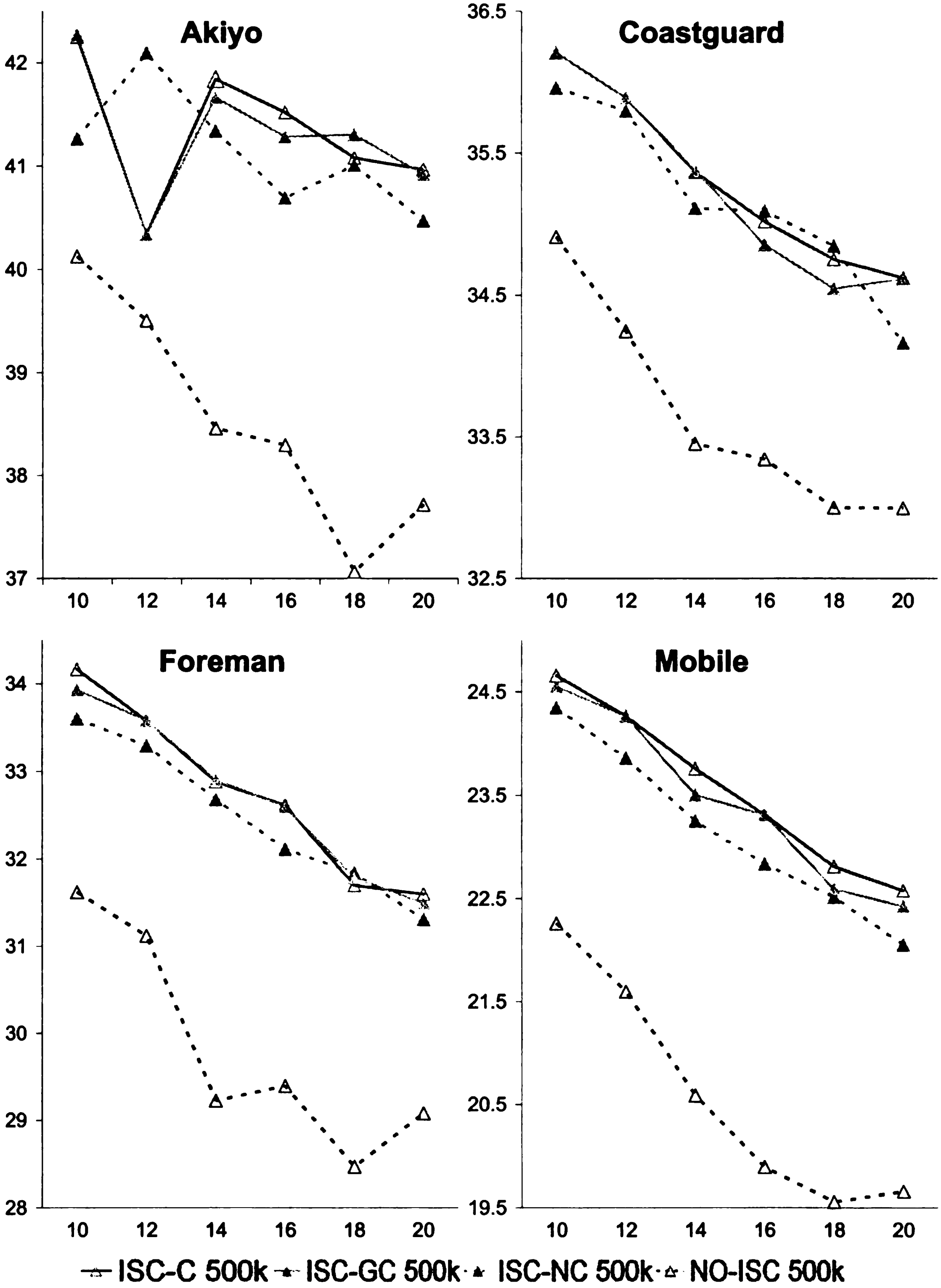


Figure 20 Average PSNR (GOV Size vs. PSNR(dB)) *@ 500kbps*

Table 7 PSNR Differences (dB): @ *500kbps* - @ *250kbps*

		10	12	14	16	18	20
ISC-C	Akiyo	**2.7269**	-0.8161	**0.6278**	**0.4076**	-0.1192	-0.0185
ISC-GC		**1.1866**	-0.9270	**0.4635**	**0.0376**	-0.0599	-0.0950
ISC-NC		**0.5363**	**1.6376**	**0.2389**	-0.4356	**0.0310**	-0.1814
NO-ISC		-0.5449	-1.1055	-2.0402	-1.7833	-1.4637	-0.9060
ISC-C	Coastguard	**0.3728**	-0.0300	-0.4178	-0.7267	-0.9147	-0.8284
ISC-GC		**0.1818**	-0.0129	-0.3814	-0.7227	-1.1446	-0.9390
ISC-NC		**0.1311**	**0.0701**	-0.4260	-0.3865	-0.5083	-1.0518
NO-ISC		-0.4437	-0.8419	-1.5190	-1.1299	-1.4293	-1.0905
ISC-C	Foreman	**0.4749**	-0.2295	-0.7225	-0.8510	-1.5208	-1.6541
ISC-GC		**0.0937**	-0.2174	-0.4077	-0.7553	-1.4945	-1.9041
ISC-NC		**0.0852**	-0.0239	-0.1164	-0.8747	-0.9077	-1.3292
NO-ISC		-1.2210	-1.4121	-2.7649	-1.9366	-2.6234	-1.4544
ISC-C	Mobile	**0.1777**	-0.2298	-0.4303	-0.6446	-0.7273	-0.6882
ISC-GC		-0.0128	-0.2458	-0.5420	-0.5544	-0.8675	-0.9304
ISC-NC		**0.0752**	-0.3674	-0.7562	-0.6121	-0.5850	-0.9178
NO-ISC		-1.0877	-1.3254	-1.9982	-2.2344	-2.4418	-1.4970

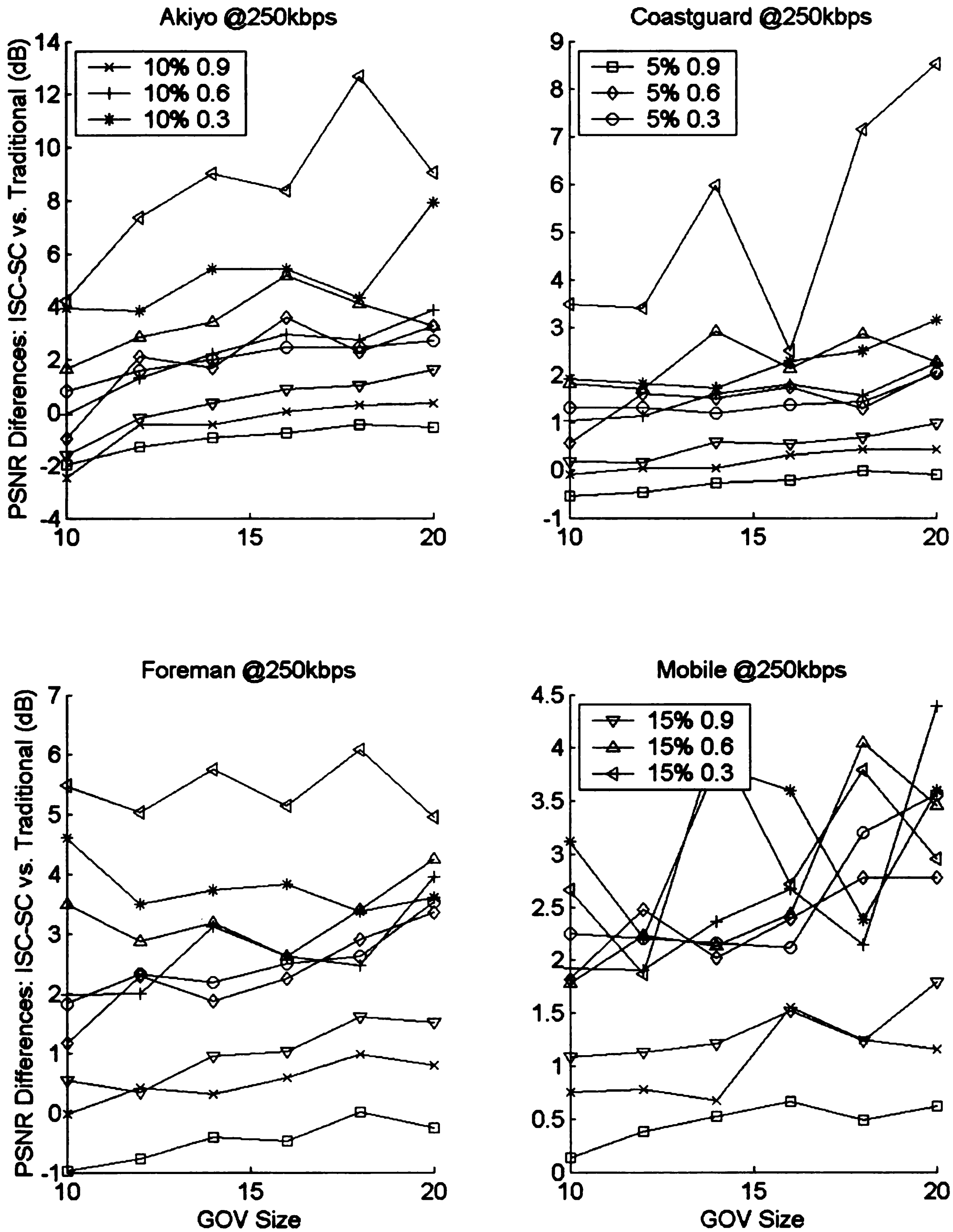


Figure 21 PSNR Differences: Sequence Specific ISC and Non-ISC *@ 250KBPS*

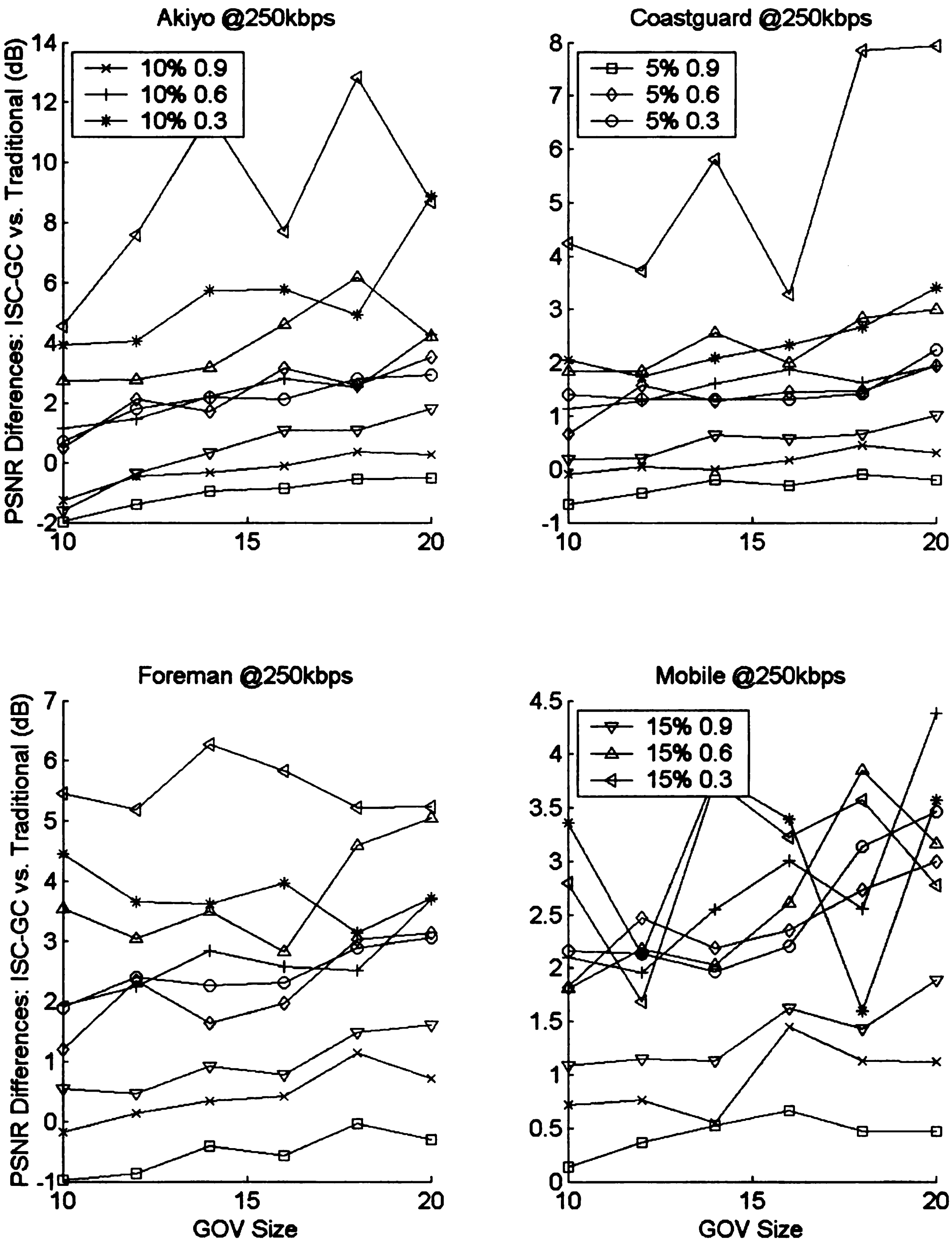


Figure 22 PSNR Differences: Generic ISC vs. Non-ISC *@ 250KBPS.*

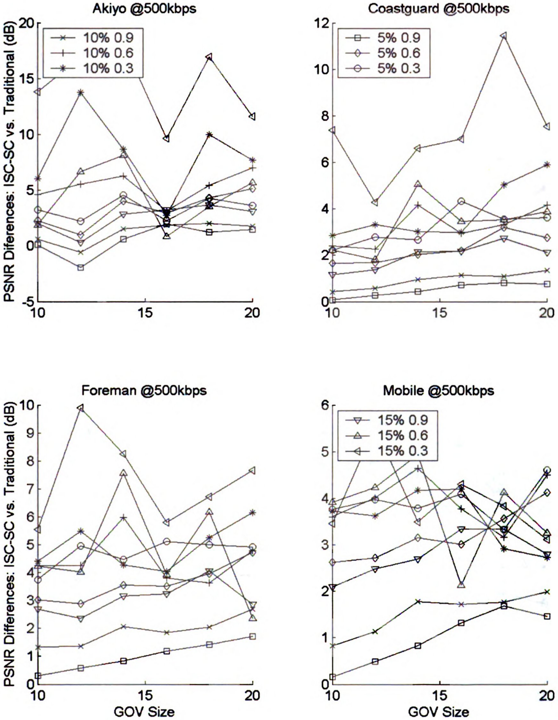


Figure 23 PSNR Differences: Sequence Specific ISC vs. Non-ISC *@ 500kbps.*

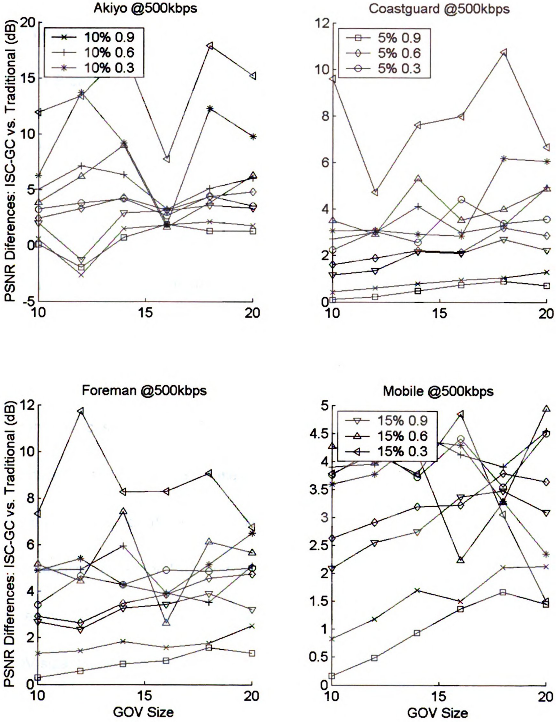


Figure 24 PSNR Differences: Generic ISC vs. Non-ISC *@ 500kbps*

4.2.3 Analysis

4.2.3.1 Bitrate and GOV size variation effects

The non-ISC cases show linear downward trend with respect to the GOV size and bitrate. This implies that such variations have negative impacts on the quality, since such changes increase the average number of packets per frame η, which in turn causes an increase in (a) the number of GOVs impacted by lost packets, (b) the average number of replaced frames, and (c) the distance between the replacement frames.

For the ISC cases, with the GOV size increment, the average PSNR shows linear trends similar to the non-ISC cases. However, the slope is rather flat when compared to the non-ISC cases. This implies that the GOV size variation has less negative impact on ISC method compared to the traditional non-ISC method. When the sequences are coded using the same coding method at the same GOV size, but with the different bitrates, e.g., *250kbps* and *500kbps*, Table 7 , shows that variation of bitrate has less impact on the PSNR values for the ISC cases than the non-ISC cases; hence this shows that ISC reduces the negative impact of increased η, the average number of packets per frame, as stated previously.

In addition, as shown in, since the average PSNR gain of ISC cases over non-ISC cases are higher, this implies that the ISC method performs better when coded at higher bitrate.

4.2.3.2 Correlation Gain Improvements

The correlation-based models, both ISC-C and ISC-GC, provide improvements over the non-correlation (ISC-NC) based scenario. In, the latter sets show improvements in PSNR gain for most of the evaluation cases, and hence demonstrate the advantages of the correlation gain computation. When comparing the two different correlation model sets, the generic correlation model shows competitive results, and it is plausible to use the generic model in cases when the actual temporal correlation for a given sequence is not feasible to compute.

4.2.3.3 Variation of Gilbert Model Parameter Pairs

ISC shows improvements on most of the evaluation cases. It is clearly seen that ISC advances the traditional method as the channel loss rate increases or the packet correlation rate decreases. This is due to the fact that ISC reduces impact of packet losses to the GOV by isolating losses to one of the two sub-sequences and decreases frame replacement distances for decoder failed frames. However, with the increment of the packet correlation constants, the frequency of the long packet loss bursts increases, hence increases the chance that both sub-sequences are impacted by the long packet loss bursts.

4.2.3.4 Evaluation Summary

Overall observation shows that the proposed ISC method improves over the traditional approach on most of the cases, especially for the sequences with high motion or low temporal correlation (Figure 8). Up to 4 dB in average PSNR improvements is observed.

This represents a very significant improvement in quality for compressed video applications. In particular, this demonstrates that ISC improves the quality of predictive coded sequences over an erasure channel by limiting losses to one of the two sub-sequences, hence minimizing the cascaded effects of lost packets, and/or decreasing the average frame replacement distance. In addition, changes in bitrate or GOV size have less impact on ISC coded sequences. Furthermore, when the non-correlation gain computed ISC (ISC-NC) sets are compared to the correlation computed sets (ISC-C and ISC-GC), the latter sets show some modest improvement in PSNR for most of the evaluation cases. Consequently, it is feasible that significant improvements can be gained by taking into consideration the channel model *only*, and hence, reducing the complexity for identifying the optimum interleaving set. Once the optimum interleaving is identified for a given channel model, this interleaving can be applied to any video sequence (i.e., without taking into consideration the particular statistical properties of the video sequence).

Chapter 5

Multi-Stream Interleaved Source Coding

From the ISC-BEC and ISC-MDP evaluations, under the same channel condition, regardless of how large the bitrate consumption is (primarily by the intra-coded frames), ISC-based coding have shown improvement in quality over traditional non-interleaving coding. This part of the dissertation extends the canonical ISC coding approach from two sub-streams to multi-stream coding. The main purpose of the proposed extension is to investigate ISC flexibility under various channel conditions with the variations of ISC-GOV refreshment rate and the number of interleaving sub-streams.

5.1 Multi-Stream Interleaved Source Coding

For multi-stream ISC, two sub-stream set search approaches, extensive search and greedy search approach, are taken. In sub-stream selection, the constraint $\sum\left\{s^{(j)}(2)-s^{(j)}(1),\cdots,s^{(j)}(N)-s^{(j)}(N-1)\right\}>N-1$ from (2) is removed to give more selection flexibility, and the new sub-stream selection algorithm is;

$$
\begin{aligned}
&s = \left\{0 \quad 1 \quad \cdots \quad M \times N - 1\right\} = \bigcup_{j=1}^{M} s^{(j)} \\
&\bigcap_{j=1}^{M} s^{(j)} = \varnothing, \qquad \forall j, size\left(s^{(j)}\right) = N
\end{aligned}
\tag{39}
$$

For extensive search algorithm, the size of the set K of all possible interleaving sets for a given GOV size is:

$$
size(K) = \prod_{i=0}^{M-2} \frac{(M \times N - i \times N)!}{(M-i)(M \times N - (i+1) \times N)!N!} \tag{40}
$$

Table 8 Number of Possible Interleaving Set, K

NON-ISC GOV SIZE	9	12	15	18	21
SIZE K, $M=3$	280	5775	126126	2858856	66512160
NON-ISC GOV SIZE	8	12	16	20	24
SIZE K, $M=4$	105	15400	2627625	48864376	96197645544

However, searching for an optimal interleaving set using extensive search approach is not always feasible due to the vast size of K.

The greedy search algorithm searches for an optimal multi-stream interleaving set based on two sub-stream interleaving. First, the algorithm searches an optimal two sub-stream interleaving set $\mathbb{S}^{(k)} = \left\{s^{(k,1)}, s^{(k,2)}\right\}$ using (19) or (30). Once the initial optimal interleaving set is found, for each sub-stream in $\mathbb{S}^{(k)}$, the following pseudo code is used to find the final optimal interleaving set $\mathbb{S}^{(k_h)}$ with predefined hierarchical level index $h_{\max}$.

```
for $h=1$ to $h_{\max}$
    if $N_{h-1}$ is even
        $N_h = N_{h-1}/2$
        using (19) or (30), find
            $\mathbb{S}^{(k_h,j)} = \left\{ s^{(k_h,1)}, s^{(k_h,2)} \right\},\ j=1,2\ s.t.\ \mathbb{S}^{(k_h,j)} \in s^{(k_{h-1},j)}$
    else
        stop
end
```

Note that the algorithm can terminate before reaching $h_{\max}$ since ISC constraint (39) requires same number of frames in each sub-stream. Hence, the size of the possible interleaving set K_h can be calculated with the following equation

$$size(K_h) = \frac{(2N)!}{2(N!)^2} + 2\sum_{h=1}^{h_{\max}} h \times \frac{(N_{h-1}/h)!}{2\left((N_{h-1}/2h)!\right)^2} \quad \text{where } N_0 = N, \quad N_{h-1}/2h = \text{int} \tag{41}$$

Table 9 Number of Possible Interleaving Set, K_h

NON-ISC GOV SIZE	12	14	16	18	20	24
SIZE $K_h, h=0$	462	6435	24310	92378	352716	1352078
SIZE $K_h, h=1$	482		24380		352968	1353002
SIZE $K_h, h=2$			24392			1353040

To find optimal interleaving solution over channel with memory, equations (31)-(38) and (16)-(18) with the MDP variables from Table 6 are used for both extensive and greedy search algorithms.

5.2 Multi-Stream ISC Evaluation and Analysis

5.2.1 Simulation Setup

For evaluation, CIF sequences of *Akiyo*, *Foreman*, *Coastguard*, and *Mobile* were coded into an *IPPP…* GOV structure using an MPEG-4 encoder. GOV sizes (un-interleaved size) of 9, 12, 15, and 18 were used to partition the evaluation sequences into three sub-streams and un-interleaved GOV sizes of 12, 16, and 20 were used for four sub-stream interleaving. The extensive search algorithm was used for three sub-stream cases and greedy algorithm was used for four sub-stream cases. Frame rate of 15 frames per second, bitrate of *250kbps* and *500kbps*, and packet size of *512 Byte* were used for encoding the sequences. Only the generic correlation (ISC-GC) algorithm was used and the network conditions were set to 5%, 10%, and 15% packet loss rates, p_1, with varying packet correlation value, ρ, of 0.3, 0.6, and 0.9. For each network condition $p_1 - \rho$ pair, ten packet loss traces were generated. Each evaluation case is fitted into these packet loss traces and the PSNR values are averaged to provide statistically satisfying results for analysis.

The simulation results are given in the following order: (a) multi-stream interleaving vs. non-interleaving performance evaluations with respect to number of frames in a GOV, (b) performance evaluation between multi-stream interleaving cases with respect to number of frames in a GOV.

5.2.2 Simulation Results

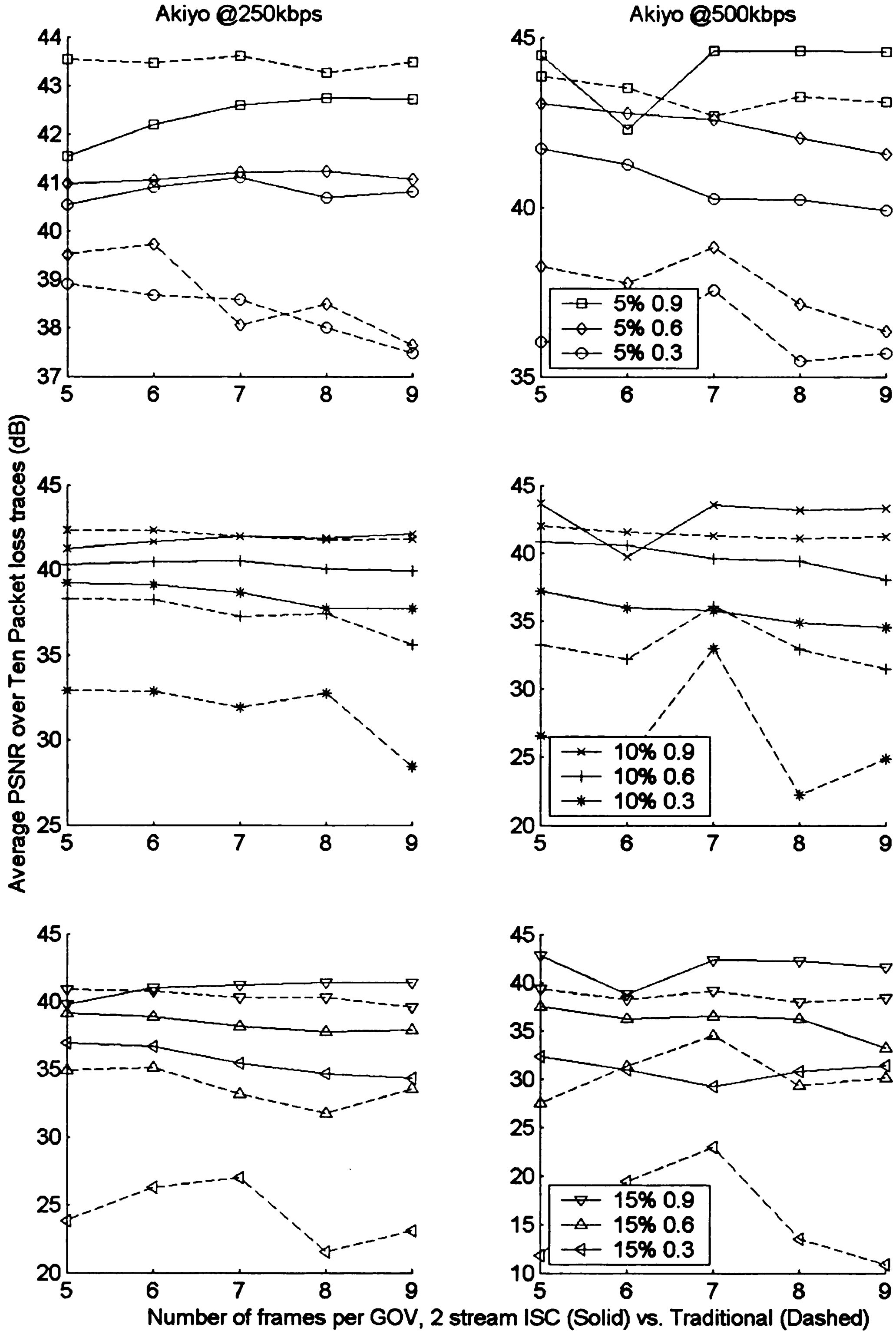


Figure 25 Average PSNR over ten packet loss traces: two sub-stream ISC vs. Non-ISC for *Akiyo*.
Number of frames per GOV based performance evaluation

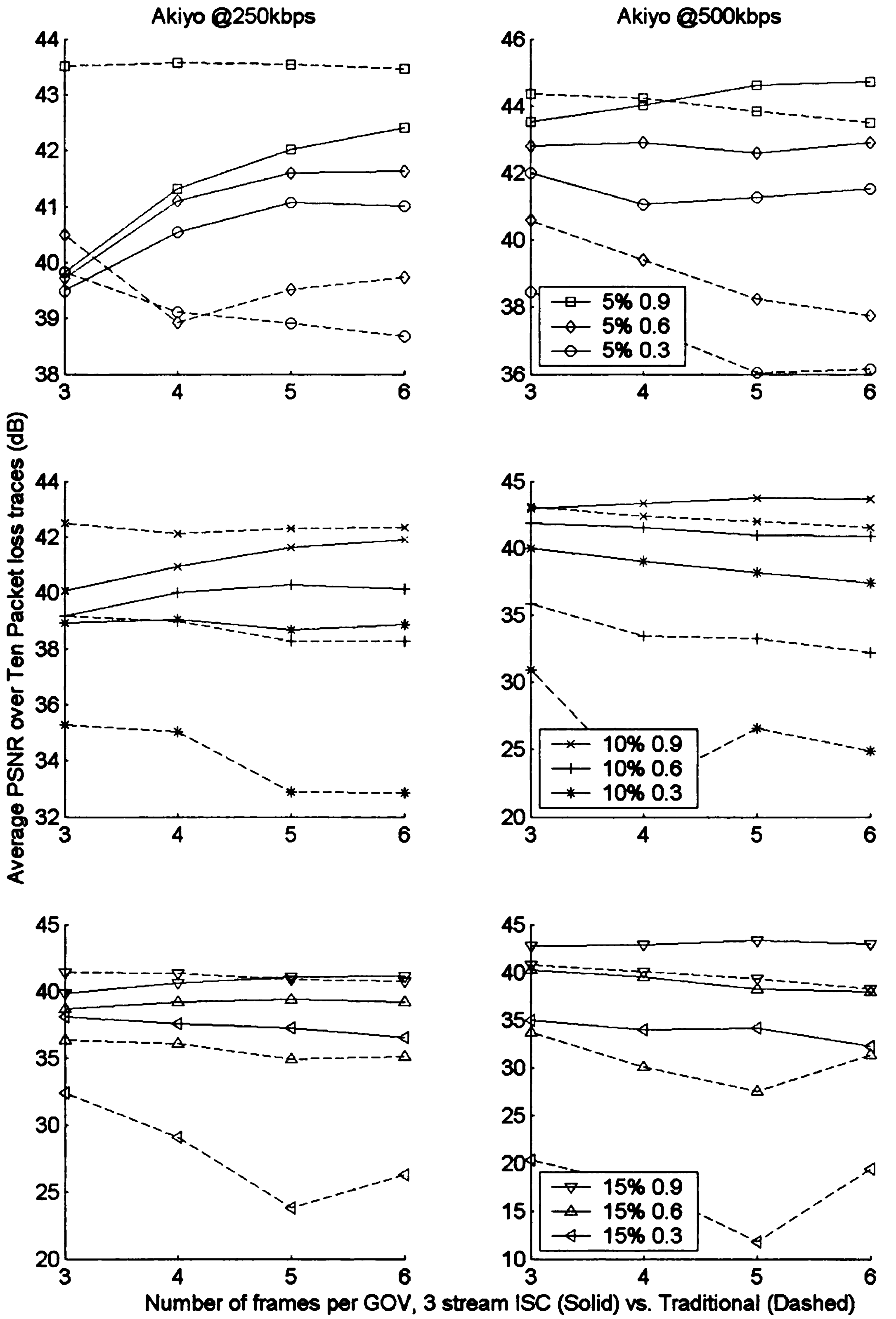


Figure 26 Average PSNR over ten packet loss traces: three sub-stream ISC vs. Non-ISC for *Akiyo*: Number of frames per GOV based performance evaluation

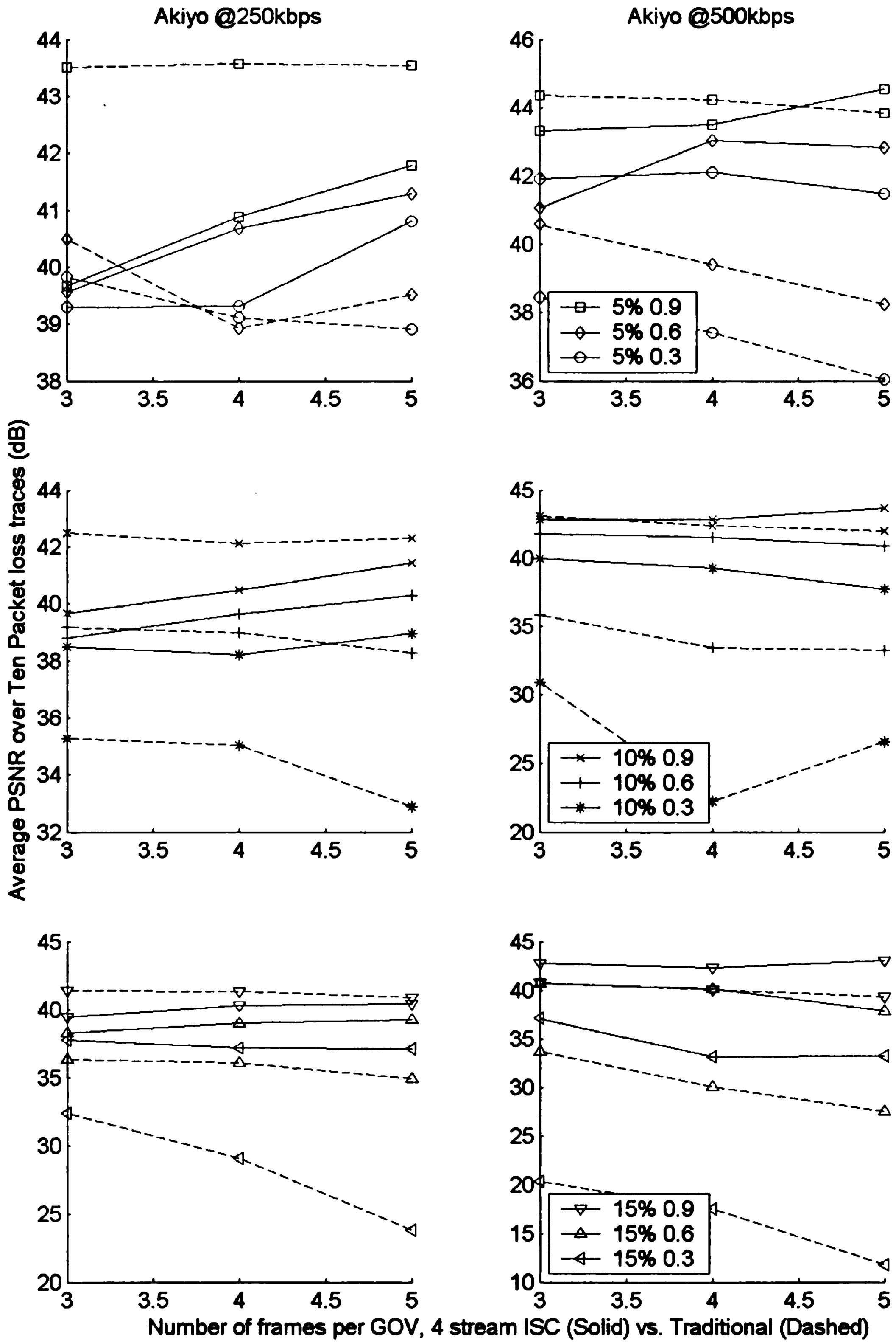


Figure 27 Average PSNR over ten packet loss traces: four sub-stream ISC vs. Non-ISC for *Akiyo*: Number of frames per GOV based performance evaluation

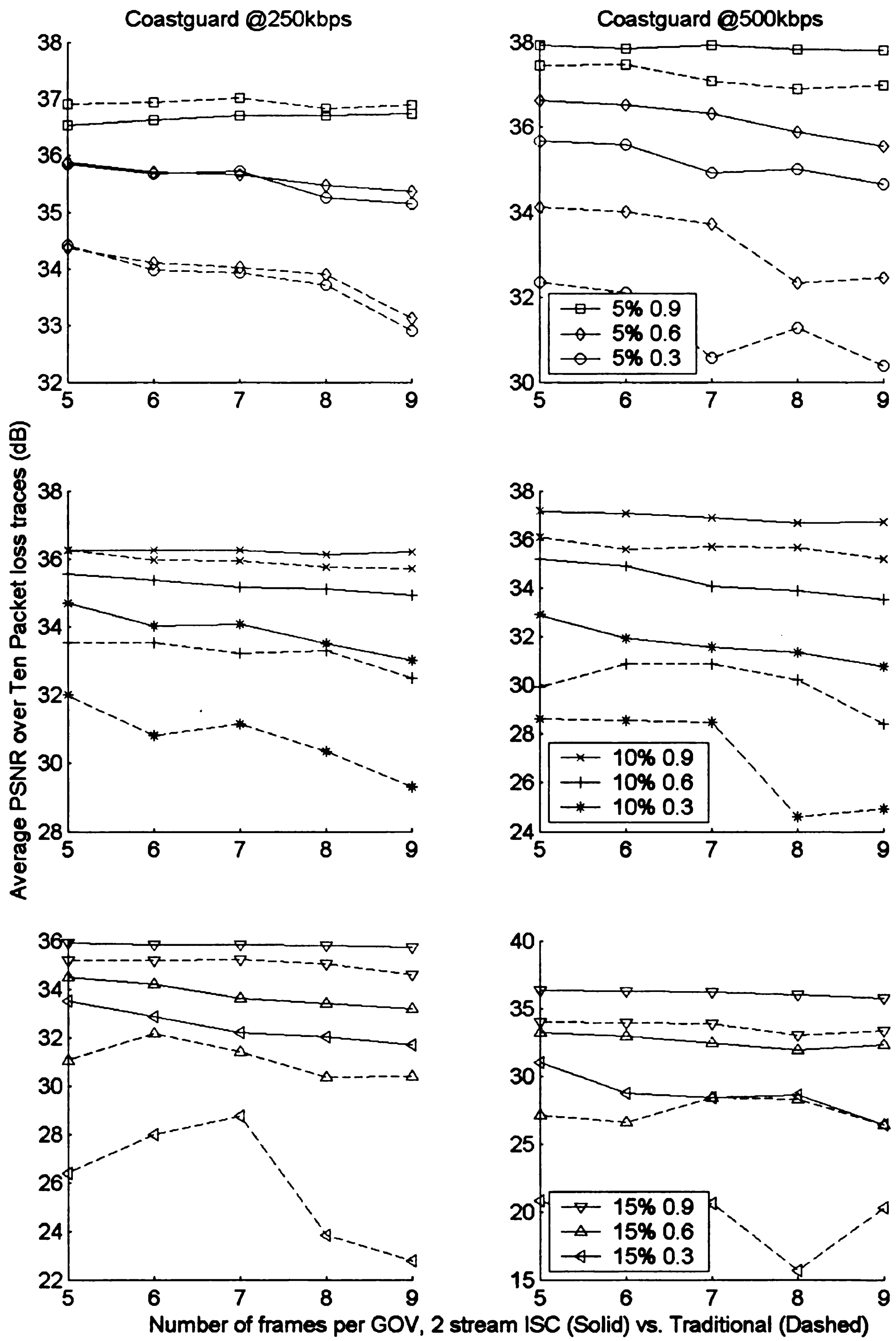


Figure 28 Average PSNR over ten packet loss traces: two sub-stream ISC vs. Non-ISC for *Coastguard*. Number of frames per GOV based performance evaluation

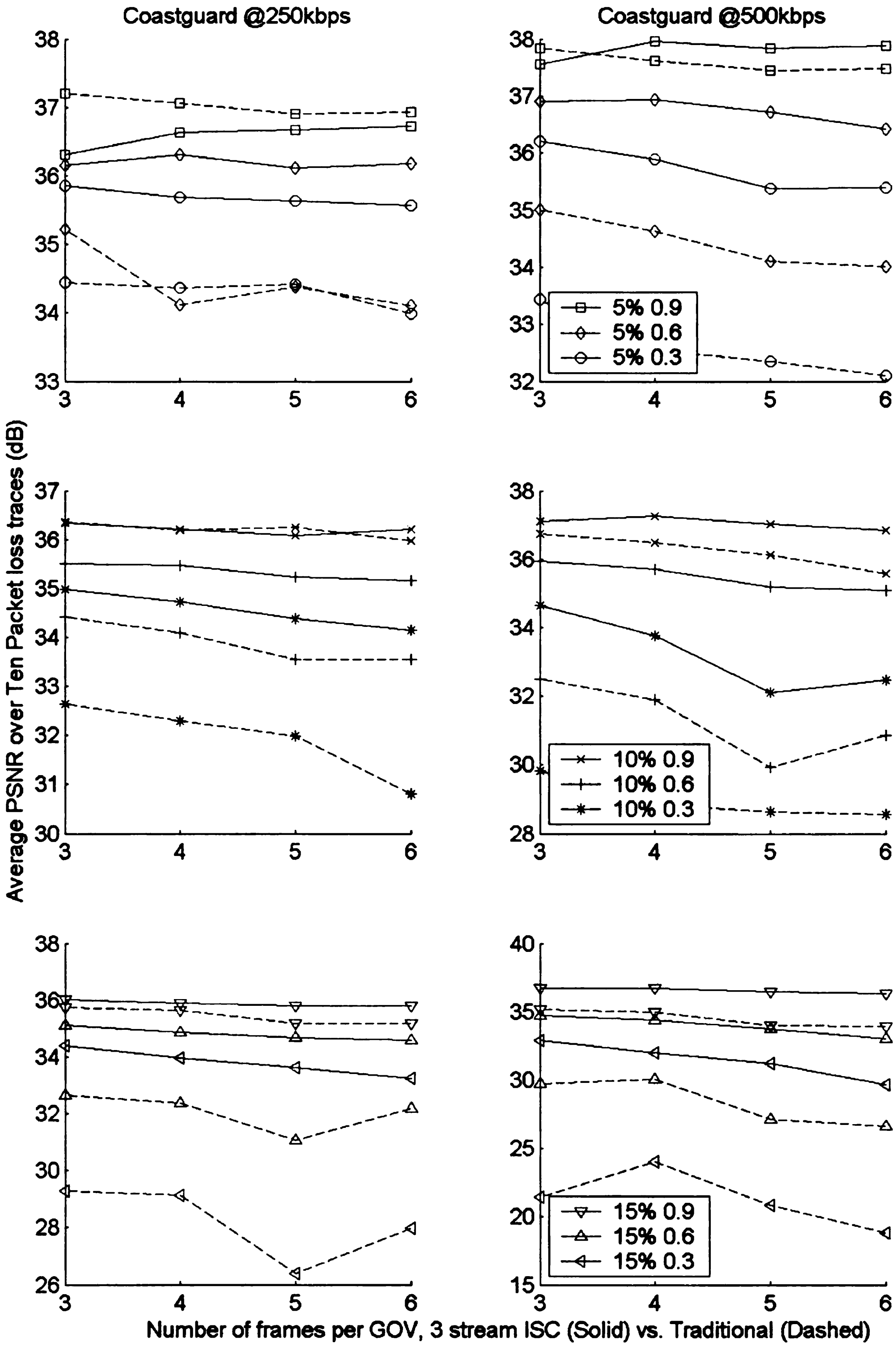


Figure 29 Average PSNR over ten packet loss traces: three sub-stream ISC vs. Non-ISC for *Coastguard*. Number of frames per GOV based performance evaluation

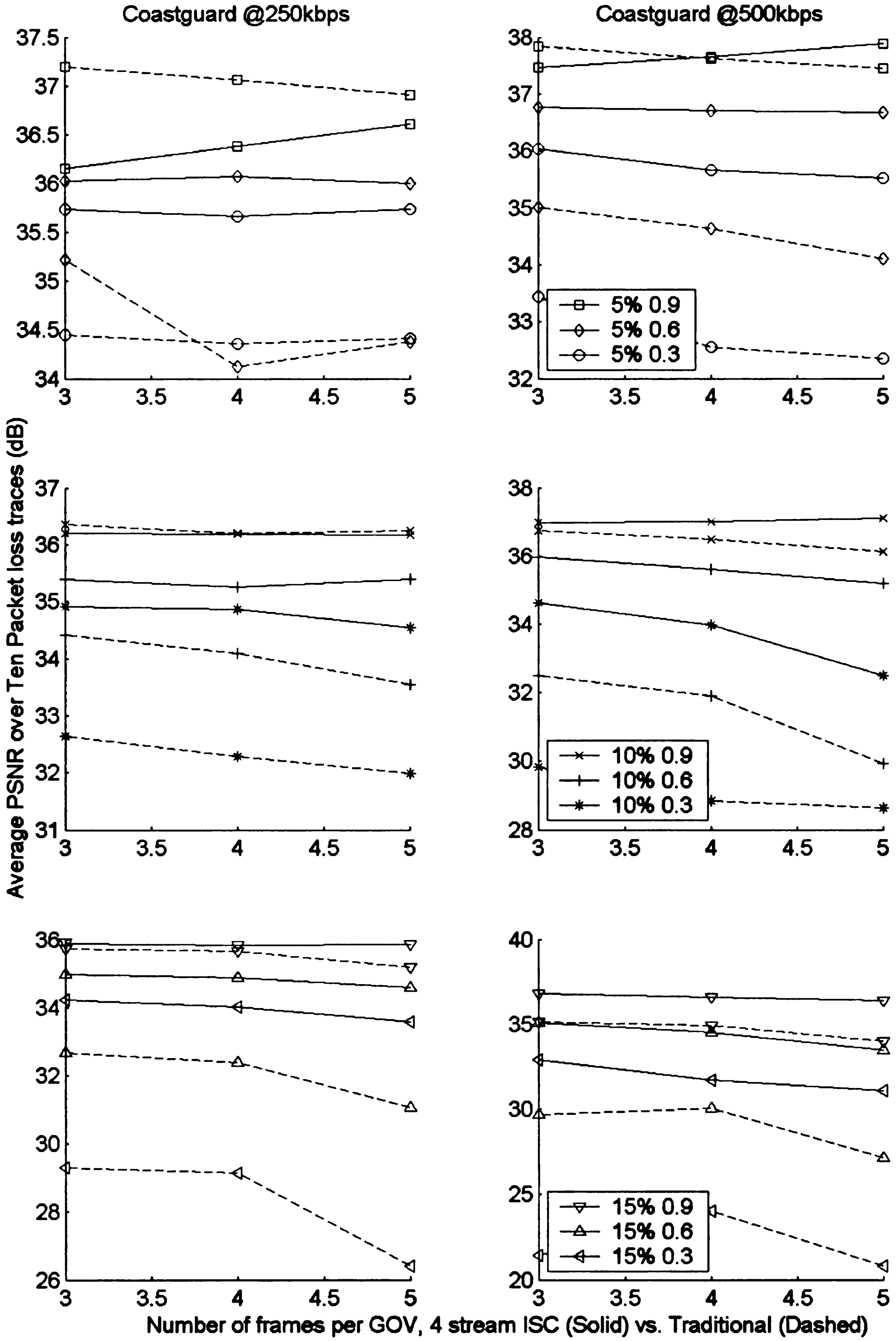


Figure 30 Average PSNR over ten packet loss traces: four sub-stream ISC vs. Non-ISC for *Coastguard*. Number of frames per GOV based performance evaluation

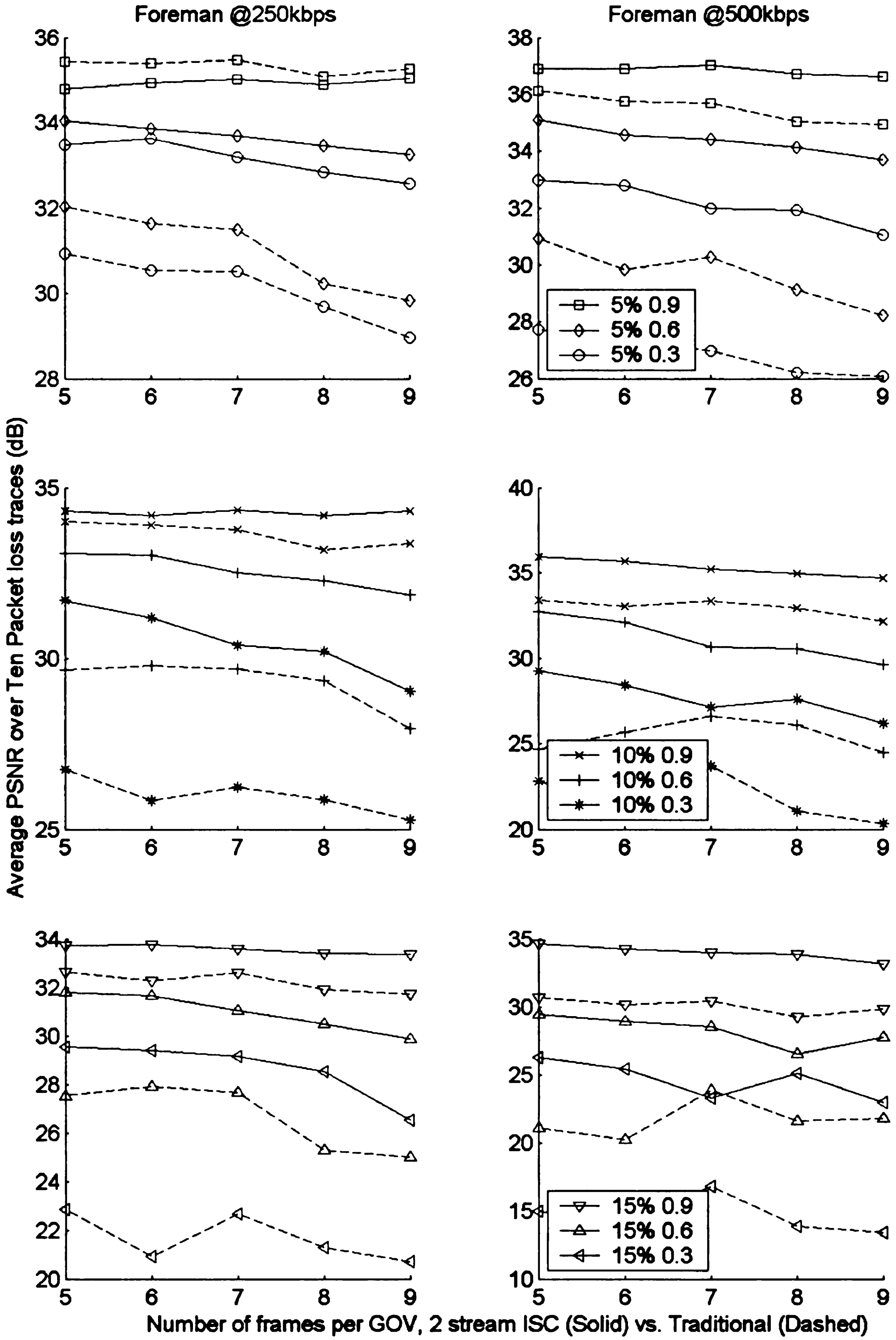


Figure 31 Average PSNR over ten packet loss traces: two sub-stream ISC vs. Non-ISC for *Foreman*. Number of frames per GOV based performance evaluation

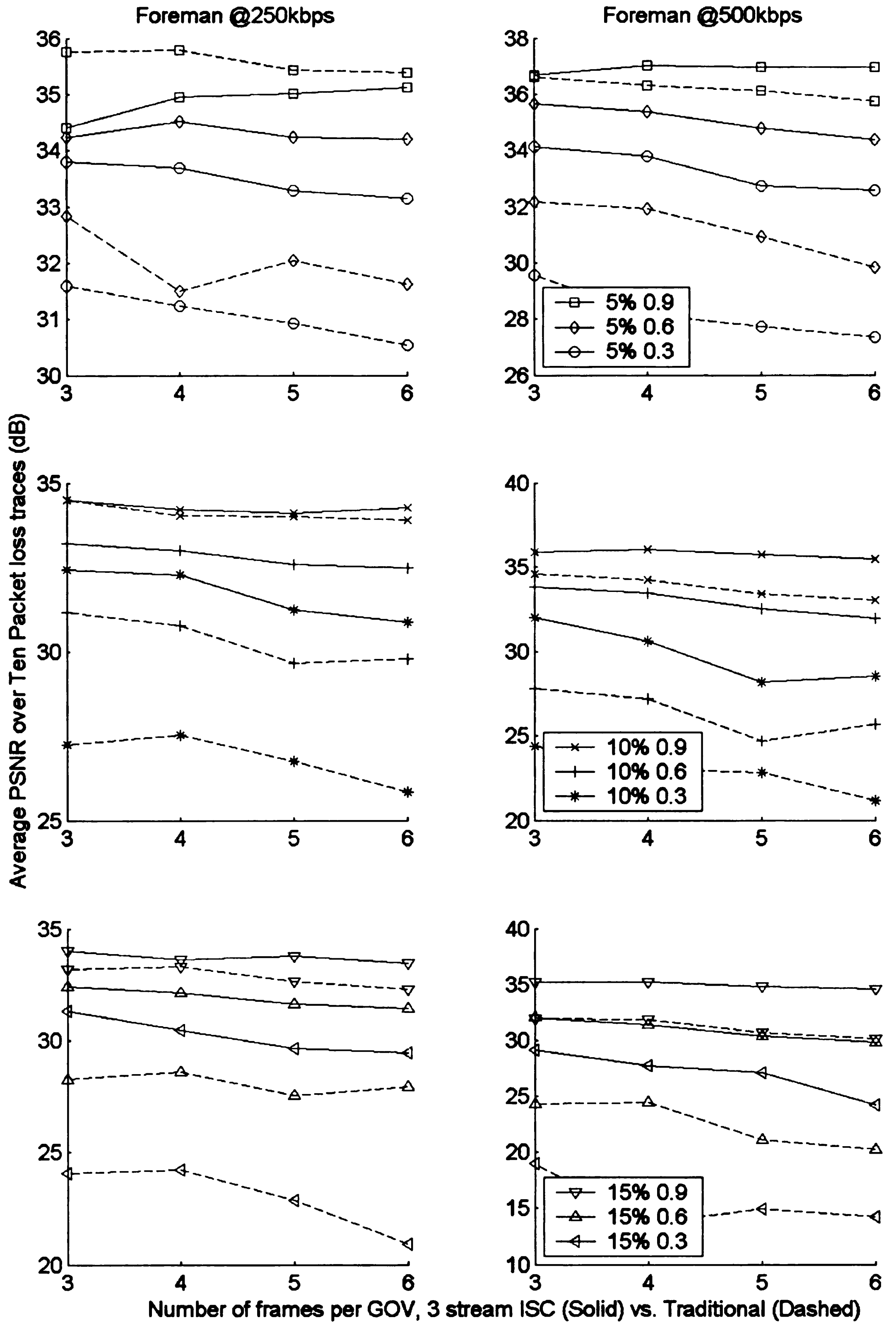


Figure 32 Average PSNR over ten packet loss traces: three sub-stream ISC vs. Non-ISC for *Foreman*. Number of frames per GOV based performance evaluation

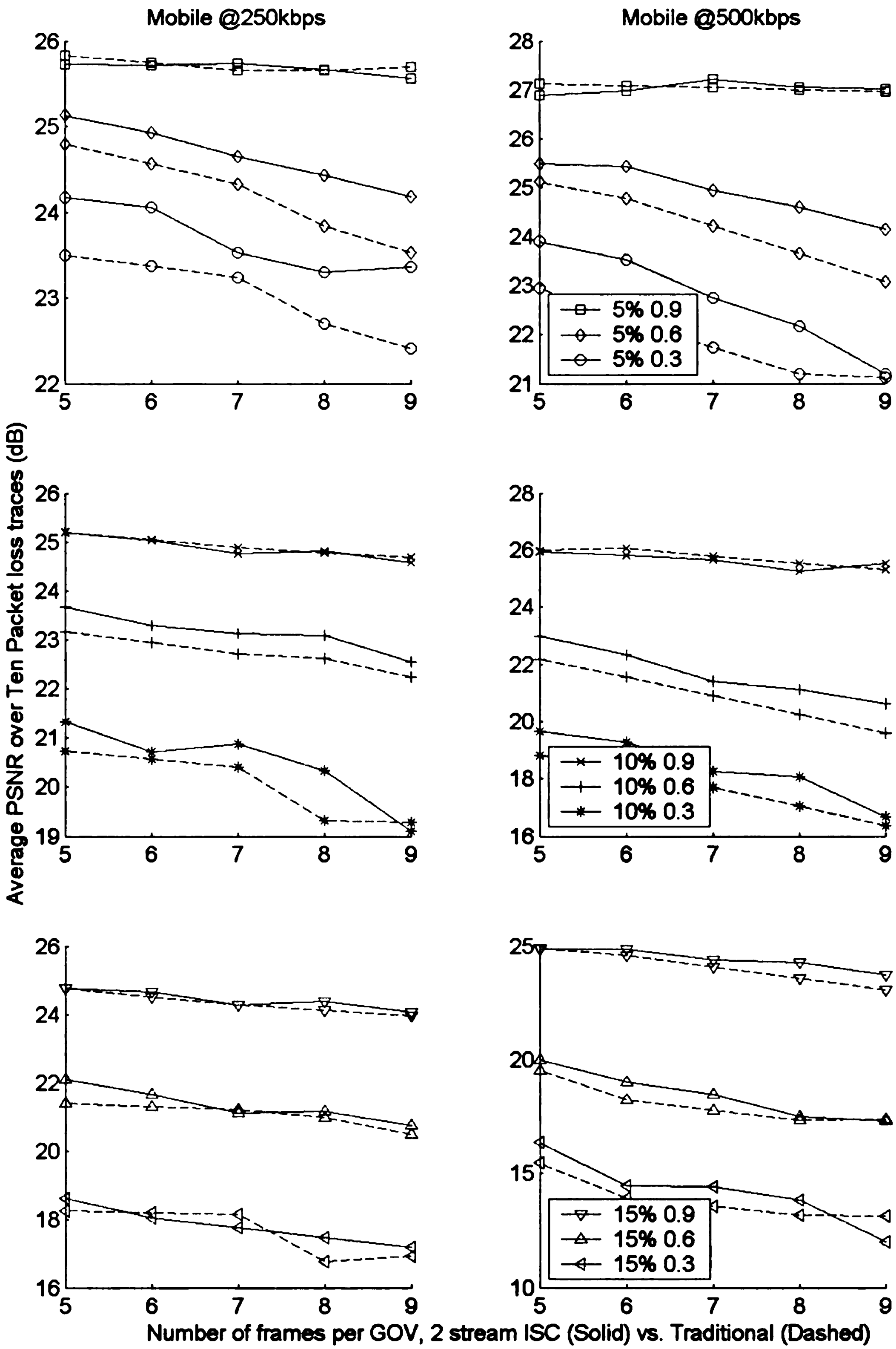


Figure 33 Average PSNR over ten packet loss traces: four sub-stream ISC vs. Non-ISC for *Foreman*. Number of frames per GOV based performance evaluation

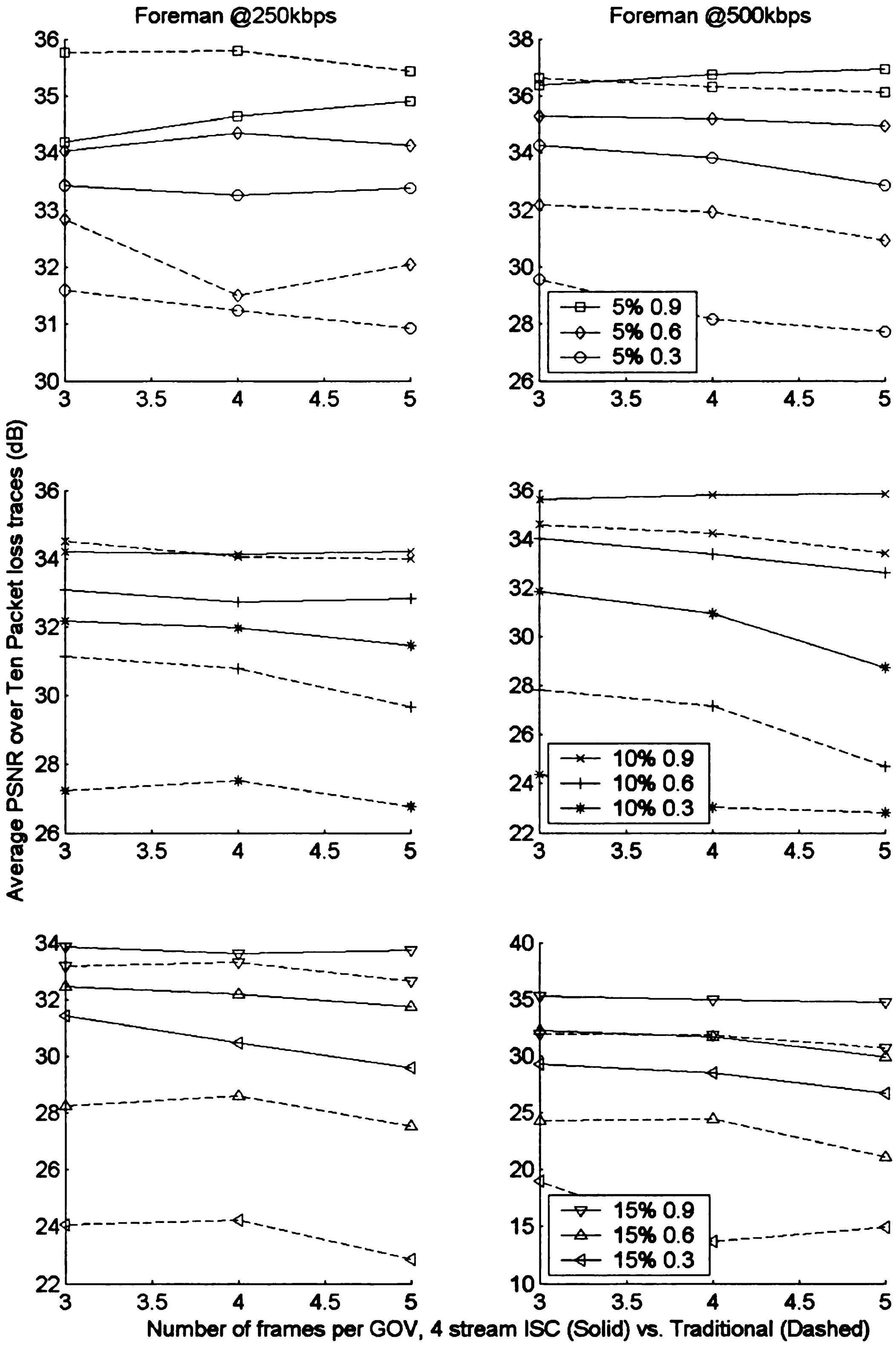


Figure 34 Average PSNR over ten packet loss traces: two sub-stream ISC vs. Non-ISC for *Mobile*. Number of frames per GOV based performance evaluation

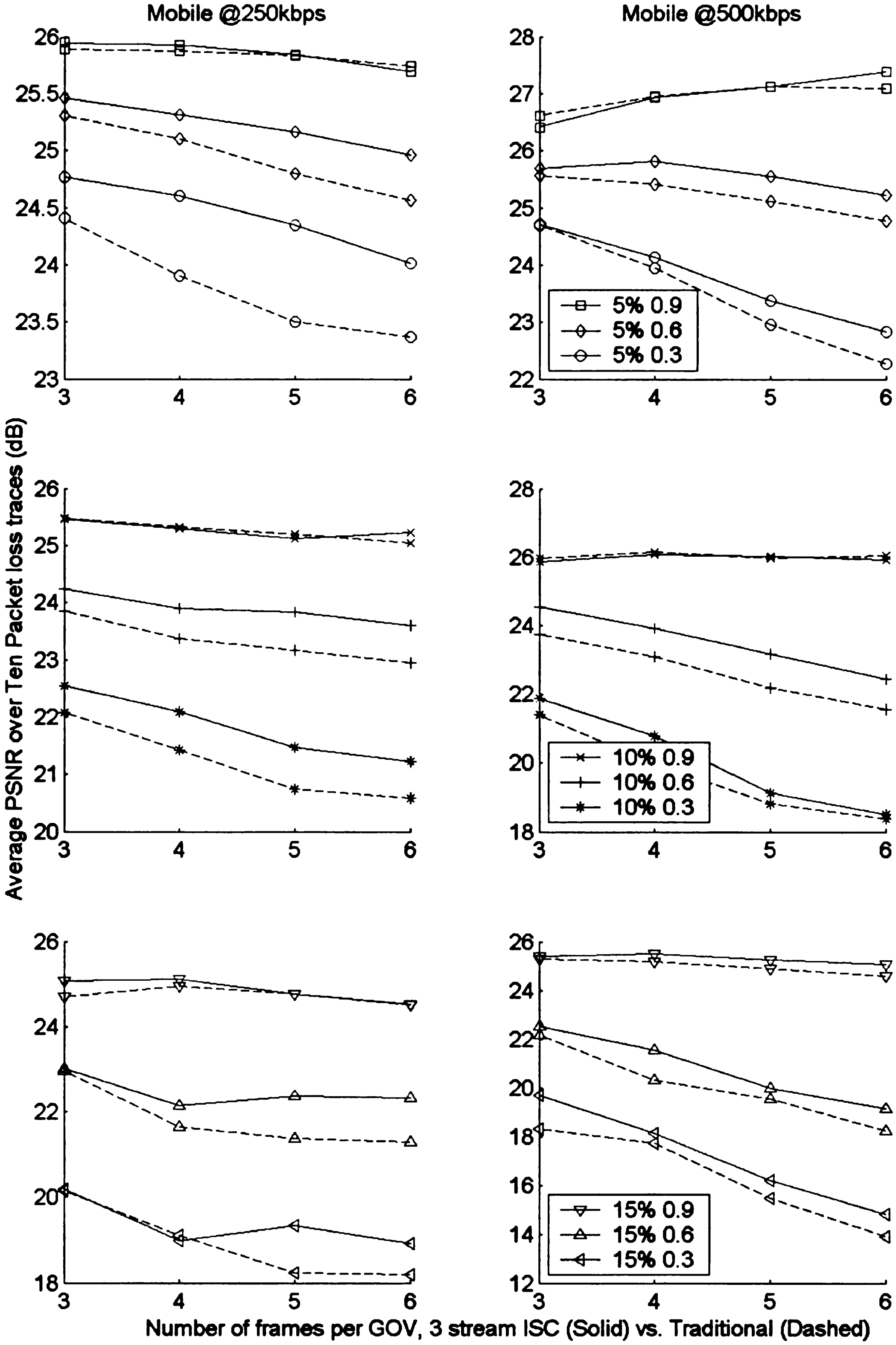


Figure 35 Average PSNR over ten packet loss traces: three sub-stream ISC vs. Non-ISC for *Mobile*. Number of frames per GOV based performance evaluation

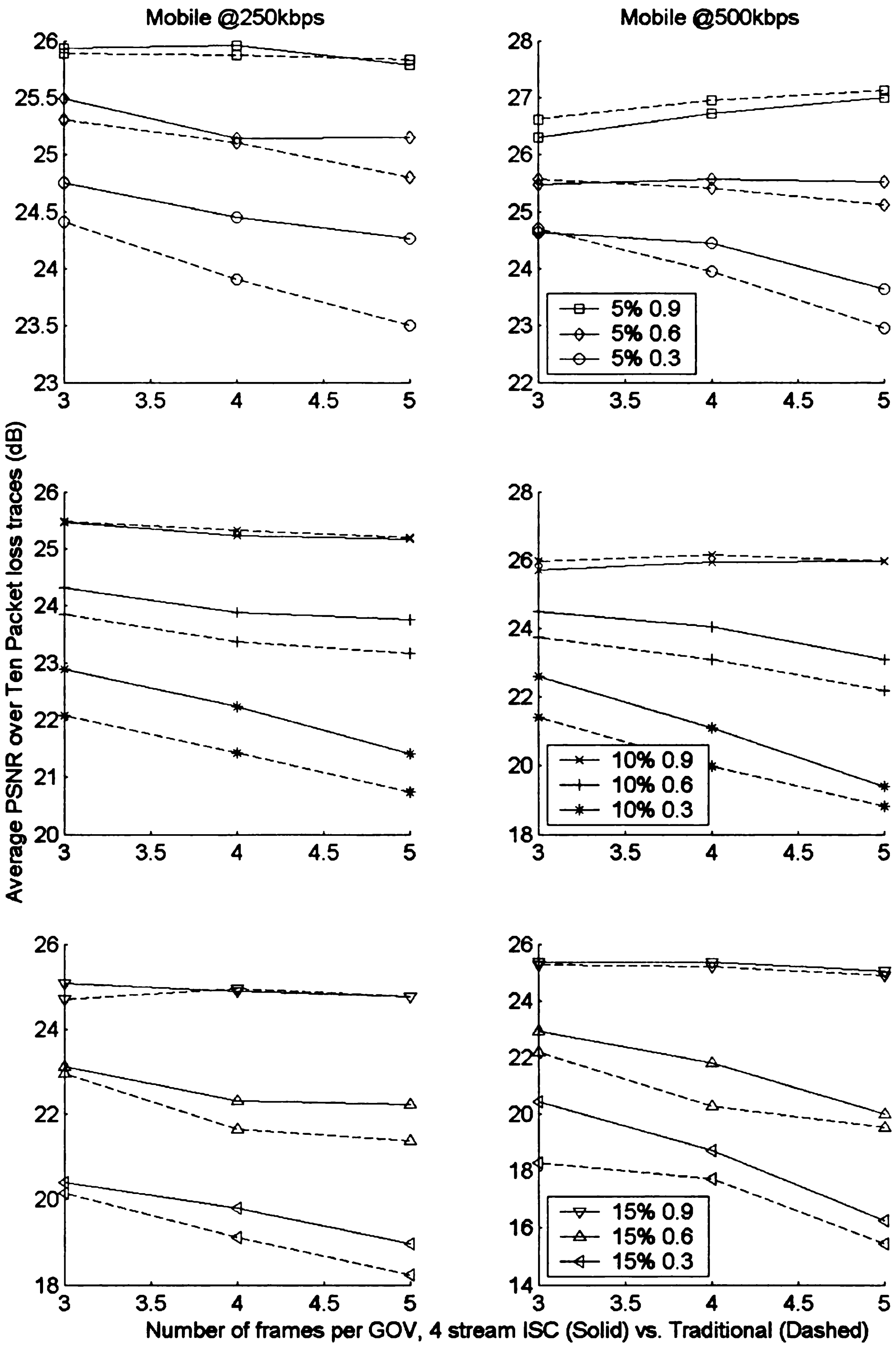


Figure 36 Average PSNR over ten packet loss traces: four sub-stream ISC vs. Non-ISC for *Mobile*. Number of frames per GOV based performance evaluation

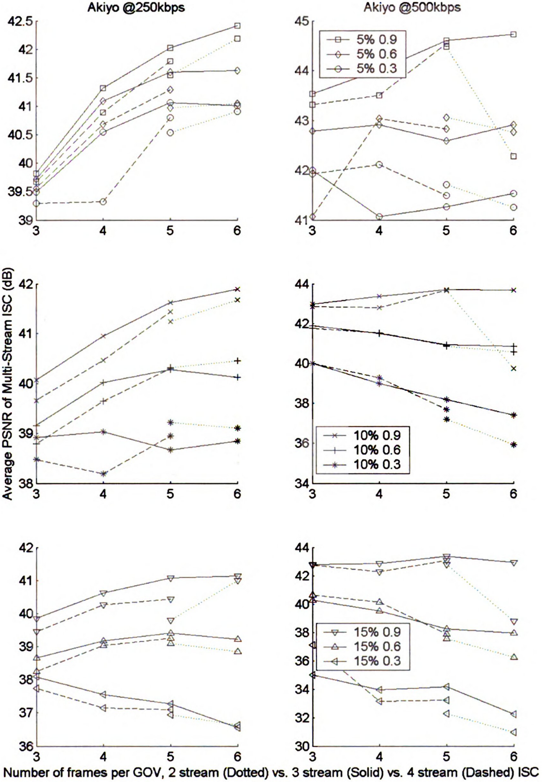


Figure 37 Average PSNR over ten packet loss traces: two, three and four sub-stream for *Akiyo*: Number of frames per GOV based performance evaluation

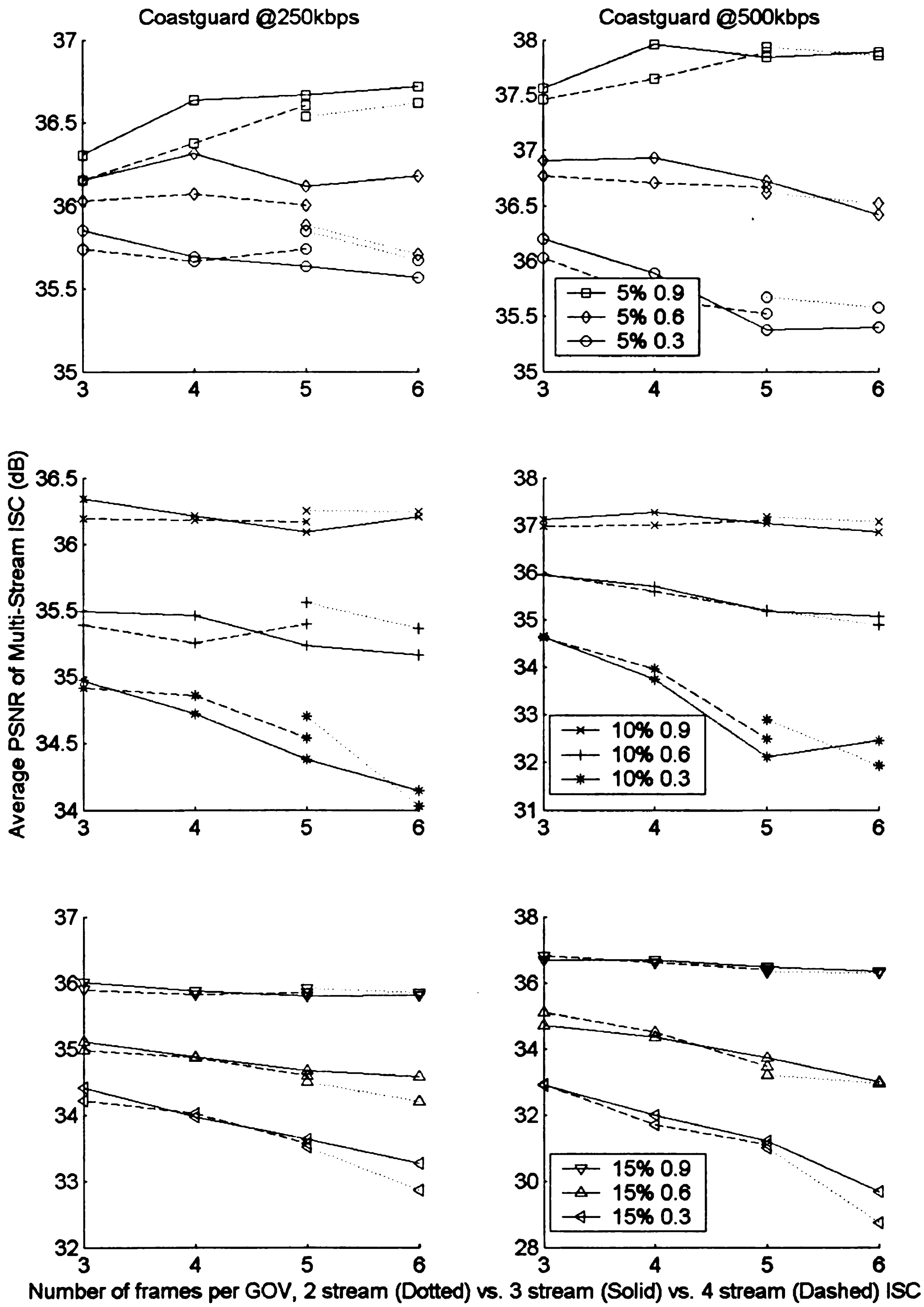


Figure 38 Average PSNR over ten packet loss traces: two, three and four sub-stream for *Coastguard*. Number of frames per GOV based performance evaluation

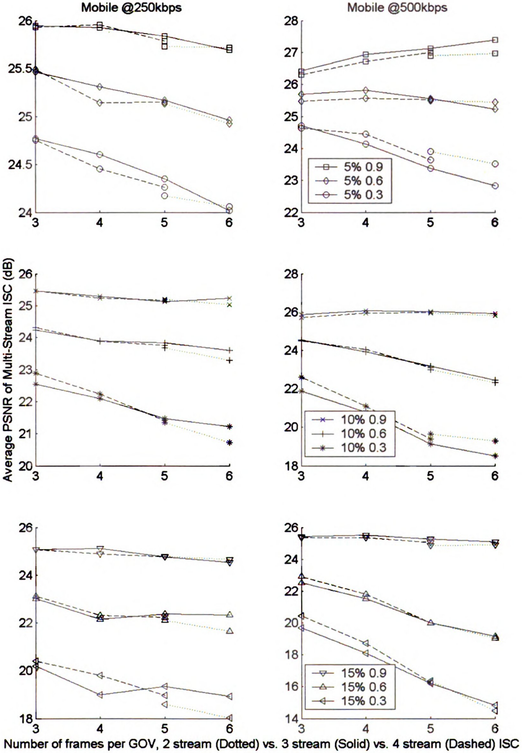


Figure 39 Average PSNR over ten packet loss traces: two, three and four sub-stream for *Foreman*. Number of frames per GOV based performance evaluation

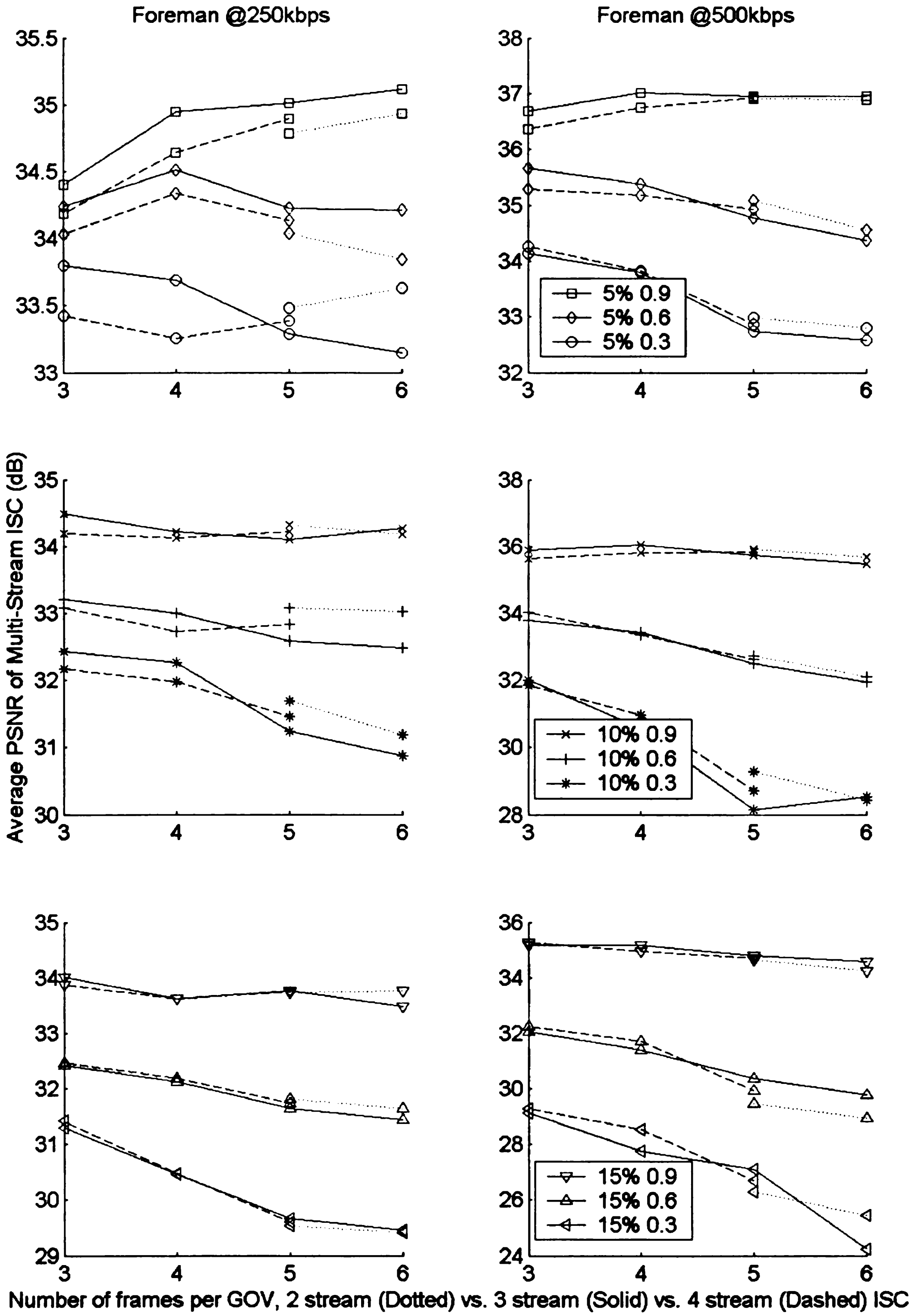


Figure 40 Average PSNR over ten packet loss traces: two, three and four sub-stream for *Mobile*: Number of frames per GOV based performance evaluation

5.2.3 Analysis

5.2.3.1 Channel condition and GOV size variation effects

When the sequences are coded into the same number of frames per GOV, ISC cases have shown performance enhancements when compared to that of non-ISC cases. Only when the sequences are coded with low bitrates and transmitted over low-loss-rate channels with very high memory, $p_1 - \rho$ pair of 5% and 0.9, non-ISC has performed better than ISC. This is due to the fact that seldom, but long burst losses can be confined into one GOV of non-ISC and the impact from losses is quickly recovered with prediction refreshment. However, even though ISC algorithm tries to confine the bursts into one of the sub-streams, due to the higher temporal prediction distortion from interleaving, the prediction refreshment without interleaving can easily outperform in such special cases. However, under the same channel condition, $p_1 - \rho$ pair , with the increment of GOV size, the performance of ISC gets closer to that of non-ISC and in some cases ISC exceeded in its performance. This is due to the fact that increasing the GOV size increases the number of packet-loss impacted frames of non-ISC cases until prediction refreshment; and the quality degradation from such impact also increases enough to exceed temporal prediction distortion from interleaving. For similar reason, under the same channel condition, when the sequence is coded into same GOV size but with higher bitrate, ISC is expected to outperform non-ISC, and yet, the simulation also show performance improvement as expected. ISC reduces packet loss impact and minimizes quality degradation, hence outperforms non-ISC in most cases. The most noticeable improvement of ISC is that the variation in quality from channel condition and bitrate variations is far less than of non-ISC. In fact, the most noticeable quality variation factor in ISC is the strength of memory. As the memory gets weaker, in other words, as the randomness of packet losses increases, it becomes harder for ISC to confine losses into minimum number of sub-streams, hence number of loss impacted frames increases which in tern degrades overall quality of the transmitted sequence. However, for non-ISC cases, since it does not have any packet loss resilient mechanism other than prediction refreshment, chances of packet loss impact can multiply with the decrement of memory strength, increment of packet loss rate, and/or increment of coding bitrate.

5.2.3.2 Multi-stream interleaving effects

When the sequences were coded into more than two sub-streams, regardless of the search method or channel condition, it was difficult to distinguish the flexibility or benefit of having multiple sub-streams when the output qualities were compared based on the number of frames per sub-stream or ISC GOV size. Therefore, considering the size of K, (Table 8 and Table 9) two stream ISC would be the most sufficient choice.

5.2.3.3 Evaluation Summary

Overall results have shown that the proposed ISC method improves over the traditional approach even with the same prediction refresh rate on most of the cases. Except for some extreme cases where the channel has low loss rate with strong memory, ISC's performance improvement was very noticeable, in some cases, up to 10dB. The most important observation made from this part of dissertation is that the most significant quality degrading factor is randomness of packet losses, or strength of memory; however, for ISC, the quality variation from memory strength variation is far less than that of traditional method, and yet, this proves again that ISC is resilient to not only long loss bursts, but also to short random bursts. In addition, as long as ISC algorithm is considered in the system, the number of sub-stream increment has almost no effect in quality improvement, therefore, it is safe to say that two sub-stream ISC would be the most efficient way to code a stream without risking long search time to find an optimal interleaving set.

Chapter 6

Forward Error Correction for Interleaved Source Coding

6.1 Forward Error Correction

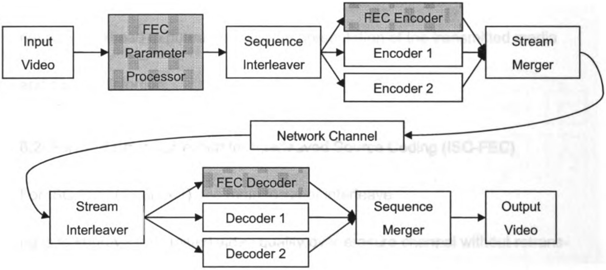


Figure 41 ISC-FEC illustration

One of the error resilient techniques that are commonly employed over packet networks is the deployment of Forward Error Correction (FEC) [6, 49-51, 53-58, 60-63]. FEC is usually used in realtime applications where retransmissions of the lost or error prone packets are not feasible. FEC recovers lost packets that occurred during transmission; however, it increases the number of redundant packets sent over the channel, which in turn could lead to an inefficient utilization of bandwidth. In particular, the FEC redundant packets reduce available bitrate for coding the original video stream, and hence this increases the overall distortion of the transmitted stream. Therefore, the key issue in the FEC scheme for the realtime media transport is measuring the acceptable number of redundancy packets to recover the lost packets in a given channel condition, while minimizing network overhead to guarantee on-time presentation of the transmitted media and coding distortion.

6.2 Forward Error Correction for Interleaved Source Coding (ISC-FEC)

For ISC-FEC (Figure 41), the main focus of interleave

ng is to improve transmitted video quality over erasure channel without retransmission. Meanwhile, the available source coding rate is decreased to a level such that the overall transmission bitrate with FEC does not exceed the total bitrate of non-FEC protected ISC (Figure 42).

Total encoding bitrate per GOV (Non-interleaving GOV size)	

||

FEC enabled encoding bitrate per GOV (Non-interleaving GOV size)	FEC Bitrate

Figure 42 Comparison of bitrate between non-FEC coding vs. FEC protected coding

Therefore, the main focus in ISC-FEC is finding a suitable bitrate for FEC redundancy packets while minimizing distortion while reducing the source encoding bitrate accordingly.

In addition, the design of ISC-FEC is intended to maximize the benefits of both the packet erasure recovery nature of FEC, and erasure concealment and short frame replacement distance of ISC. An example of ISC-FEC is illustrated in Figure 43.

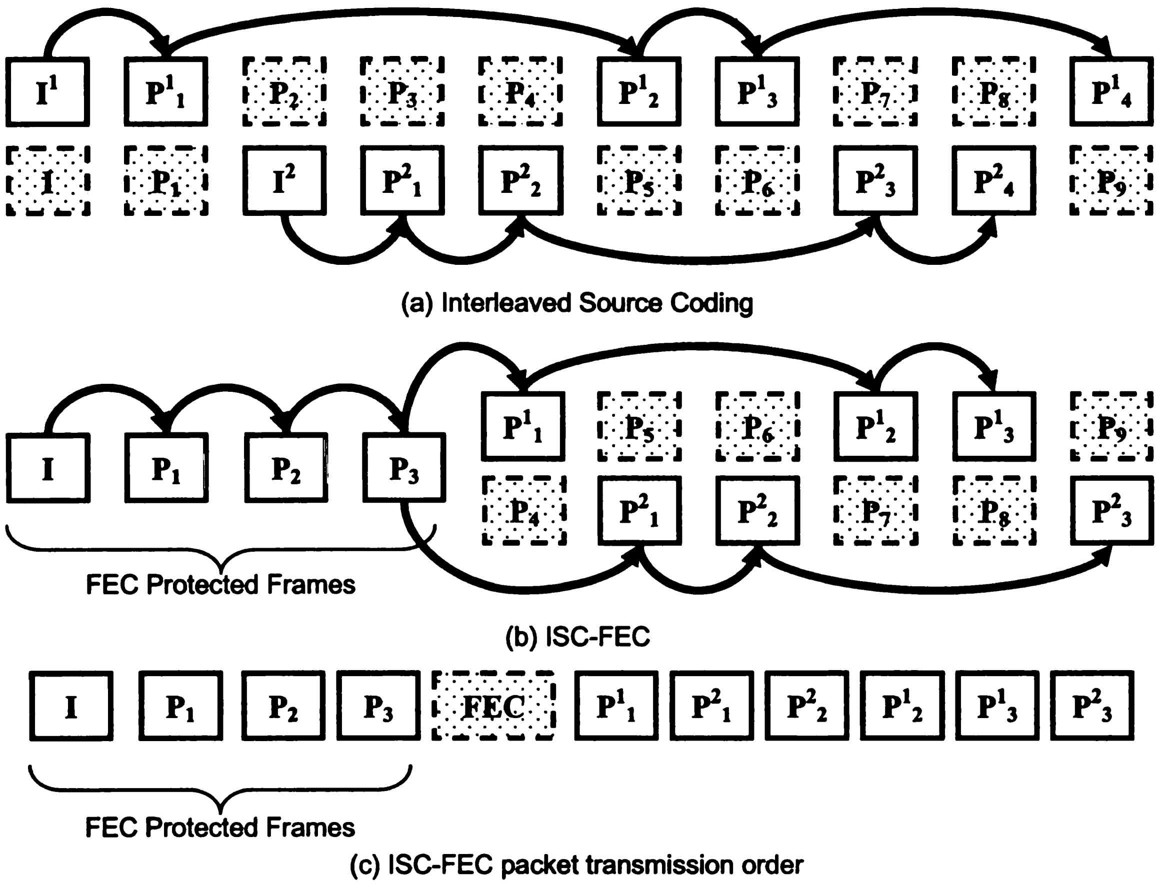


Figure 43 ISC-FEC design scheme and packet transmission illustration

Different from the original ISC design, ISC-FEC sub-streams share a common intra- coded frame I - frame and some inter- coded frames P - frame based on the assumption that FEC will protect those frames from losses. Hence the length of FEC redundancy packets and the number of FEC protected frames vary depending on the packet loss rate and maximum distortion variation allowed with FEC. From the observations in Chapter 5, since multi-stream interleaving performances are very close to each other, the frames that are not protected with FEC will be interleaved into two sub-streams. There is a risk involved with this design, if FEC fails, the quality degradation risk is almost the same as that of non-ISC coding. Therefore, when computing the number of R-D optimized FEC protected frame, this number should be carefully selected so that it does not exceed the FEC recovery rate. The design will consider maximum distance separable (MDS) codes, i.e., Reed Solomon codes, for FEC[49-59]. In MDS code, usually defined as $[n,k]$, where k is the number of message packets and $h = n - k$ is the number of FEC redundancy packets. If any k packets are received out of the n packet transmission block, the system will be able to recover any lost packets in k. Hence h/n is the loss recovery rate of the given system. However, if the number of lost packets are greater than h, the risk factor of ISC-FEC, the system won't be able to recover the lost packets, hence, ISC-FEC will use any received message packets to decode some frames that are not affected by packet losses, and yet, partial decoding is still not an option as in the original ISC. In addition, the rate assigned to the FEC-protected portion of the video stream and the rate allocated to each interleaving sub-stream is determined based on the number of frames in each corresponding part such that the overall (total) bitrate remains the same regardless the pattern of ISC-FEC.

6.2.1 Rate-Distortion Optimized ISC-FEC

Since ISC finds optimal interleaving pattern specific to channel condition and sequence, ISC-FEC assumes that R-D function can also be found empirically as a preprocess information. A curve fitting method with the Minimum Mean Square Estimator (MMSE) is used to obtain a R-D function for a given sequence.

$$\underset{\{a,b,c\}}{\arg\min}\left[MSE\left\{D(R), a\times\exp(-R)+c, \forall R\right\}\right] \tag{42}$$

$$D(R) = a\times\exp\left(-R^{b}\right)+c \tag{43}$$

With a sequence specific R-D function, ISC-FEC computes the number of FEC redundancy packets for MDS with predefined variable, e.g., number of frames in a GOV (un-interleaved size), frame rate, bitrate, packet loss rate, and maximum coding distortion (%). The computation process to find number of FEC packets, FEC protected frames, as well as FEC protected packets is shown in the following pseudo code:

```
Compute total number of packets per second, totPkt,  using current bitrate, bitrate / packetsize
Compute current distortion, iDistortion using current bitrate (43)
fDistortion = iDistortion * (1 + ΔDistortion /100)
using fDistortion new coding bitrate, compute nbitrate, by reversing (43)
if nbitrate < bitrate
  k = ⌊newbitrate / packetsize⌋ : Number of message packets per second
  h = totPkt − k : Number of FEC packets per second
  η = ⌈k / framerate⌉: Average number of packets per frame
  totPkt_GOV = η × N_GOV : Total number of packets in a GOV
  h_GOV = (h / framerate) × N_GOV : Number of FEC packets per GOV
  If h_GOV ≥ 1 : break point 1
     n_GOV = ⌊h_GOV / p_1⌋ : Maximum number of packets protected over channel with packet loss rate p_1
     If n_GOV < totPkt_GOV : break point 2
        k_GOV = n_GOV − h_GOV : Number of FEC protected packets
        N_FEC = ⌊(k_GOV / η)⌋ : Number of frames protected with FEC in a GOV
        If N_GOV − N_FEC = odd
           N_FEC = N_FEC − 1 : Make the remainder of GOV into even number for two sub-streams without sacrificing recovery rate
        end
        If N_FEC > 0 : break point 3
           Return N_FEC, n_GOV, h_GOV
end
```

There are break points set up in the code: *break point 1* stops the process if the number of FEC packets, h_{GOV}, is less one since there is no full packet length FEC redundancy packet. It is possible to pad the packet with zero to make a full length FEC redundancy packet, however, if the bits used in zero padding add up over time, it could result the coded video to exceed assigned bitrate. Therefore, to prevent such ripple effect, FEC parameter computation process terminates in such case. *Break point 2* is set to stop process if the total number of FEC coded packets in a GOV, n_{GOV}, is greater than total number of packets assigned to a GOV. In such case, all the frames in a GOV can be protected with FEC, therefore, interleaving is not necessary. This usually happens when the maximum distortion variation allowance is set to high. *Break point 3* is set before returning the computed values to the Sequence Separator. It is set to stop the process if number of FEC protected frame, N_{FEC}, is equal to zero. Since the system encodes the remainder of non-FEC protected frames in a GOV into two sub-streams, the number of frames associated with interleaving must be in even number. To satisfy such constraint, N_{FEC} is reduced by one frame, which in some cases, become zero. It is possible to increase N_{FEC} by one frame, however, this can cause the system to exceed the FEC recovery rate.

6.2.2 ISC-FEC

When FEC parameter computing process returns values, N_{FEC}, n_{GOV}, and h_{GOV}, ISC-MDP goes through the optimal interleaving set search process for the remainder of the frames in a GOV that are not protected with FEC with following:

$$\begin{gathered} s = \left\{ 0 \quad 1 \quad \cdots \quad N_{GOV} - N_{FEC} - 1 \right\} = \bigcup_{j=1}^{2} s^{(k,j)} \\ \bigcap_{j=1}^{2} s^{(k,j)} = \varnothing, \qquad \forall j, size\left(s^{(k,j)}\right) = \left(N_{GOV} - N_{FEC}\right)/2 \end{gathered} \tag{44}$$

Since these interleaving patterns references to the last frame in FEC protected portion of the GOV, the interleaving set is redefined as:

$$\begin{array}{c} s^{*(k,1)}(n) = \{0 \quad s^{(k,1)}(n) - s^{(k,1)}(0) + 1\} \\ s^{*(k,2)}(n) = \{0 \quad s^{(k,2)}(n) + 1\} \end{array}, \quad \text{for } 0 \le n \le \left(N_{GOV} - N_{FEC}\right)/2 \tag{45}$$

Once the set is defined, ISC-FEC process finds an optimal interleaving set for the frames in the remainder of the GOV with the ISC-MDP equations, (31)-(38), and the decision criteria and equations, (19)-(30)

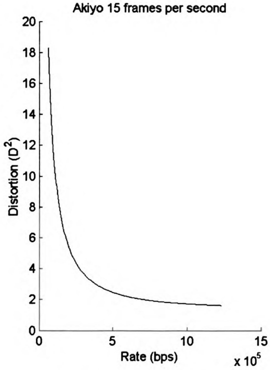
For encoding and decoding of the interleaving portion of the GOV, the system uses interleaving set parameter from (44) instead of (45).

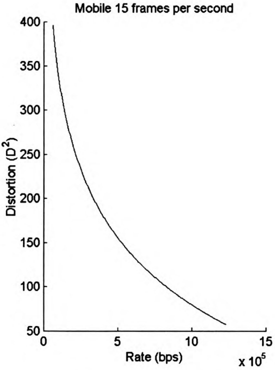
6.3 ISC-FEC Evaluation and Analysis

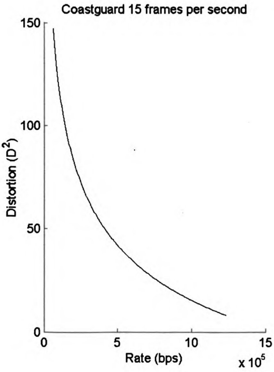
6.3.1 Simulation Setup

For evaluation, CIF sequences of *Akiyo*, *Foreman*, *Coastguard*, and *Mobile* were coded into an *IPPP…* GOV structure using an MPEG-4 encoder. GOV sizes (un-interleaved size) of 10, 12, 14, 16, 18, and 20 were used to protect and partition the evaluation sequences. Frame rate of 15 frames per second, bitrate of *250kbps* and *500kbps*, and packet size of *512 Byte* were used for encoding the sequences. Network conditions were set to 5%, 10%, and 15% packet loss rates, p_1, with varying packet correlation value, ρ, of 0.3, 0.6, and 0.9. For each network condition $p_1 - \rho$ pair, ten packet loss traces were generated. For FEC parameter computation, maximum distortion variation of 1~5% were used. Since ISC-FEC R-D optimized protection and interleaving pattern search algorithm targets to a specific sequence, sequence specific correlation model is used to find an optimal interleaving set for ISC. In addition, the ISC-FEC results were compared with sequence specific ISC, ISC-SC. Each evaluation case is fitted into these packet loss traces and the PSNR values are averaged to provide statistically satisfying results for analysis.

6.3.2 Simulation Results







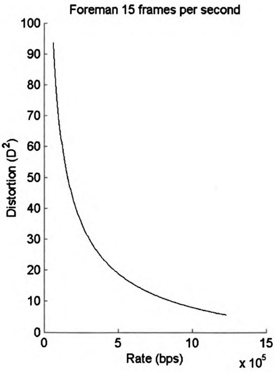


Figure 44 Rate-Distortion of the evaluation sequences

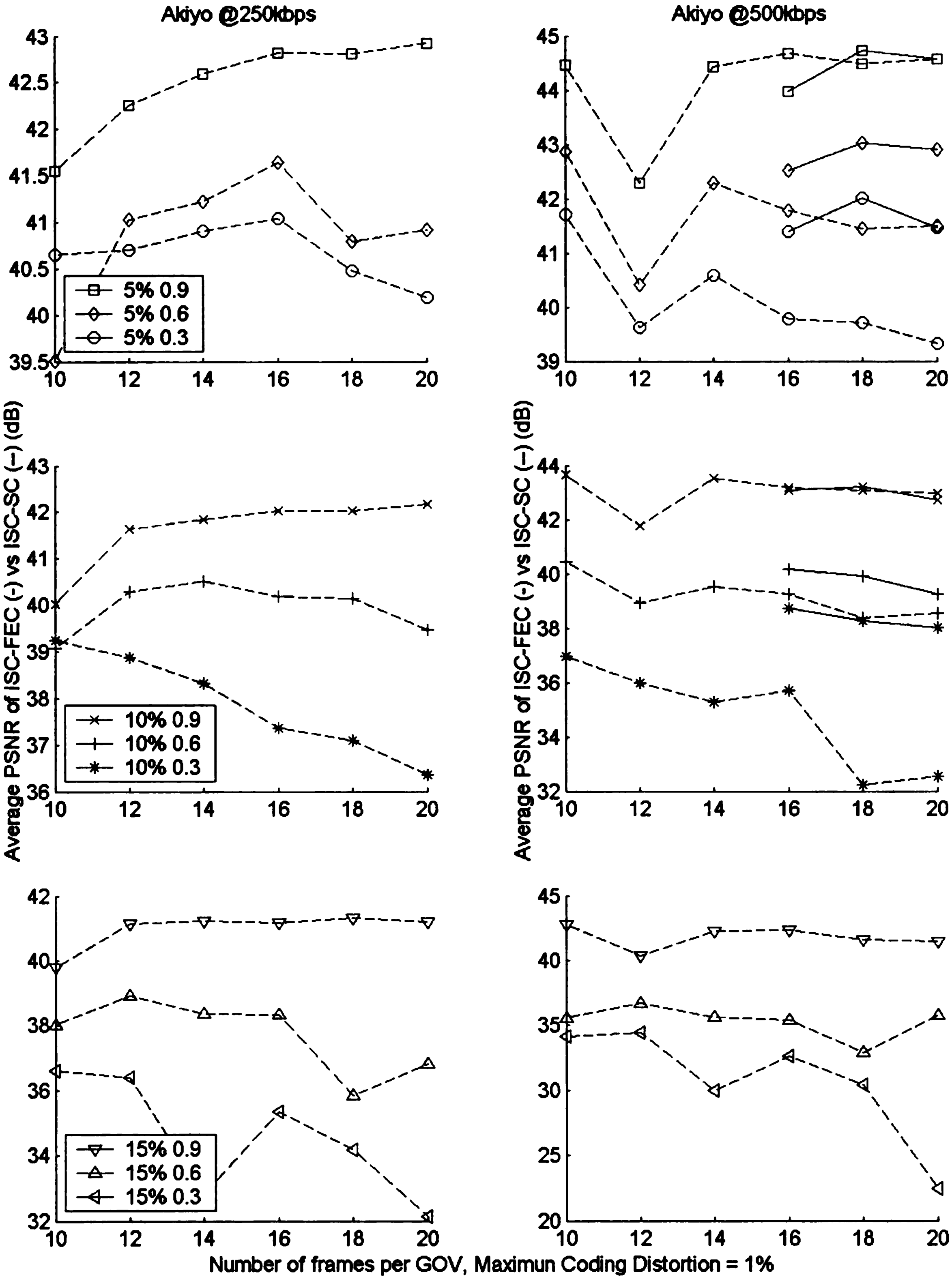


Figure 45 ISC-FEC for Akiyo, Maximum coding distortion = *1%*

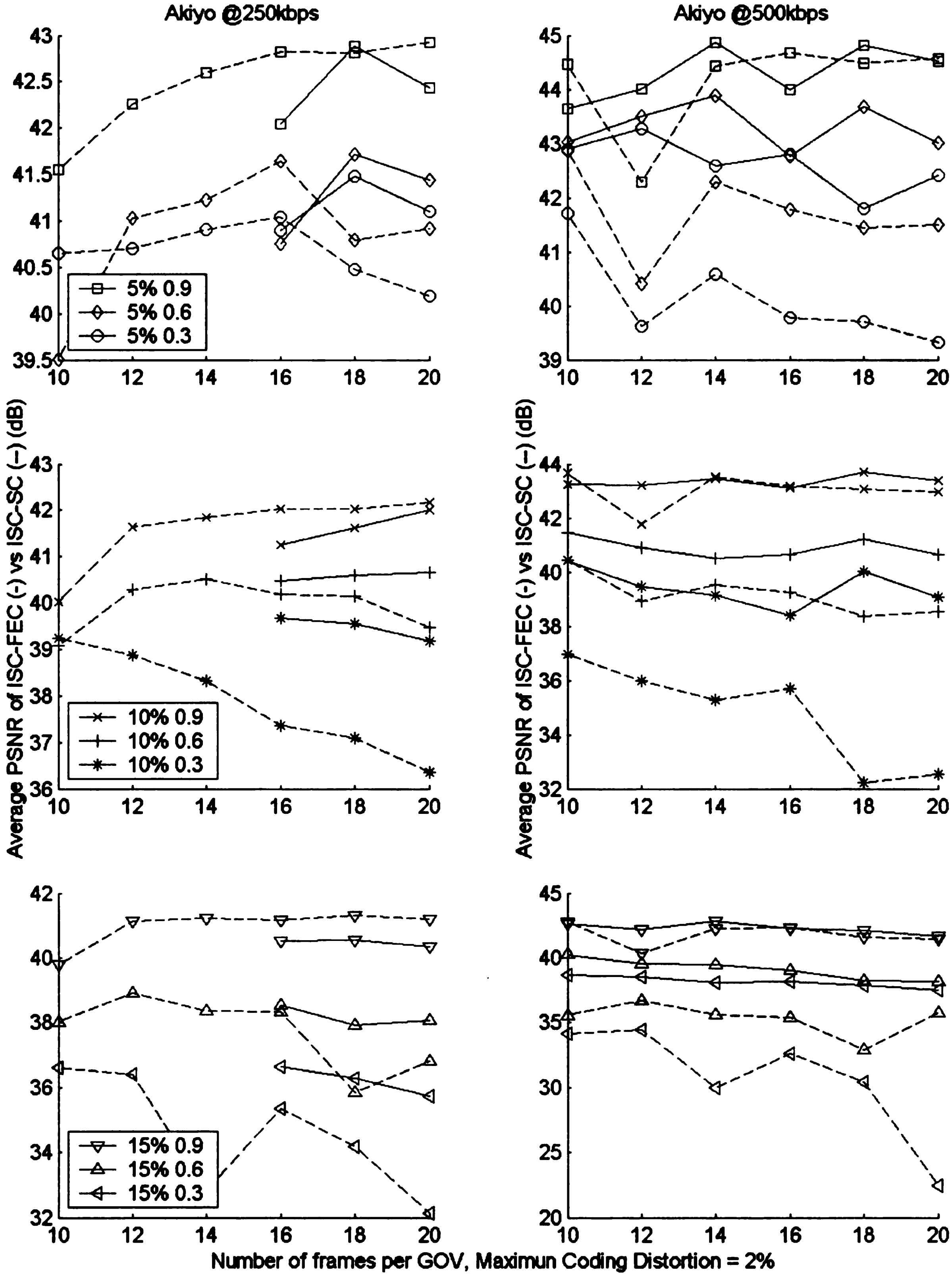


Figure 46 ISC-FEC for Akiyo, Maximum coding distortion = *2%*

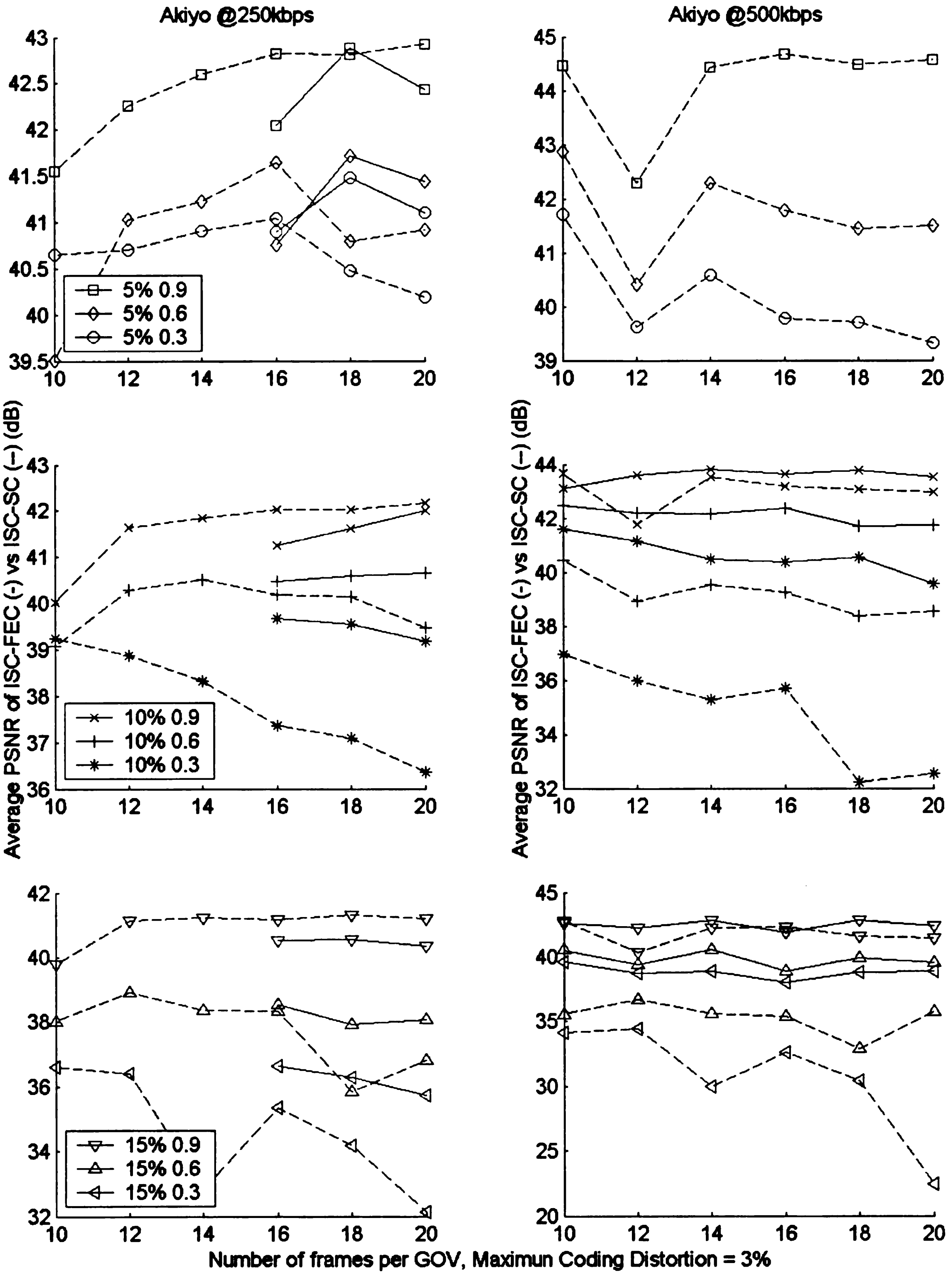


Figure 47 ISC-FEC for Akiyo, Maximum coding distortion = *3%*

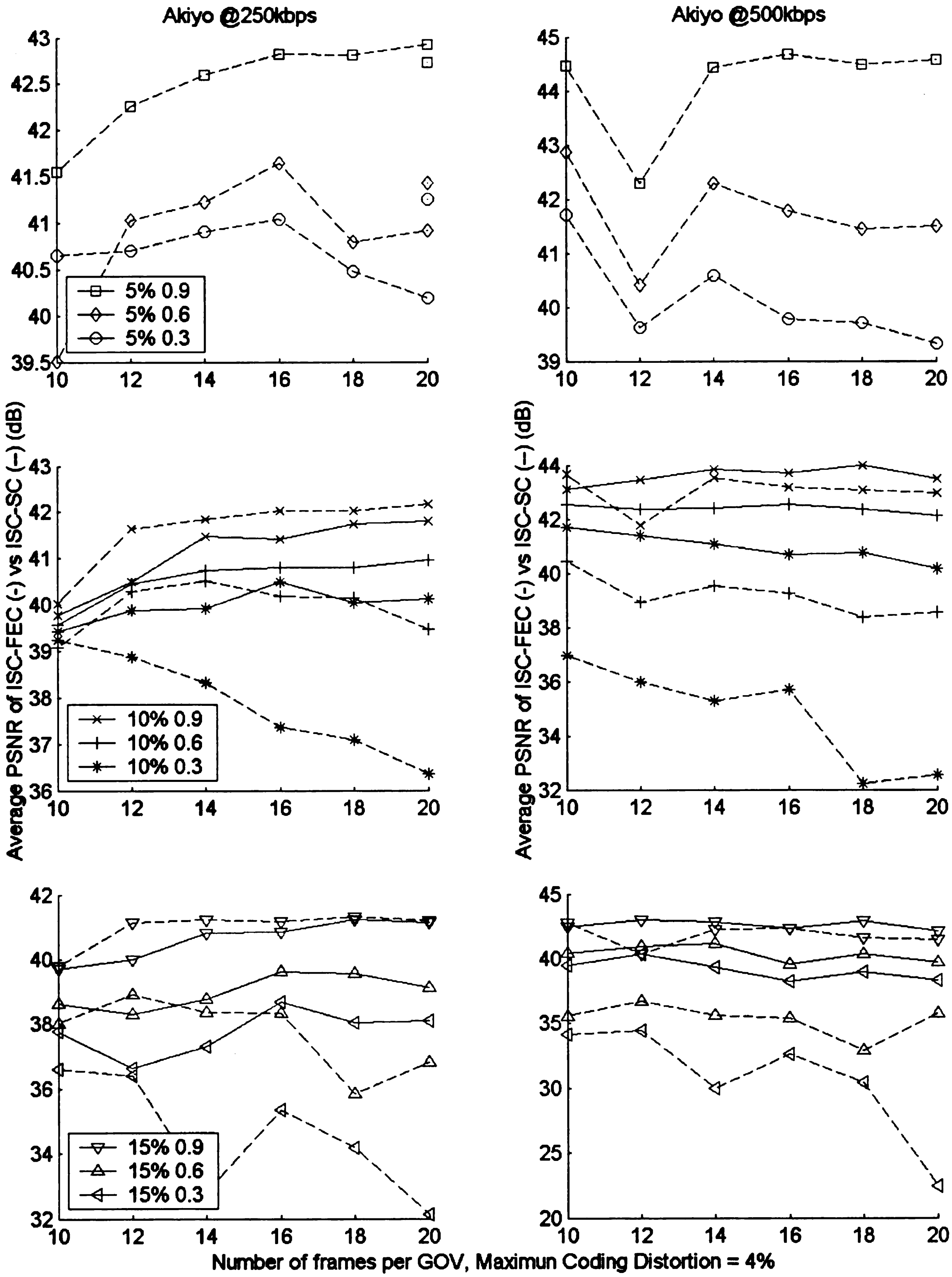


Figure 48 ISC-FEC for *Akiyo*, Maximum coding distortion = *4%*

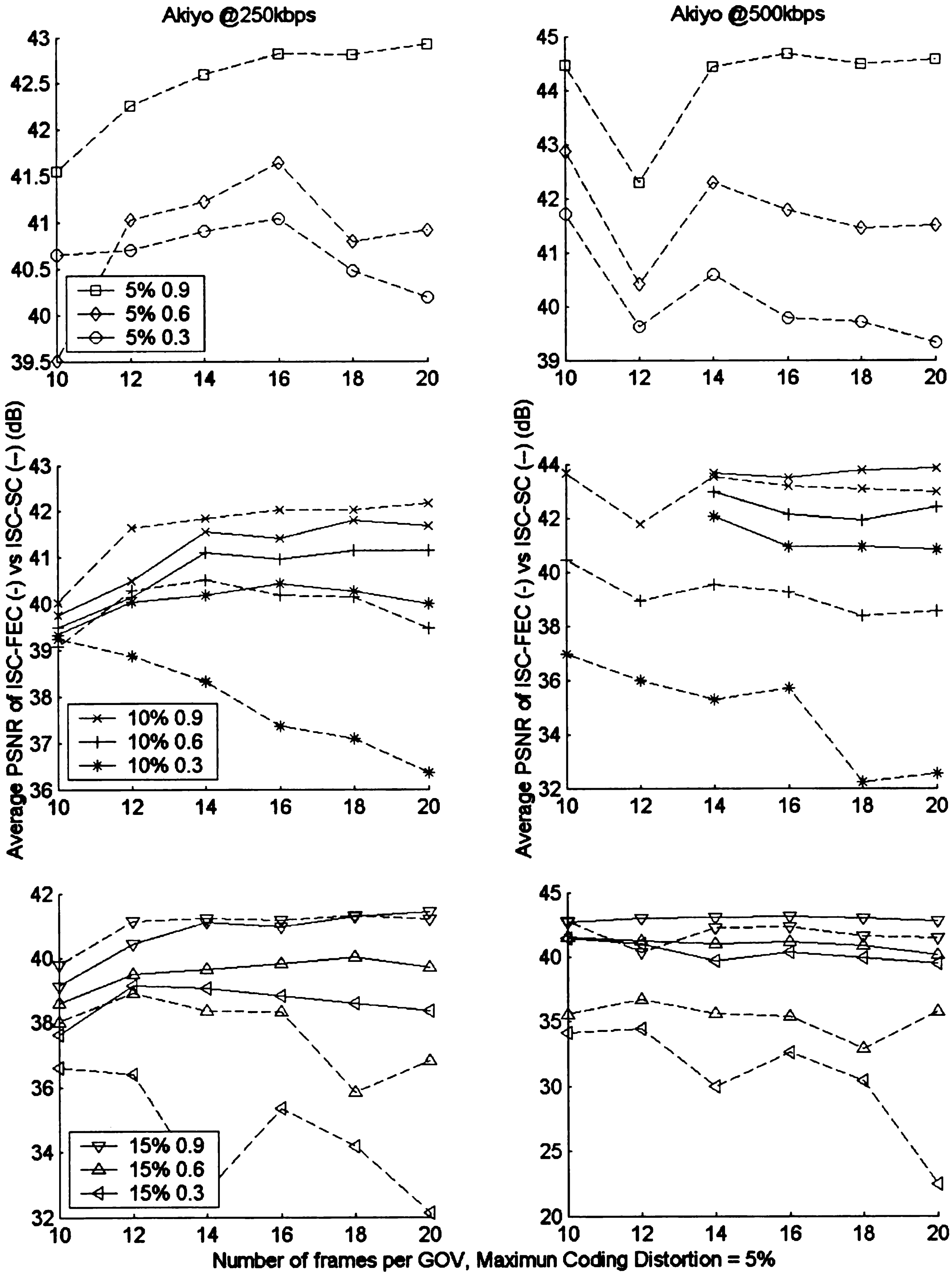


Figure 49 ISC-FEC for *Akiyo*, Maximum coding distortion = *5%*

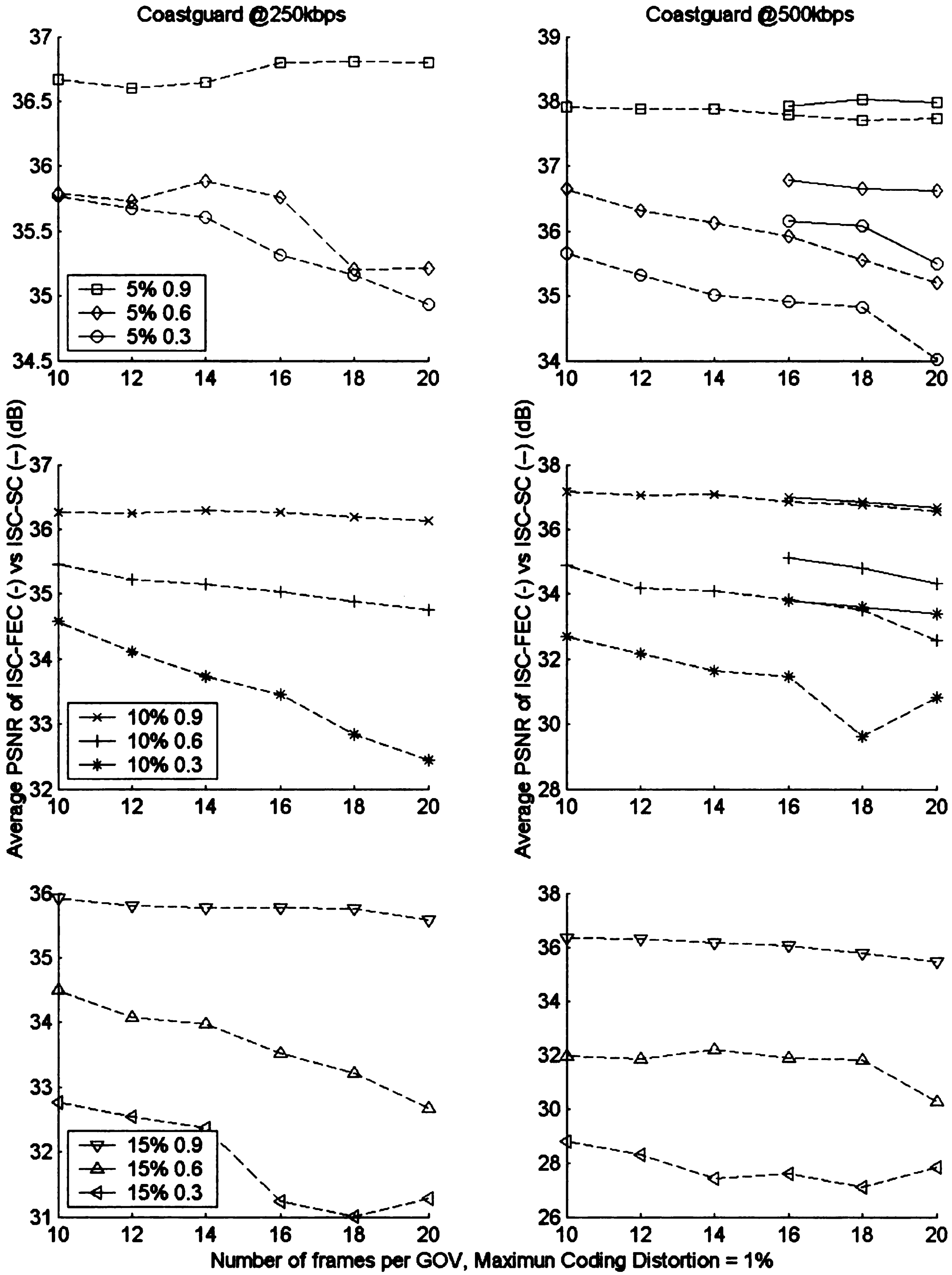


Figure 50 ISC-FEC for *Coastguard*, Maximum coding distortion = *1%*

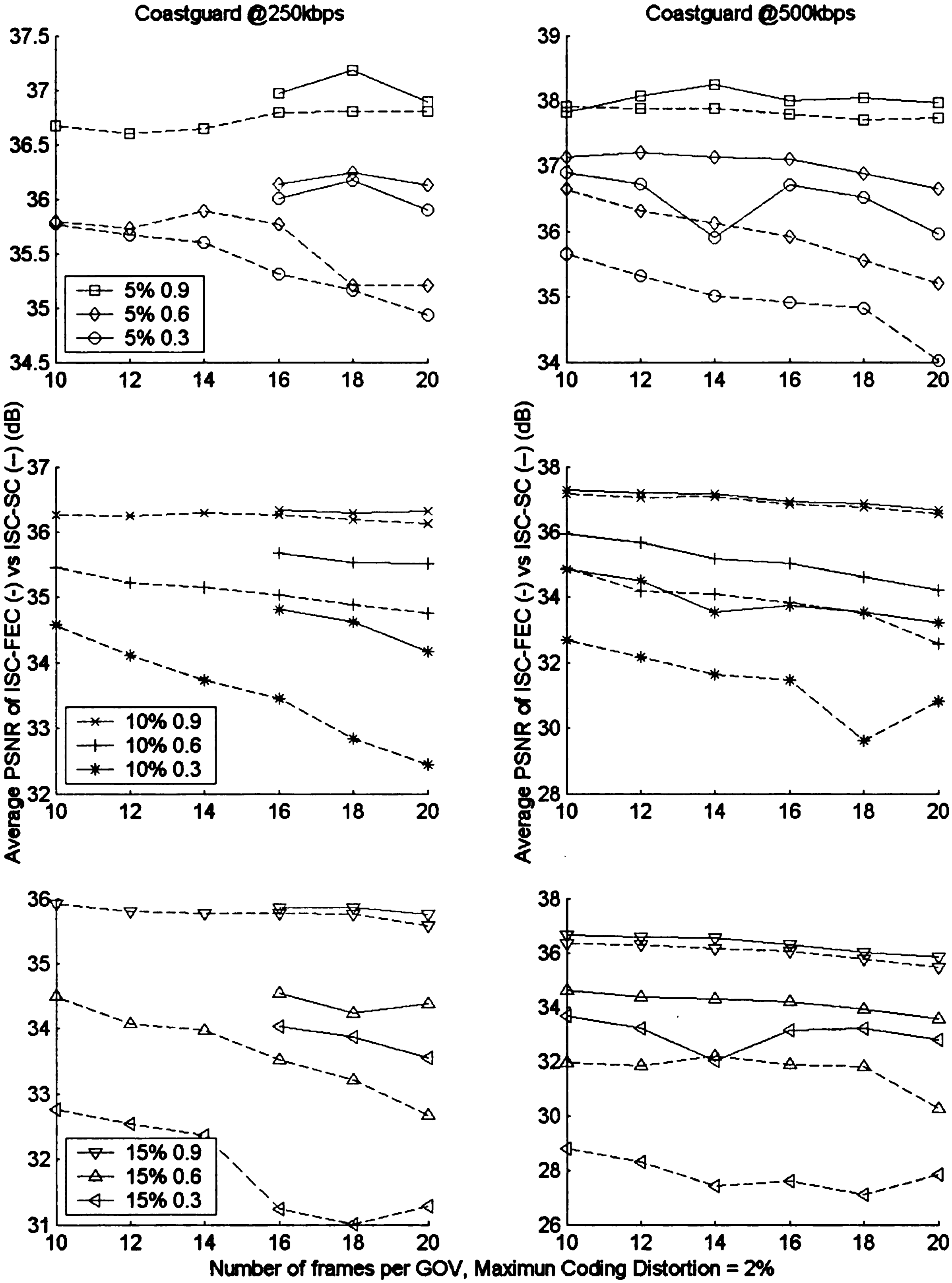


Figure 51 ISC-FEC for *Coastguard*, Maximum coding distortion = *2%*

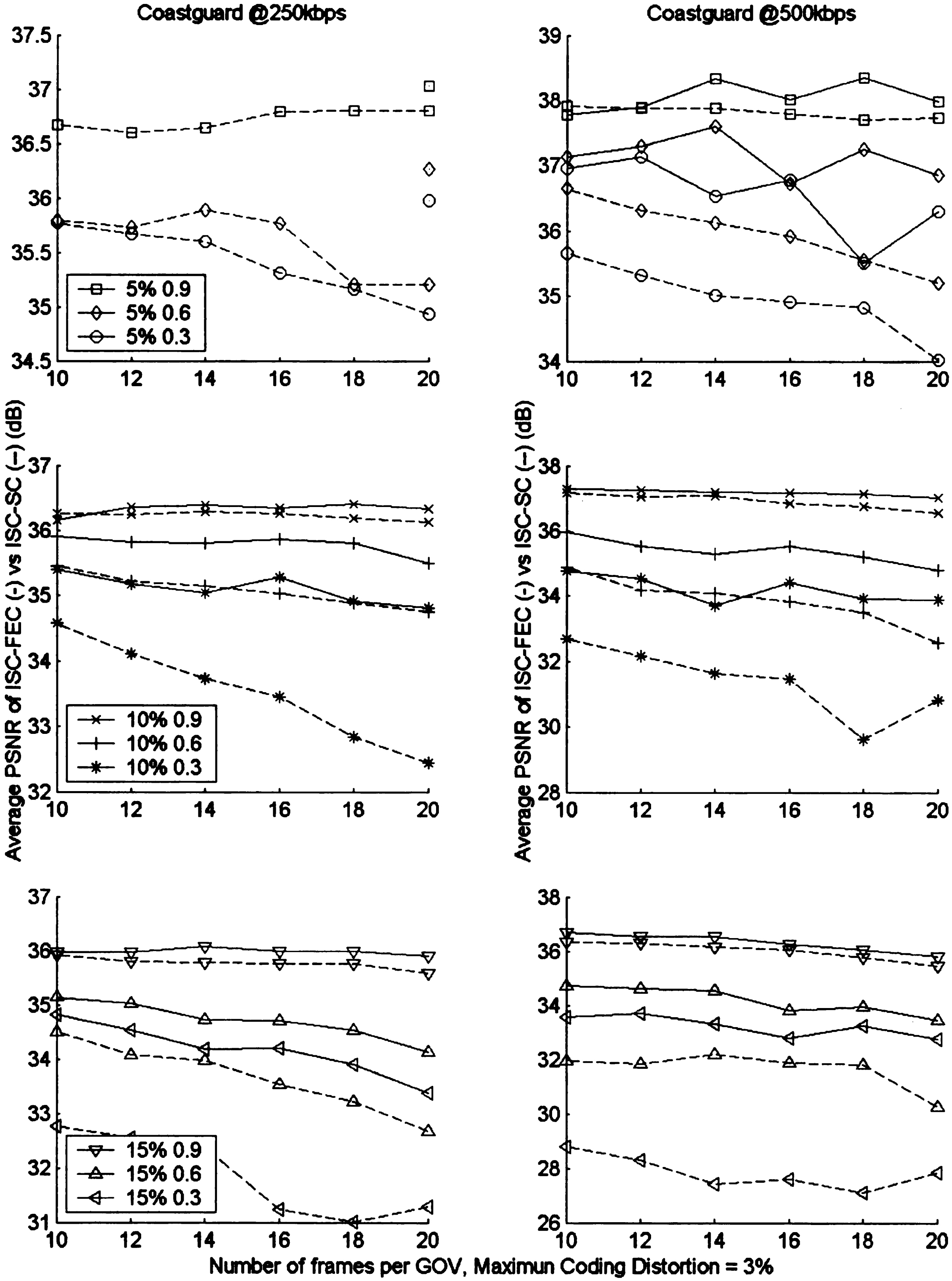


Figure 52 ISC-FEC for *Coastguard*, Maximum coding distortion = *3%*

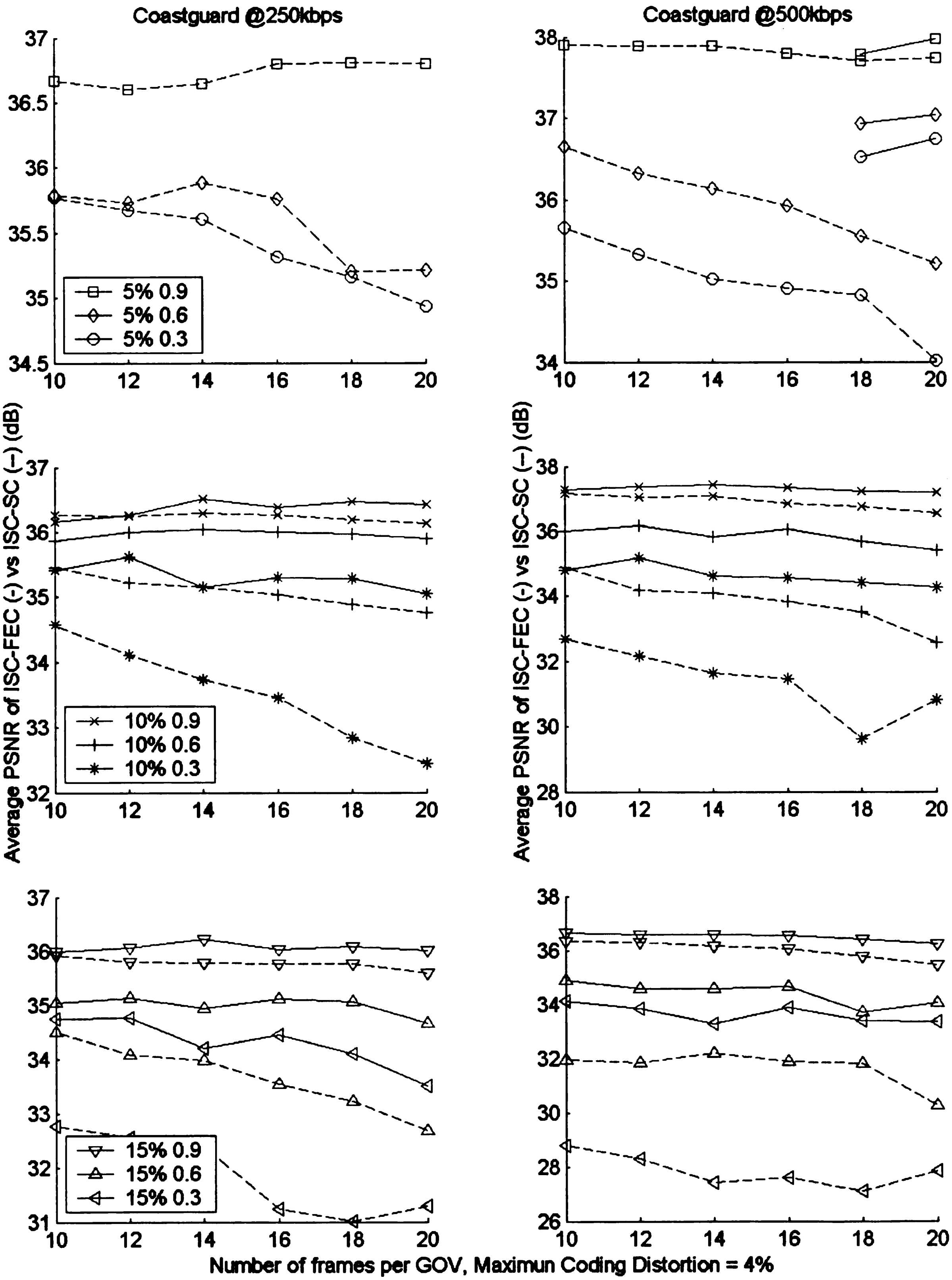


Figure 53 ISC-FEC for *Coastguard*, Maximum coding distortion = *4%*

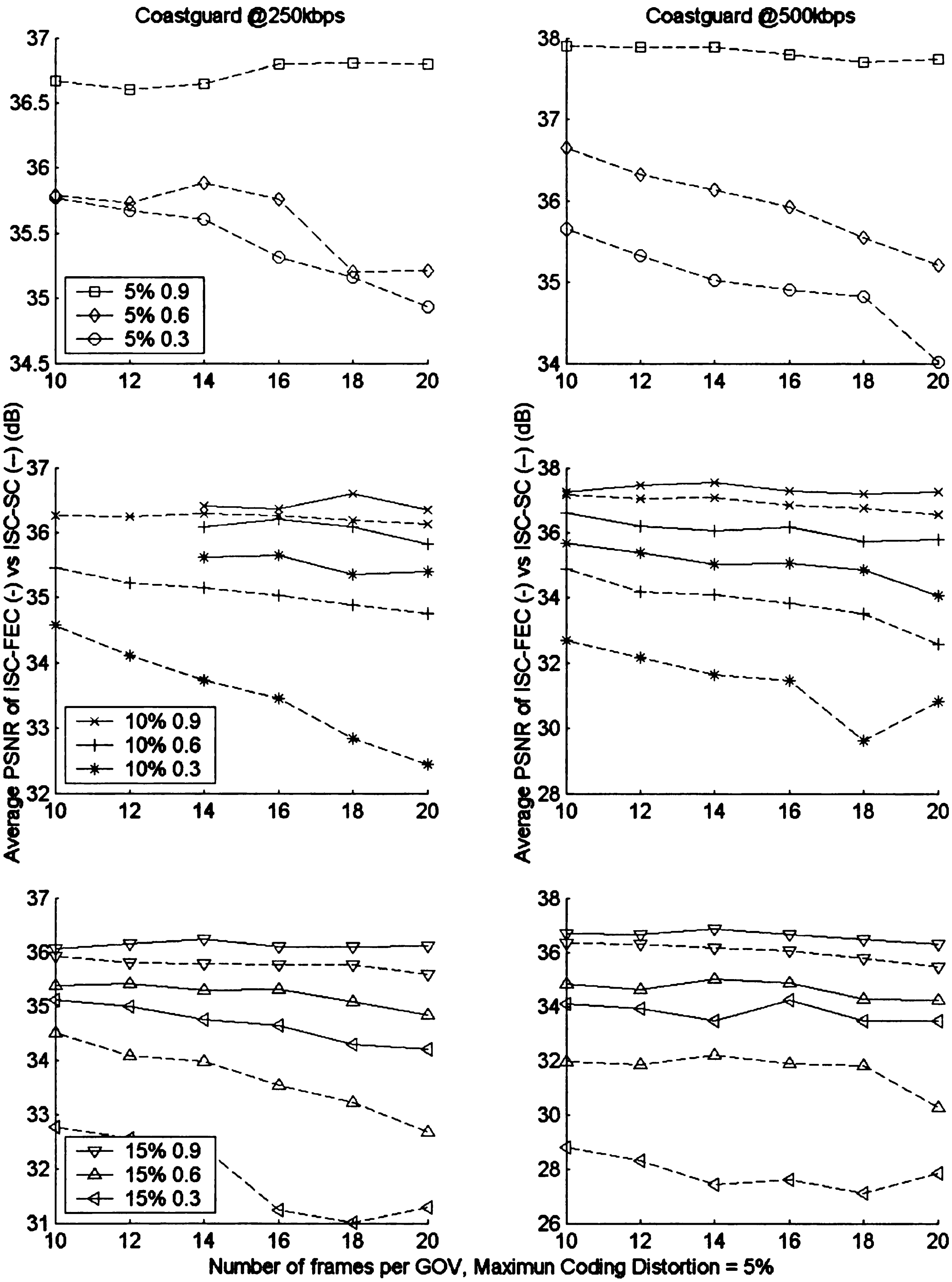


Figure 54 ISC-FEC for *Coastguard*, Maximum coding distortion = *5%*

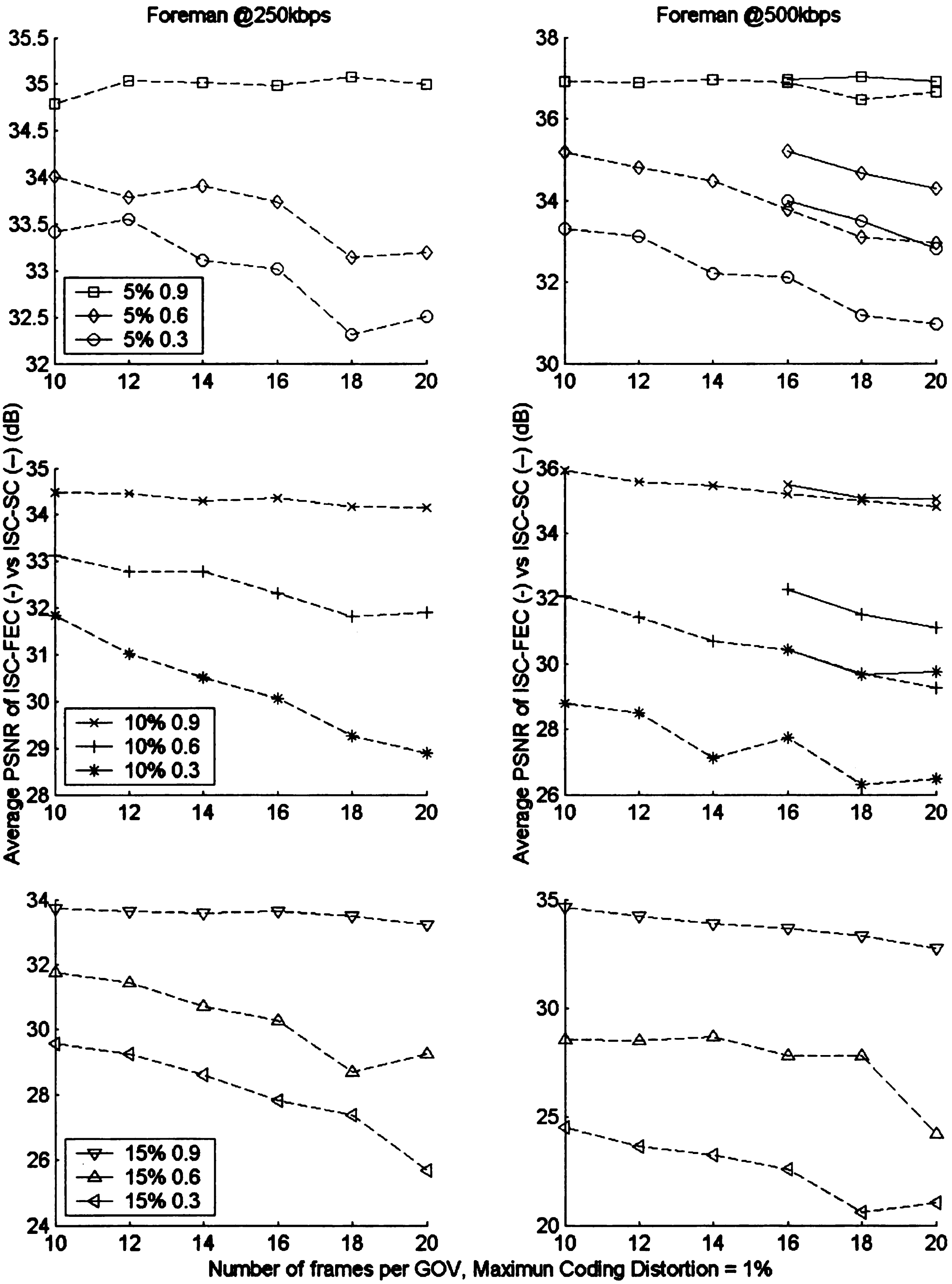


Figure 55 ISC-FEC for *Foreman*, Maximum coding distortion = *1%*

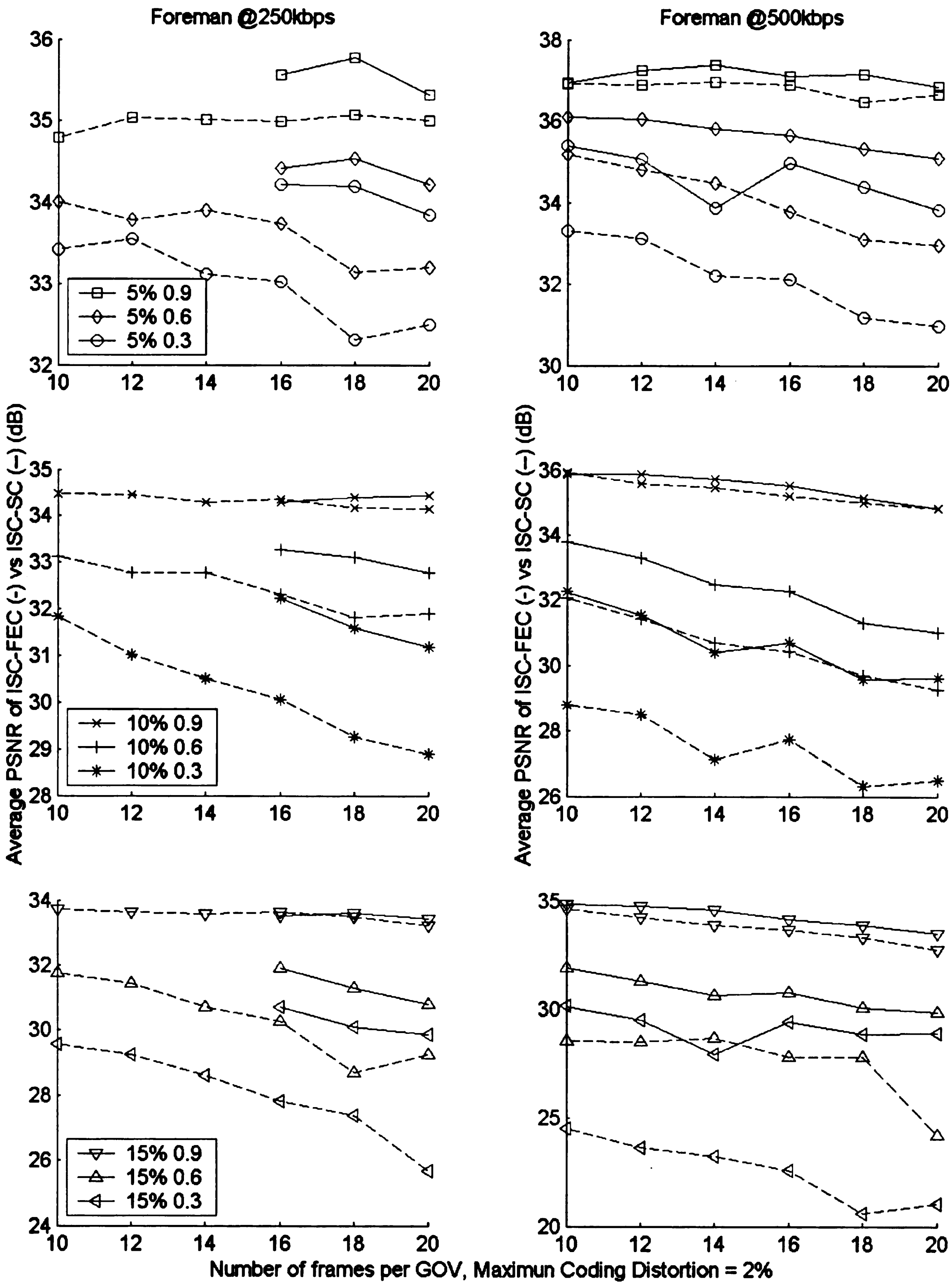


Figure 56 ISC-FEC for *Foreman*, Maximum coding distortion = *2%*

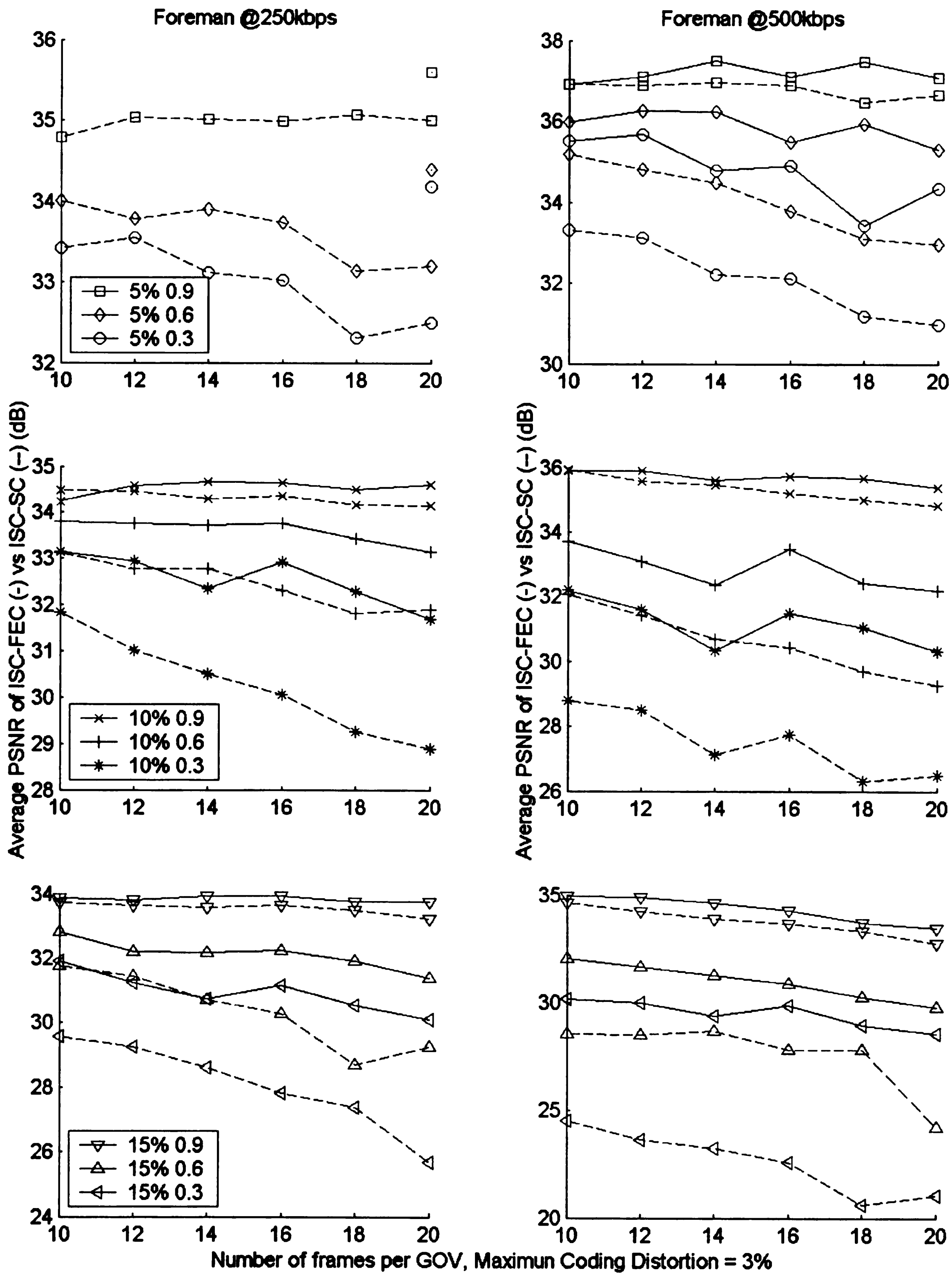


Figure 57 ISC-FEC for *Foreman*, Maximum coding distortion = *3%*

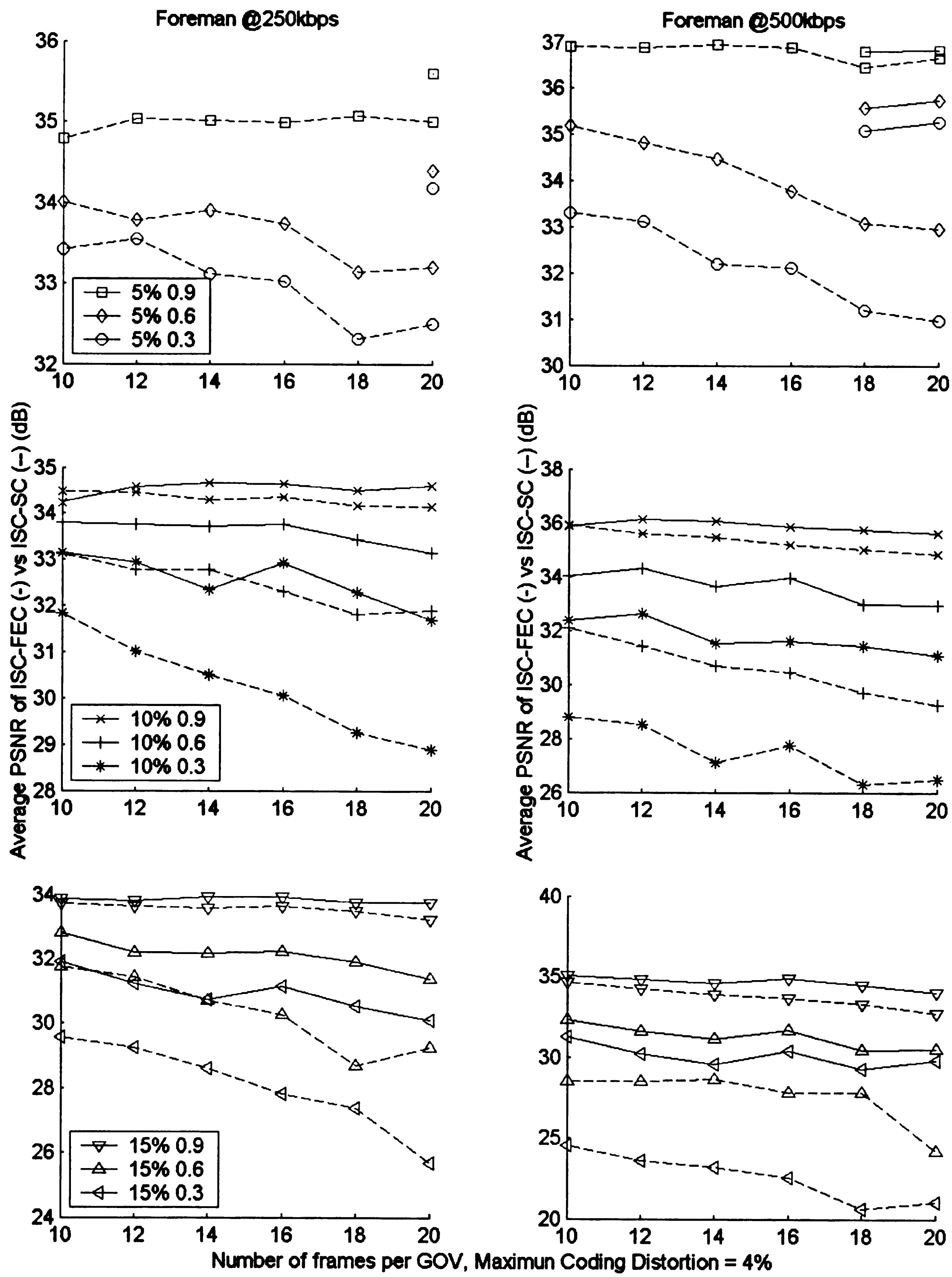


Figure 58 ISC-FEC for *Foreman*, Maximum coding distortion = *4%*

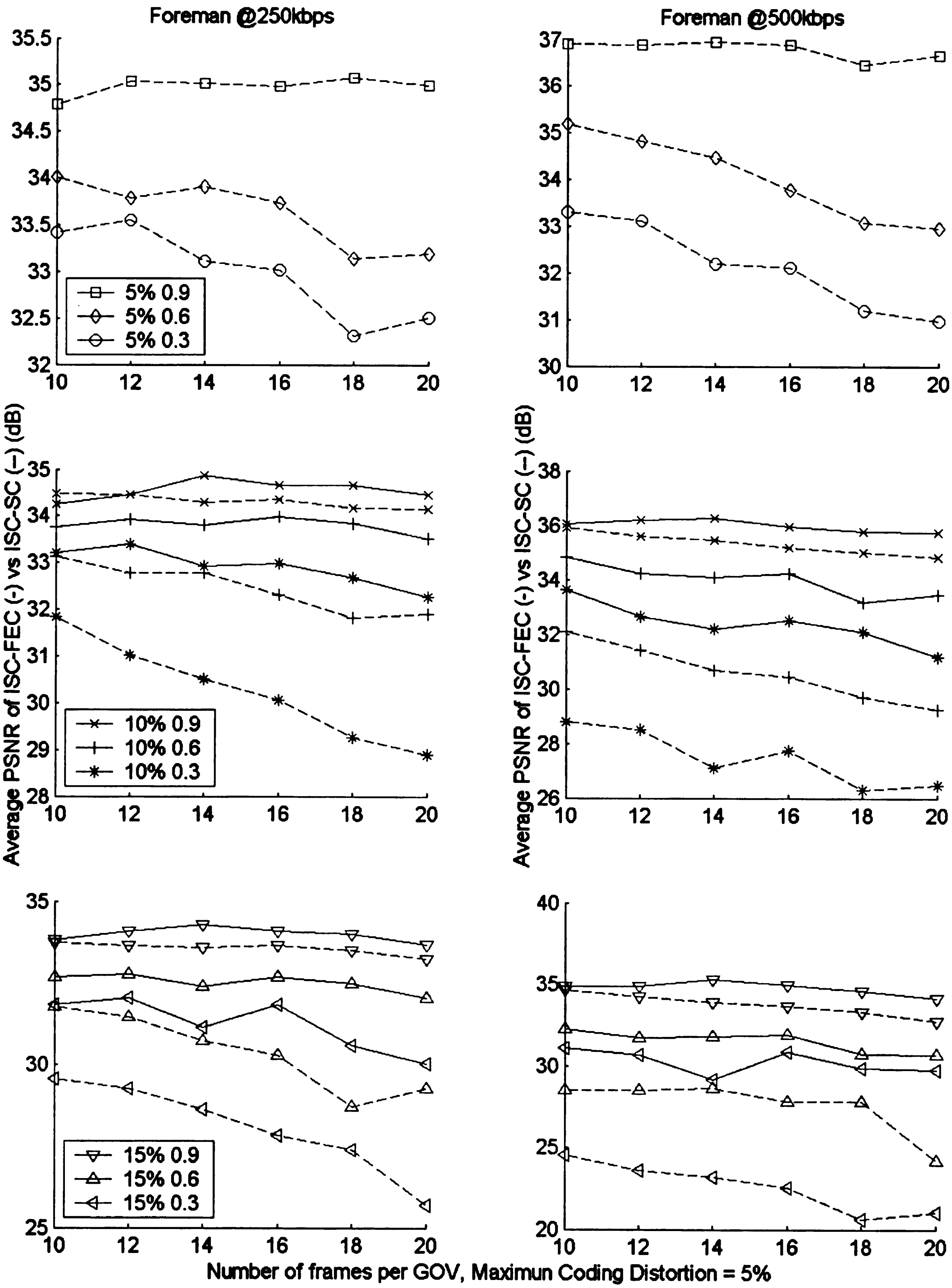


Figure 59 ISC-FEC for *Foreman*, Maximum coding distortion = *5%*

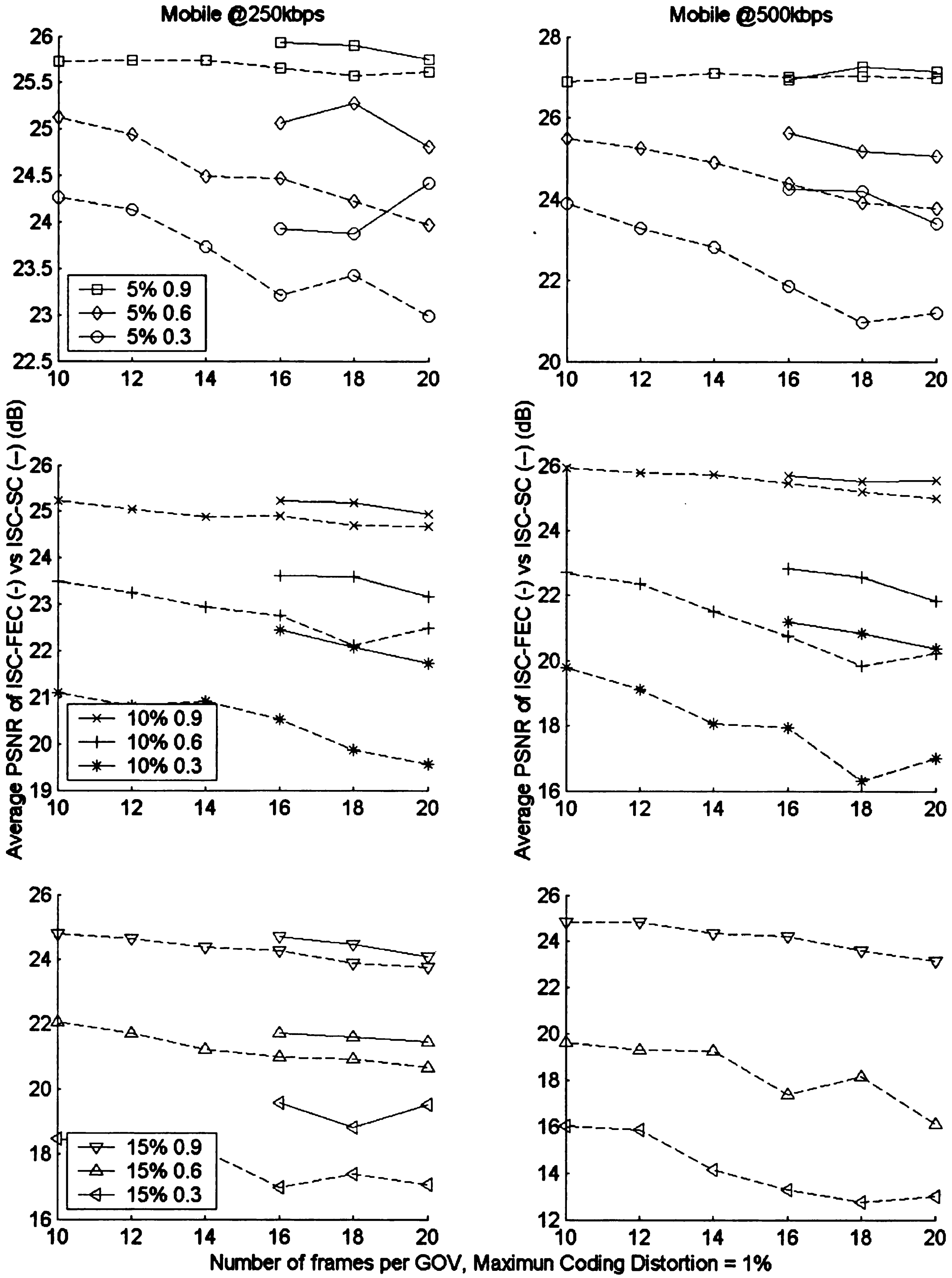


Figure 60 ISC-FEC for *Mobile*, Maximum coding distortion = *1%*

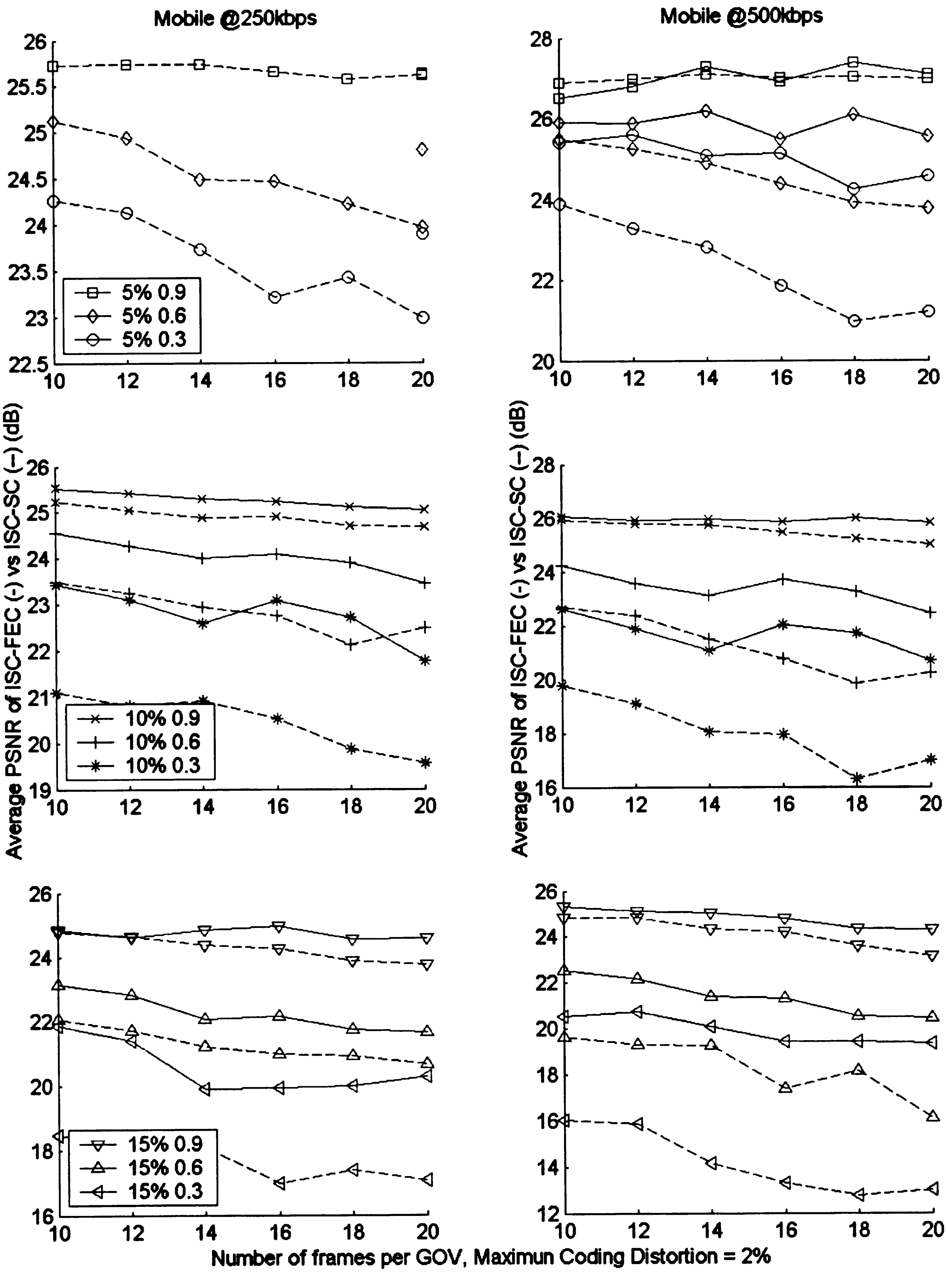


Figure 61 ISC-FEC for *Mobile*, Maximum coding distortion = *2%*

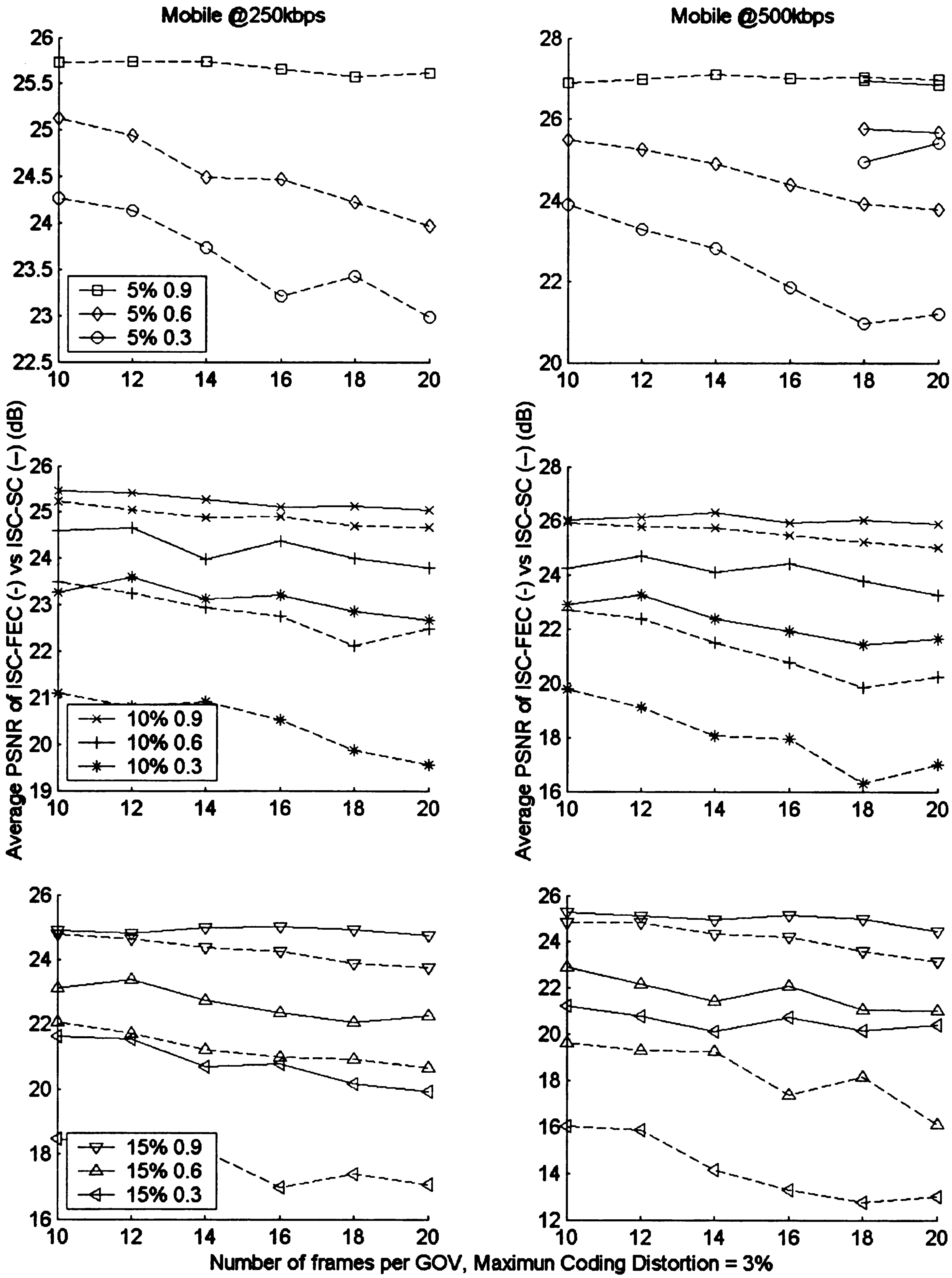


Figure 62 ISC-FEC for *Mobile*, Maximum coding distortion = *3%*

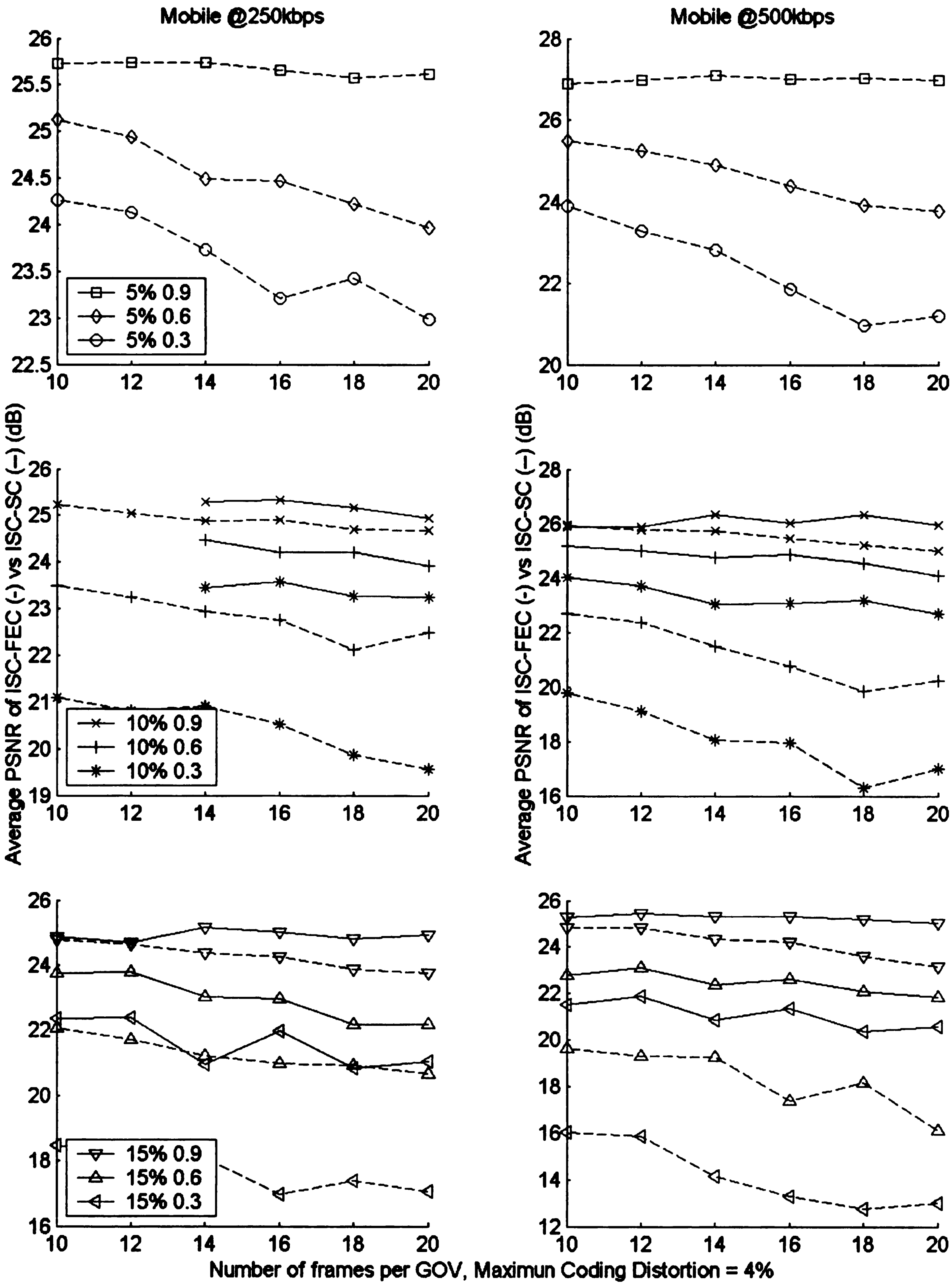


Figure 63 ISC-FEC for *Mobile*, Maximum coding distortion = *4%*

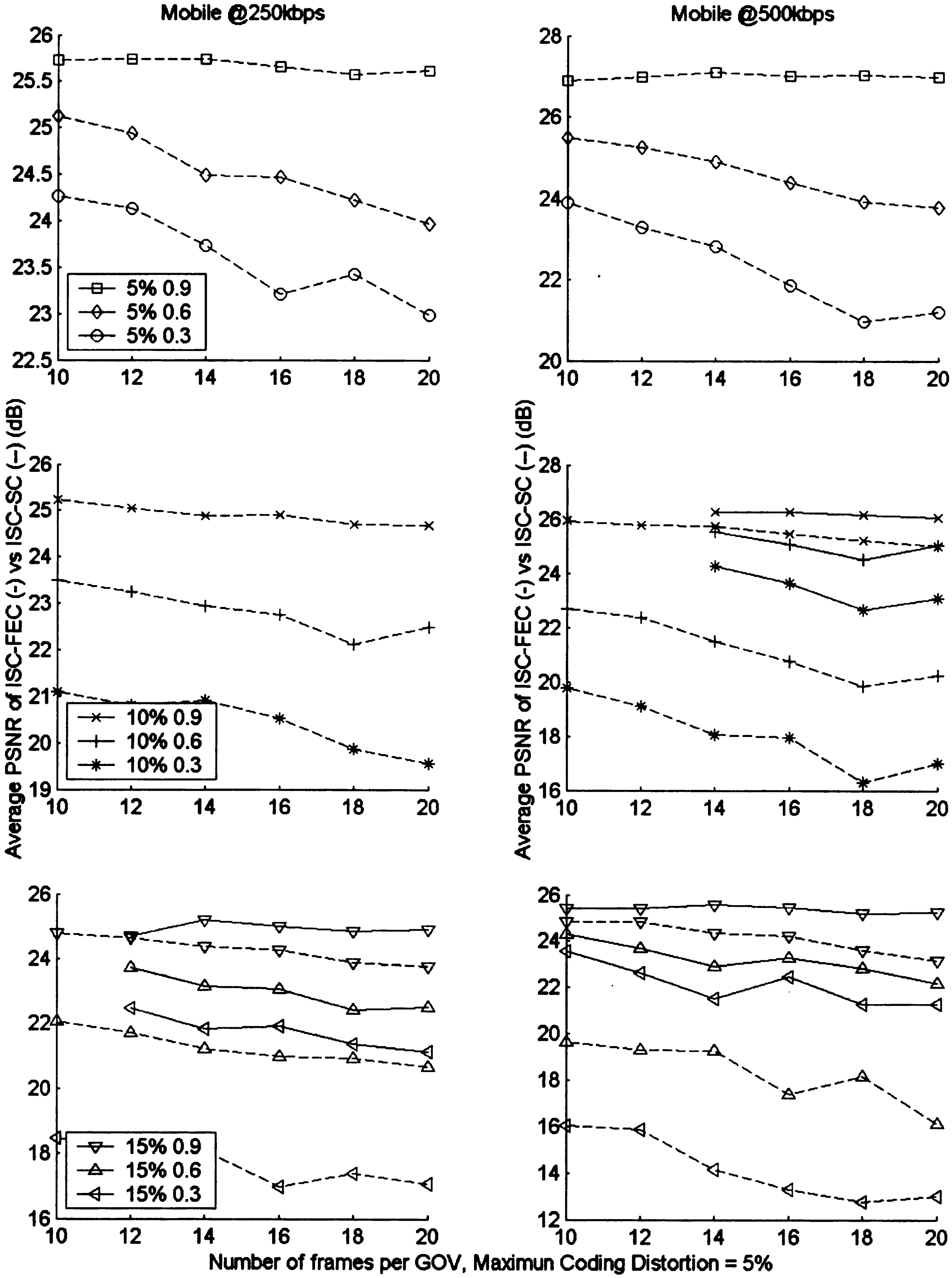


Figure 64 ISC-FEC for *Mobile*, Maximum coding distortion = *5%*

6.3.3 Analysis

6.3.3.1 Channel condition and GOV size variation effects

When FEC parameter search process successfully returned R-D optimized ISC-FEC encoding parameters, following observations are made: when the channel shows strong memory, $\rho = 0.9$, ISC-FEC shows quality improvement over ISC-SC regardless of packet loss rate or GOV size in most of the cases. This is due to the fact that the long packet loss burst in FEC protected portion can exceed FEC recovery rate, hence results ISC-FEC to lose following non-FEC protected, interleaved frames. However, as the memory strength gets weaker, in other words, as the randomness of packet loss increases, ISC-FEC shows noticeable improvement over ISC-SC. As the chance of FEC protected portion to be fully decoded increases, the cascaded loss effect on interleaving portion decreases; hence, the total number of frames that can be successfully decoded is higher than that of ISC-SC.

6.3.3.2 Distortion Variation

With the increment of maximum coding distortion allowance, all the evaluation cases have shown performance improvement compare to the ones with less coding distortion allowance. In addition, overall performance compared with ISC-SC has also improved. Even though coding distortion increment decreases per frame quality in lossless decoding, the increased number of successfully decoded frame in ISC-FEC and the overall lossless decoding quality from them can easily overcome the effect from coding distortion increment. In fact, if a sequence is coded under the same channel condition and/or with the same total GOV size, increased coding distortion in ISC-FEC reduces the number of unprotected frames in a GOV. This results smaller sum of temporal prediction distances in the interleaving portion of ISC-FEC compared with ISC-SC; hence can also increase overall quality of decoded stream.

6.3.3.3 Evaluation Summary

Overall observation shows that the proposed ISC-FEC method improves performance of ISC-SC, even with the risk of total GOV loss or coding distortion, except for some cases where the channel has very strong memory. Similar to the observation made in previous parts, the most significant quality degrading factor is randomness of packet losses, or strength of memory. However, with the adaptation of FEC, the effect of memory is reduced, and yet provides more flexibility to ISC-MDP for different channel models.

Chapter 7

Conclusions and Future Work

7.1 Conclusion

In this dissertation, Interleaving Source Coding framework for packet video over erasure channel has introduced and investigated. The main purpose of ISC framework is to provide packet loss resilient video-coding for predictive video sequences over single erasure channel by reducing the frequency and impact of the cascaded effect of packet erasures and related propagation of decoding errors. In design, ISC finds optimal interleaving pattern that separates single video stream into multiple sub-streams based on the channel conditions, packet loss rate and packet correlation ratio.

Simulations have shown that ISC advances traditional method for most of the cases; however, the performance highly depends on the channel's loss rate and the packet correlation. Since most of the packet losses over the best-effort network channel are caused by the buffer overflow of the routers on the path, hence the high frequency, short burst (low packet correlation) losses, or isolated single losses (BEC channel) are more likely to be observed.

In addition to the canonical ISC approach, the proposed ISC-FEC method also has shown performance improvement. Similar to the observation made in previous parts, the most significant quality degrading factor is randomness of packet losses, or strength of memory. However, with the adaptation of FEC, the effect of memory is reduced, and yet provides more flexibility to ISC-MDP for different channel models.

In summary, adopting various models of Interleaved Source Coding to the real-time streaming services over the best-effort IP network would be beneficial.

7.2 Future work

Future extension of ISC framework would be sole R-D optimized FEC protection for *Intra-* coded frame. This will require embedding R-D optimized FEC process into the reward computation process of ISC-MDP framework. With this extension, we expect to find the advantage of *Intra-* coded frame protection algorithm compared with multi-frame protection algorithm of ISC-FEC.

References

[1] Borella, M.S., et al. *Self-similarity of Internet packet delay*. in *IEEE International Conference on Communications (ICC)*. 1997. Montreal, Canada.

[2] Leland, W.E., et al., *On the self-similar nature of Ethernet traffic (extended version)*. IEEE/ACM Trans. Netw., 1994. **2**(1): p. 1-15.

[3] van der Schaar, M. and H. Radha. *Packet-loss resilient Internet video using MPEG-4 fine granularscalability*. in *IEEE International Conference on Image Processing (ICIP)*. 2000. Vancouver, Canada.

[4] Yajnik, M., J. Kurose, and D. Towsley. *Packet loss correlation in the MBone multicast network*. in *IEEE Global Internet Miniconference, part of GLOBECOMM*. 1996. London.

[5] Yajnik, M., S. Moon, and J. Kurose. *Measurement and modeling of the temporal dependence in packet loss*. in *INFOCOM '99. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies*. 1999. New York, NY.

[6] Karande, S., M. Wu, and H. Radha. *Network embedded FEC (NEF) for video multicast in presence of packet loss correlation*. in *IEEE International Conference on Image Processing (ICIP)*. 2005.

[7] Van Der Schaar, M., H. Radha, and C. Dufour. *Scalable MPEG-4 video coding with graceful packet-loss resilience over bandwidth-varying networks*. in *IEEE International Conference on Multimedia and Expo*. 2000. New York, NY.

[8] Lee, H. and H. Lee, *A Packet-Loss Recovery Scheme Based on the Gap Statistics.* LNCS Information Networking, 2005. **3391**: p. 627-34.

[9] Chung, S.-H., et al. *Detection and analysis of packet loss on underutilized enterprise network links.* in *Workshop on End-to-End Monitoring Techniques.* 2005.

[10] Bi, J., Q. Wu, and Z. Li. *Packet delay and packet loss in the Internet.* in *International Symposium on Computers and Communications.* 2002.

[11] Clark, A.D. *Modeling the Effects of Burst Packet Loss and Recency on Subjective Voice Quality.* in *IP Telephony Workshop.* 2001.

[12] Xue, F., V. Markovski, and L. Trajkovic. *Packet loss in video transfers over IP networks.* in *IEEE International Symposium on Circuits and Systems (ISCAS).* 2001. Sydney, Australia.

[13] Feamster, N. and H. Balakrishnan. *Packet Loss Recovery for Streaming Video.* in *Internal Packet Video Workshop.* 2002.

[14] Pyun, J.-Y., et al. *Packet loss resilience for MPEG-4 video stream over the Internet.* in *International Conference on Consumers Electronics (ICCE).* 2002.

[15] Su, Y.-C., C.-S. Yang, and C.-W. Lee, *The analysis of packet loss prediction for Gilbert-model with loss rate uplink.* Inf. Process. Lett., 2004. **90**(3): p. 155-9.

[16] Claypool, M. and Y. Zhu, *Using Interleaving to Ameliorate the Effects of Packet Loss in a Video Stream,* in *International Conference on Distributed Computing Systems.* 2003, IEEE Computer Society.

[17] Liu, Y. and M. Claypool. *Using Redundancy to Repair Video Damaged by Network Data Loss*. in *IS&T/ACM/SPIE Multimedia Computing and Networking (MMCN)*. 2000. San Jose, CA, USA.

[18] Radha, H., M.v.d. Scharr, and Y. Chen, *The MPEG-4 FGS video coding method for multmedia streaming over IP*. IEEE Transaction on Multimedia, 2001. **3**(1): p. 53-68.

[19] Radha, H., et al., *Scalable Internet video using MPEG-4*, in *Signal Processing: Image Communication*. 1999. p. 95-126.

[20] P. de Cuetos, M. Reisslein, and K.W. Ross, *Evaluating the Streaming of FGS--EncodedVideo with Rate--Distortion Traces*. 2003, Institut Eurécom Technical Report.

[21] Lin, S. and Y. Wang. *Error resilience property of multihypothesis motion-compensated prediction*. in *IEEE International Conference on Image Processing (ICIP)*. 2002.

[22] Girod, B., *Efficiency analysis of multihypothesis motion-compensated prediction*. IEEE Transaction on Image Processing, 2000. **9**(2): p. 173-83.

[23] Apostolopoulos, J.G. *Error-Resilient Video Compression Through the Use of Multiple States*. in *International Conference on Image Processing (ICIP)*. 2000. Vancouver, BC, Canada.

[24] Apostolopoulos, J.G. and S.J. Wee. *Unbalanced Multiple Description Video Communication Using Path Diversity*. in *IEEE International Conference on Image Processing (ICIP)*. 2001. Thessaloniki, Greece.

[25] Reibman, A.R., et al. *Multiple description coding for video using motion compensated prediction* in *IEEE International Conference on Image Proc-*

essing (ICIP). 1999. Kobe, Japan.

[26] Barrenechea, G., et al. *Multiple description source coding and diversity routing: A joint source channel coding approach to real-time services over dense networks*. in *International Packet Video Workshop*. 2003.

[27] Begen, A.C., Y. Altunbasak, and O. Ergun. *Multi-path selection for multiple description encoded video streaming*. in *IEEE International Conference on Communications (ICC)*. 2003.

[28] Franchi, N., et al., *Multiple description video coding for scalable and robust transmission over IP*. IEEE Transaction on Circuits and Systems for Video Technology, 2005. **15**(3): p. 321-34.

[29] Khayam, S.A. and H. Radha. *Markov-based Modeling of Wireless Local Area Networks*. in *ACM Mobicom International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*. 2003.

[30] Konrad, A., et al., *A Markov-based channel model algorithm for wireless networks*. Wireless Networks, 2003. **9**(3): p. 189-99.

[31] Levenshtein, V.I. *Efficient Reconstruction at the Output of a Discrete Memoryless Channel*. in *IEEE International Symposium on Information Theory (ISIT)*. 2000. Sorrento, Italy.

[32] Brian A. Banister, B.B., Thomas R. Fischer. *Robust Video Transmission over Binary Symmetric Channels with Packet Erasures*. in *Data Compression Conference, DCC 2002*. 2002.

[33] Adler, M. and F. Fich, *The complexity of end-to-end communication in memoryless networks*, in *Proceedings of the eighteenth annual ACM symposium on Principles of distributed computing*. 1999, ACM: Atlanta,

Georgia, United States.

[34] Gallager, R.G., *Discrete stochastic processes*. 1996, Boston: Kluwer Academic Publishers. xiii, 271 p.

[35] Puterman, M., *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. 1994, New York, NY: John Wiley & Sons, Inc.

[36] Wakamiya, N., et al., *Dynamic Quality Adaptation Mechanisms for TCP-friendly MPEG-4 Video Transfer*, in *International Workshop on Quality of Service in Multiservice IP Networks*. 2003, Springer-Verlag.

[37] Bertesekas, D.P., *Dynamic Programming and Optimal Control*. 1995: Athena Scientific.

[38] *INFORMATION TECHNOLOGY – CODING OF AUDIO-VISUAL OBJECTS – Part 2: Visual*, in *ISO/IEC JTC 1/SC 29/WG 11 , ISO/IEC 14496-2:2001*. 2001: Sydney.

[39] Beynier, A. and A.-I. Mouaddib, *A polynomial algorithm for decentralized Markov decision processes with temporal constraints*, in *International joint conference on Autonomous agents and multiagent systems*. 2005: Netherlands.

[40] Lane, T. and L.P. Kaelbling. *Approaches to macro decompositions of large markov decision process planning problems*. in *SPIE Conference on Mobile Robotics*. 2001. Newton, MA, USA.

[41] *CODING OF MOVING PICTURES AND AUDIO*, in *ISO/IEC JTC 1/SC 29/WG 11 N8706*. 2007: Marrakech, Morocco.

[42] Ebert, J.P. and A. Willig, *A Gilbert-Elliot Bit Error Model and the Efficient*

Use in Packet Level Simulation, in *TKN Technical Reports Series*. 1999, Technical University Berlin: Berlin.

[43] Yee, J.R. and E.J. Weldon, Jr., *Evaluation of the performance of error-correcting codes on a Gilbert Channel.* IEEE Transaction on Multimedia, 1995. **43**(8): p. 2316-23.

[44] Chande, V. and N. Favardin. *Progressive Image Transmission over Compund Packet Erasure Channels*. in *IEEE International Symposium on Information Theory (ISIT)*. 2000. Sorrento, Italy.

[45] Kizhakkemadam, S.N., et al. *Tradeoff between source and channel coding for erasure channels*. in *IEEE International Symposium on Information Theory (ISIT)*. 2005. Adelaide SA, Australia.

[46] Liang, Y.J., J.G. Apostolopoulos, and B. Girod. *Model-based delay-distortion optimization for video streaming using packet interleaving*. in *Asilomar Conference on Signals, Systems and Computers*. 2002.

[47] Chande, V. and N. Farvardin. *Progressive Image Transmission over Compound Packet Erasure channels*. in *IEEE International Symposium on Information Theory (ISIT)*. 2000. Sorrento, Italy.

[48] Yousefi'zadeh, H., H. Jafarkhani, and F. Etemadi, *Distortion-Optimal Transmission of Progressive Images over Channels with Random Bit Errors and Packet Erasures*, in *Proceedings of the Conference on Data Compression*. 2004, IEEE Computer Society.

[49] Wolf, J.K., *ECC Performance of Interleaved RS Codes with Burst Errors.* IEEE Transaction on Magnetics, 1998. **34**(1): p. 75-9.

[50] Su, Y.-C., C.-S. Yang, and C.-W. Lee. *Optimal FEC Assignment for Scal-*

able Video Transmission over Burst Error Channel with Loss Rate Uplink. in *International Packet Video Workshop*. 2003. Nantes, France.

[51] Youshi, X. and Z. Tingting, *Variable shortened-and-punctured Reed-Solomon codes for packet loss protection.* Broadcasting, IEEE Transactions on, 2002. **48**(3): p. 237-45.

[52] Chin, S.K. and R. Braun. *Improving video quality using packet interleaving, randomisationand redundancy*. in *IEEE Conference on Local Computer Networks*. 2001. Tampa, FL, USA.

[53] Matsuzono, K., et al., *Analysis of FEC Function for Real-Time DV Streaming* Sustainable Internet, Lecture Notes in Computer Science, 2007. **4866/2007**: p. 114-22.

[54] Huitema, C., *The case for packet level FEC*, in *TC6 WG6.1/6.4 Fifth International Workshop on Protocols for High-Speed Networks V*. 1997.

[55] Malik, R., A. Munjal, and S. Chakraverty. *Adaptive Forward Error Correction (AFEC) based Streaming using RTSP and RTP*. in *Advanced International Conference on Telecommunications - International Conference on Internet and Web Applications and Services (AICT-ICIW)*. 2006.

[56] Stankovic, V. and Z. Xiong, *Packet Erasure Protection for Multicasting*, in *Conference on Data Compression*. 2004.

[57] Mudugamuwa, U. and B.G. Evans, *Impact of packet-level FEC and existing link level reliability mechanisms to deliver MBMS content over WCDMA mobile satellite systems.* International Journal of Satellite Communications and Networking, 2006. **24**(2): p. 169-84.

[58] Vanhaverbeke, F., et al. *On the Application of FEC to Counter Packet*

Loss Caused by Buffer Overflow. in *28th Symposium on Information Theory in the Benelux*. 2007. Enschede, Netherlands.

[59] Luby, M.G., et al., *Efficient Erasure Correcting Codes.* IEEE Transaction on Information Theory, 2001. **47**(2): p. 569-84.

[60] Nafaa, A., T. Ahmed, and A. Mehaoua. *Unequal and Interleaved FEC Protocol for Robust MPEG-4 Multicasting Over Wireless LANs*. in *IEEE International Conference on Communications*. 2004.

[61] Choi, T.-u., et al., *An Adaptive Periodic FEC Scheme for Internet Video Applications*, in *Proceedings of the Thyrrhenian International Workshop on Digital Communications: Evolutionary Trends of the Internet*. 2001, Springer-Verlag.

[62] Hayasaka, M., M. Gamage, and T. Miki. *Cell Loss Recovery using Adaptive LDPC-based FEC Method for Real Time Services over Best Effort Networks*. in *Joint Conference of the 10th Asia-Pacific Conference on Communications and the 5th International Symposium on Multi-Dimensional Mobile Communications*. 2004.

[63] Hasimoto-Beltran, R. and A.A. Khokhar, *Dynamic Adaptive Forward Error Control for Image Transmission over Lossy Networks*, in *International Conference on Information Technology: Coding and Computing (ITCC)*. 2000.