Journal of Zhejiang University SCIENCE
ISSN 1009-3095
http://www.zju.edu.cn/jzus
E-mail: jzus@zju.edu.cn

H.264/AVC error resilience tools suitable for 3G mobile video services*

LIU Lin (柳 林)†, YE Xiu-zi (叶修梓), ZHANG San-yuan (张三元), ZHANG Yin (张 引)

(*Institute of Computer Science, State Key Lab of CAD&CG, Zhejiang University, Hangzhou 310027, China*)

†E-mail: liulin@zju.edu.cn

Received Jan. 10, 2005; revision accepted May 20, 2005

Abstract: The emergence of third generation mobile system (3G) makes video transmission in wireless environment possible, and the latest 3GPP/3GPP2 standards require 3G terminals support H.264/AVC. Due to high packet loss rate in wireless environment, error resilience for 3G terminals is necessary. Moreover, because of the hardware restrictions, 3G mobile terminals support only part of H.264/AVC error resilience tool. This paper analyzes various error resilience tools and their functions, and presents 2 error resilience strategies for 3G mobile streaming video services and mobile conversational services. Performances of the proposed error resilience strategies were tested using off-line common test conditions. Experiments showed that the proposed error resilience strategies can yield reasonably satisfactory results.

Key words: Error resilience tools, H.264/AVC, 3G mobile video services
doi:10.1631/jzus.2005.AS0041 **Document code:** A **CLC number:** TP37

INTRODUCTION

Video service on 3G systems

The emerging 3G systems support both circuit-switched service and packet-switched service, with video service being an important part of such services. There are three major categories for video services over 3G systems:

(1) Multimedia messaging services (MMS) (3GPP TS 23.140, 2004), also supported by existing old systems such as GPRS.

(2) Packet-switched pre-coded streaming service (3GPP TS 26.233, 2004; 3GPP TS 26.234, 2004), in which video streams are pre-coded and stored.

(3) Circuit-switched (3GPP TS 26.110, 2002) and packet-switched (PCS) (3GPP TR 26.235, 2004) conversational services, for example, video telephony and videoconference.

Terminal requirements for 3G systems

Terminals for 3G circuit-switched systems are called 3G-324M terminals, which are based on ITU-T H.324 (ITU-T, 2002) with Annex C (sometimes referred to as H.324/M), and some modifications defined in (3GPP TS 26.111, 2004) are made on H.324 to be suitable for 3G application. For 3G-324M terminals, the support for H.264 of 3G-324M terminals is optional. When supported, H.264 codecs shall support baseline profile level 1.0 (JVT, 2003). A successful 3G-324M terminal will have to function well at bandwidths as low as 32 kbps and in potentially high error rate environments.

But terminals for 3G packet-switched multimedia terminals (3G PS terminals) (3GPP TR 26.235, 2004) should support H.264/AVC Baseline Profile at Level 1b (JVT, 2004) without requirements on output timing conformance (Annex C of JVT (2003)). It is notable that not all 3G PS terminals support redundant slices in H.264 stream, because of the potential end-to-end delay. So it is better for the streaming server to provide 2 types of H.264 streams for 3G terminals, one with redundant slices, the other with-

*Project supported by the National Natural Science Foundation of China (Nos. 60473106 and 60333010), China Ministry of Education (No. 20030335064), and China Ministry of Science and Technology (No. 2003AA4Z1020)

out redundant slices. Capability exchanges before stream transport will decide which streams to provide for the 3G PS terminals.

Commonly 3G-324M terminals are designed for downward compatibility with old wireless systems, such as GSM and GPRS, and 3G PS terminals are designed for 3G systems. In addition, 3G PS terminals must have the ability for interoperation with 3G-324M terminals: in such case a media gateway functionality supporting 3G-324M–IM Subsystem inter-working should be supplied.

Link performance of 3G systems

For mobile environment, 3G systems have a bandwidth of up to 384 kbps; for stationary environment, 3G systems have a bandwidth of up to 2 Mbps. But 3G mobile systems are error-prone systems, and the packet-loss rate is very high. The packet-loss model of 3G mobile systems is described in (Roth *et al*., 2001). It has also been stated that when packets length increases, the packet-loss probability increases significantly (Stockhammer *et al*., 2003). So in a wireless system, the MTU size is typically considerably smaller, most researchers assume that MTU size is around 100 bytes (Wenger, 2003).

Previous work on mobile video services of H.264/AVC

Two important mobile video services were presented by Stockhammer *et al*.(2003) and Wenger (2003). Stockhammer *et al*.(2003) described in detail the tools that are likely to be used in wireless environment, including R-D optimization, slices structure, error concealment and feedback channel. Wenger (2003) described the application of H.264 coded video over IP networks, using RTP as the real-time transport protocol, and the H.264 error resilience tools and introduced the draft specification of the RTP payload format.

For error resilience tools, there are more literatures. An error recovery scheme based on feedback channel and the use of multiframe is used in conversational services for wireline environment (Yu *et al*., 2004). And an improved error concealment algorithm based on H.264/AVC non-normative decoder is given in Su *et al*.(2003). The performance of various H.264 error concealment techniques is evaluated in (Chiarcluce *et al*., 2004).

All these papers give results of overall research on H.264/AVC over wireline or wireless video services, but when the 3G mobile system is considered, the situation becomes different, because more detailed requirements are described for 3G mobile terminals in 3GPP/3GPP2. This paper focuses on the selection of H.264/AVC error resilience tools suitable for 3G mobile video services, and detailed experiment results are given for comparison.

H.264 ERROR RESILIENCE TOOLS FOR MOBILE VIDEO SERVICES

H.264/AVC provides many error resilience tools, but some of them are not suitable for mobile video services. In this section, the error resilience tools suitable for mobile video services are discussed.

Slice structure

The slice structure is first used in MPEG-2 for error resilience, and is adopted by the standards followed. Due to the restriction of MTU size (Wenger, 2003), slice structure for the H.264 video encoder in 3G mobile services is more important. The encoder has to fragment the payloads to accord with the MTU demand. In H.264, users can choose different slice structure according to their need. This is called Flexible Macro-block Ordering (FMO) technology.

Error concealment

Error concealment utilizes the received data to recover lost data, and commonly the spatial information and temporal information are used in the error recovery procedure. Error concealment is important in error-prone environment, especially for 3G mobile video decoders.

JVT reference software utilizes an error concealment method based on spatial information where optimal SAD difference exists between edge pixels of lost area and neighbouring edges pixels of received blocks.

There are also many other error concealment methods, but all of them are based on spatial information or temporal information.

Intra block refreshing

H.264/AVC uses intra block refreshing to stop error propagation and drift due to the macroblock loss. H.264/AVC uses intelligent intra-block refreshing by R-D (rate-distortion) control, as described in (Wenger, 2003). For real-time and conversational video services, it is not advisable to insert I-frames due to bit-rate constraint and the resulting long delay, so, intra block refreshing is very important for removing artifacts caused by error and inter-prediction drift.

Feedback channel

Feedback channel can be efficiently used with long-term memory motion-compensated prediction for error resilience (Wiegand *et al*., 2003); and can also be used for packets retransmission. The receiver can send ACK messages or NAK messages to the sender to report whether the RTP packets have been received or not. In the case of conversational applications, the encoder can select the corrected-received frames as the reference frames, and in the case of stream applications, the sender can retransmit the lost packets.

Some efficient error resilience tools, such as data partition, are not supported by H.264/AVC baseline profile, so they cannot be used in mobile video services. The encoder or decoder impliedly provides other error resilience tools such as parameter sets and error diction, but they are not listed here.

PROPOSED ERROR RESILIENCE STRATEGIES FOR 3G MOBILE SERVICES

Error resilience strategies for MMS

Due to the feature of MMS, feedback channel of the terminals to the server is valid and retransmission of the lost packet is supported by 3G systems. So the encoder server can focus on the compression efficiency, and error resilience is not the most important part of the H.264 video systems. The MTU size for MMS application can be properly increased, but according to Stockhammer *et al.*(2003), the MTU size should not exceed 200 bytes due to the high packet-loss rate. Because MMS service is relatively simple, in this work, MMS service was not tested by experiments.

Error resilience strategies for streaming service

When streaming service is provided by 3G systems, in most cases the video is pre-coded in the server, and several seconds of delay is tolerable. So the mechanism of feedback between the server and the terminals can only provide the ability of packets retransmission, the encoder with multiple frame reference is not suitable. For streaming service, slice structure feedback channel is required; data partition and error concealment are optional; and intra block refreshing is optional because the decoder can ensure that it correctly received all the packets.

Error resilience strategies for conversational service

Conversational service on wireless environment is a challenging task, because mobile terminals must have the real-time ability of both encoding and decoding. Commonly the computing ability and memory capability of mobile terminals is restricted by their hardware, so we must consider the computing complexity and memory requirement for error resilience algorithm.

Error concealment is the most important tool for conversational service, because feedback channel and RTP packets retransmission are not allowed due to the implicit time delay. Other tools that can help error concealment are also important. Here slice structure with FMO is used. It is better that macro-blocks are sliced by interleave mode, because this can make full use of spatial correlation and reduce the bit rate requirement. Moreover, intra block refreshing should be supported to remove the error drift caused by the loss of inter-prediction macro-block. Redundant slice will not be supported, because it entails increased bit rate.

EXPERIMENTS RESULTS

Here we consider only streaming services and conversational services. The common off-line test conditions are described in Table 1. The proposed error resilience strategies for mobile video services are listed in Table 2. The first item without error resilience tool is used for comparison. The error concealment tool is used by all 3 services. Streaming service uses only error resilience tool of feedback channel. Conversational service uses most error resilience tools.

The encoded sequences are IPPP… sequences (I frame only at the first frame), and with the bit rate control of 64 kbps, and with the initial *QP*=24, All the sequences are encoded by H.264/AVC baseline profile level 1b, all the sequences are encoded and decoded with JM software version JM 9.2, but we fix the decoder bugs caused by error concealment modules. Packets loss is generated by off-line simulation software in Roth *et al*.(2001). When a frame cannot be recovered due to RTP packets loss, direct frame copy from previous frame was performed. The experiment results were evaluated by average PSNR of luminance Y value representing the whole effect, and the minimum PSNR of luminance Y value representing the worst conditions.

Other important encoding/decoding parameters includes:

For the encoder, ProfileIDC=66, LevelIDC=10, constrained_set1_flag=1. For the decoder, the MAX_NUM_SLICES should be given relatively bigger value (such as 200), because there will be much more slices for I frames.

Experiment results for mobile streaming video services are listed in Table 3 and Table 4. Table 3 lists mobile streaming services without feedback channel and shows that mobile streaming services without efficient feedback channels will bring intolerable RTP packets losses. Even error concealment and frame-copy cannot stop the error drift; and the received RTP packets cannot be correctly decoded, so we cannot give the Y PSNR value. Table 4 lists mobile streaming services with a feedback channel, and every lost RTP packet being retransmitted once. Table 4 data show that feedback channel can efficiently remove the RTP packets loss, so that the PSNR value is satisfied.

The detailed experiment results for mobile con-

Table 1 Common test conditions suggested by Roth *et al*.(2001)

No.	Bit rate (kbit/s)	Length (s)	Bit error pattern	Applications
1	64	60	9.3e-3	Streaming
2	64	60	2.9e-3	Streaming
3	64	180	5.1e-4	Conversational
4	64	180	1.7e-4	Conversational
5	128	180	5.0e-4	Conversational
6	128	180	2.0e-4	Conversational

Table 2 Error resilience strategies for mobile video services

	Without error resilience	Streaming service	Conversational service
Slice mode	Fixed 100 bytes in slice	Fixed 100 bytes in slice	Fixed 100 bytes in slice
Error concealment	Yes	Yes	Yes
Redundant slices	No	No	No
FMO	No	No	Interleave mode
MB line intra refreshing	No	No	Yes
Feedback channel	No	Yes	No

Table 3 Experiment results of mobile streaming video services without error resilience

Video sequences	Video format	Frame rate (fps)	Total frames	RTP packets	Bit error pattern	Lost packets	Ave. Y PSNR	Min. Y PSNR
Foreman	QCIF	15	250	1504	1 without retransmission	179	N/A	N/A
Foreman	SQCIF	30	250	793	1 without retransmission	81	N/A	N/A

Note: Bit error pattern of 1 without retransmission means the bit error pattern with the retransmission mode set by 0

Table 4 Experiment results of mobile streaming video services with error resilience strategy

Video sequences	Video format	Frame rate (fps)	Total frames	RTP packets	Bit error pattern	Lost packets	Ave. Y PSNR	Min. Y PSNR
Foreman	QCIF	15	250	1504	1	0	35.54	33.14
Foreman	SQCIF	30	250	793	1	0	36.33	33.21

versational video services are given in Fig.1, and the test conditions and overall result are listed in Table 5. Fig.1a gives the comparison result of Carphone sequences, and Fig.1b gives the comparison result of Foreman sequences. In Fig.1, the upper curve is the PSNR value of decoded sequences with proposed error resilience strategy, and the lower curve is the PSNR value of decoded sequences without error resilience. Appropriate error resilience strategy can help achieve better video quality. Without appropriate error resilience, when a RTP packet is lost, the PSNR value falls rapidly and has a permanent adverse effect on the following frames. Our proposed system with appropriate error resilience strategy can prevent falling of PSNR value and efficiently remove error drift to obtain satisfactory PSNR value. Other error resilience tools may also be used for this case, but the computation loads for mobile terminals have to be considered as well.

Table 5 Experiment results of mobile conversational video services with/without error resilience strategy

Video sequences	Video format	Frame rate (fps)	Total frames	With error resilience	RTP packets	Bit error pattern	Lost packets	Ave. Y PSNR	Min. Y PSNR
Carphone	QCIF	15	380	Yes	2526	3	31	32.12	25.01
Carphone	QCIF	15	380	No	2325	3	29	30.75	22.53
Foreman	QCIF	15	380	Yes	2511	3	29	31.18	23.41
Foreman	QCIF	15	380	No	2284	3	31	27.63	21.22

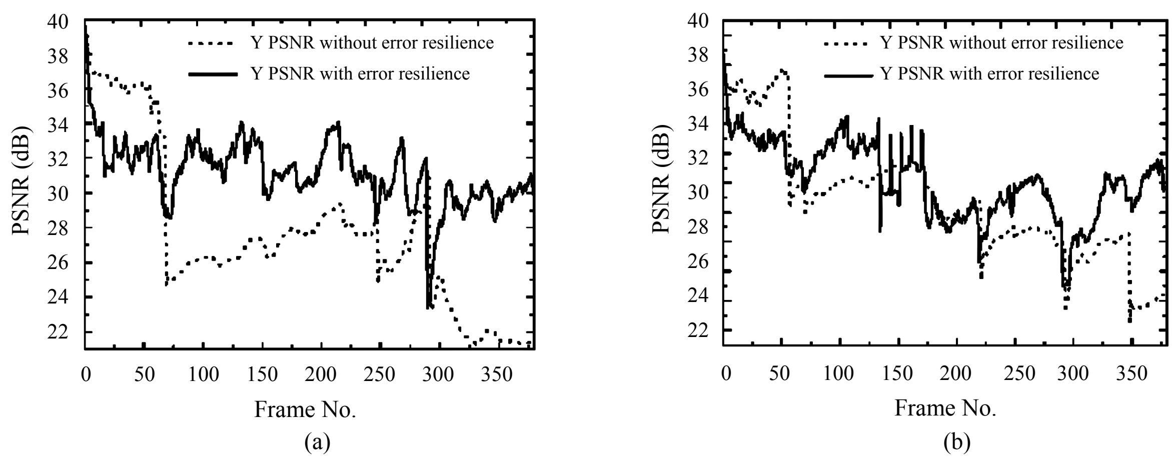

Fig.1 Experiment results of Carphone sequences (a) and Foreman sequences (b) (380 frames)

CONCLUSION

The emergence of 3G technology makes video transmission in wireless environment possible. The latest 3GPP/SGPP2 standards require that the 3G mobile terminals should support H.264/AVC. The function of 3G mobile terminals is restricted by their hardware, and 3G mobile terminals support only parts of the error resilience tools of H.264/AVC. This paper analyses suitable error resilience tools for 3G mobile video services and present 2 error resilience strategies for streaming video services and conversational services. Experiments showed that the proposed error resilience strategies can avoid most RTP packets loss in streaming service, and can stop error drift due to packet loss in conversational service. Compared with 3G mobile services without error resilience, the PSNR value is reasonably improved.

References

3GPP TS 26.110, 2002. Codec for Circuit Switched Multimedia Telephony Service, General Description. V5.0.0.
3GPP TS 23.140, 2004. Multimedia Messaging Service (MMS), Functional Description, Stage 2. V6.8.0.
3GPP TS 26.111, 2004. Codec for Circuit Switched Multimedia Telephony Service, Modifications to H.324.

V6.0.0.
3GPP TS 26.233, 2004. Transparent End-to-end Packet Switched Streaming Service (PSS), General Description. V6.0.0.
3GPP TS 26.234, 2004. Transparent End-to-end Packet-switched Streaming Service (PSS), Protocols and Codecs. V6.1.0.
3GPP TR 26.235, 2004. Packet Switched Conversational Multimedia Applications, Default Codecs. V6.2.0.
Chiaraluce, F., Ciccarelli, L., Gambi, E., Spinsante, S., 2004. Performance Evaluation of Error Concealment Techniques in H.264 Video Coding. Picture Coding Symposium 2004.
ITU-T (ITU-T Recommendation H.324), 2002. Terminal for Low Bitrate Multimedia Communication.
JVT (Joint Video Team) of ISO/IEC and ITU-T VCEG, 2003. Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC), JVT-G050r1.
JVT (Joint Video Team) of ISO/IEC and ITU-T VCEG, 2004. ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, Draft Text of H.264/AVC Fidelity Range Extensions Amendment, JVT-L047.
Roth, G., Sjoberg, R., Liebl, G., Stockhammer, T., Varsa, V., Karczewicz, M., 2001. Common Test Conditions for RTP/IP over 3GPP/3GPP2, ITU-T SG16 Doc, VCEG-M77.
Stockhammer, T., Hannuksela, M.M., Wiegand, T., 2003. H.264/AVC in wireless environments. *IEEE Transactions on Circuits and Systems for Video Technology*, **13**(7):657-673.
Su, L., Zhang, Y., Gao, W., Lu, Y., 2004. Improved Error Concealment Algorithms Based on H.264/AVC non-normative Decoder. Proceedings of the 2004 IEEE International Conference on Multimedia and Expo (ICME'2004).
Wenger, S., 2003. H.264/AVC over IP. *IEEE Transactions on Circuits and Systems for Video Technology*, **13**(7):645-656.
Wiegand, T., Farber, N., Stuhlmuller, K., Girod, B., 2003. Error-resilient video transmission using long-term memory motion-compensated prediction. *IEEE J. Select. Areas Commun.*, **18**:1050-1056.
Yu, H.B., Wang C., Yu, S., 2004. A novel error recovery scheme for H.264 video and its application in conversational services. *IEEE Transaction on Comsumer Electronics*, **50**(1):329-334.