Key Technology Research and Analysis of H.264/AVC Suitable for Wireless Network

YUN-PENG LIU\textsuperscript{1,2}, SAN-YUAN ZHANG\textsuperscript{1}, SHU-CHANG XU\textsuperscript{1}, YIN ZHANG\textsuperscript{1}

\textsuperscript{1}College of Computer Science and Technology  
Zhejiang University  
388 Yuhangtang Road, Hangzhou  
CHINA  
\textsuperscript{2}College of Computer Science and Information Technology  
Zhejiang Wanli University  
8 SouthQianhu Road, Ningbo  
CHINA

Abstract: - The emergence of H.264/AVC video compression standard provides better support functions for wireless network. The characteristics of limited bandwidth, high error rate of wireless channel and the constraints of computing and storage capacity of most mobile devices make the key technology of video codec in error resilience strategies, rate control, mode decision and scalable extension more important. In this paper, we first analyze the main and popular strategies in the key technology, and then we propose a method to choose and combine these strategies to validate the results and performance by experiment. Experiment shows the rational use of these technologies improves the quality of video transmission and provides reference for the video service evaluation.

Key-Words: - H.264, Wireless network, Video transmission, Error resilience, Rate control, Mode decision

1 Introduction

As the next-generation video coding standard, compared to the previous coding standard, H.264/AVC codec performance has a great increase. Higher codec efficiency and stronger ability to adapt to network and to error resilience make it more suitable for wireless transmission. As wireless communications technology development, especially the emergence of the 3G communication system makes the video communication gradually becoming the major service and important part of the mobile telecommunication, and the services are more diversified, such as: wireless video broadcasting, wireless video dialogue, wireless VOD and wireless video surveillance etc.

2 Problem Formulation

Rate problem: In order to make the ever-changing video bit stream fit to the relatively fixed bandwidth requirements, so rate control is an essential and indispensable part in different video compression standards. Traditional rate control method is to use of the buffer to adjust the video output bit stream, use of the time delay to make transmission channels adapt to the natural variable rate of compressed video bit stream. Now commonly used by the traditional rate control algorithms are VBR and CBR. Different form the traditional stable cable channel, wireless channel has a much higher bit error rate, and the bit error rate will soar by the increase of transmission packet [1]. So the packet size is restricted in the wireless channel. In general, the Maximum Transfer Unit (MTU) is about 1500 bytes in size in cable channel, and not more than 254 bytes in wireless channel [2]. This problem is not very obvious in cable channel because of the bigger IP packet, so the rate control strategies are very different in the two kinds of channels [3].

Complexity problem: One key reason why H.264/AVC can get a great increase in encoding performance is that a variety of different coding modes in motion estimation (ME) are introduced. The currently adopted encoding mode of H.264/AVC is that encoder must check up the encoding mode in every macro block so as to determine the optimal
mode. The full search motion estimation applying the all models is estimated to have more than 80% of the whole calculation. This whole search method makes it difficult to attain real-time communication requirements \cite{4}.

Scalability problem: Also, the wireless network environment, which is made of various types of wireless network, is heterogeneous, such as 3G/GPRS, WLAN, WPAN, wireless Ad Hoc network etc. How to fully utilize the bandwidth and resources of the existing wireless network, to provide the different video services based on the different demand and process capacity, to realize the service flexibility and ensure the service quality will become an import issue \cite{5}.

For the problems mentioned above, there are all kinds of new research techniques and results to solve one side in a certain degree. However, it is few to combine all these improved methods for an experiment. In this paper, an innovation is that we choose and combine the advanced methods to test and analyze by 3G offline simulation environment and JVT reference software, and it provides references for evaluation of video service.

In the following description, the structure of H.264/AVC video streaming transmission in the wireless network will be introduced first, then the key technology, the related discussion and solution will be presented, and the experiment is tested and analyzed, finally, we give a conclusion and refer the further research work.

3 Problem Solution

3.1 Structure of H.264/AVC video streaming transmission

As shown in Fig. 1, H.264/AVC can be divided into two different semantic levels: VCL and NAL. VCL is responsible for the effective video signal compression, and NAL gives a definition of the interface between video compression information and the outside. And NAL is responsible for packaging each slice of VCL into a NAL unit. In the process of compression the synchronization and error identification are processed. After NAL decoder receives NAL units, it will check the synchronization and error identification to ensure the correct decoding.

![Fig. 1 H.264/AVC Semantic Hierarchical Structure](image1)

In order to adapt different types of transmission systems, NAL unit must be re-packaged in accordance with different transport protocols in the transport layer. When packets switching, NAL units usually are packaged and transferred in the form of RTP package. As the network of “All IP” trend, most multimedia application will apply in the form of packets switching in the next wireless network.

The data of H.264/AVC NAL are packaged RTP protocol packets, and then transfer in IP network to CDMA2000 system as an example for real-time packet data service shown as Fig. 2. In order to reduce the relay of retransmission, UDP protocol is adopted.

![Fig. 2 NAL Unit Package in CDMA2000 System](image2)

3.2 Error resilience strategies

There are two reasons on errors in wireless network, one is the bits inversion, and the other is that RTP packets are discarded by the gateway or router. The latter data cannot recover, but the former data can recover by means of FEC etc in the applications which need not a high real time. Considering the delay issue in the real-time application, RTP packages are discarded directly \cite{1}. In this way, the
main errors are caused by RTP packages loss and the problem of synchronization loss.

Judging from the current view of the latest study, the use of forward error correction (FEC) measures such as convolutional code or Turbo code can effectively reduce the bit error rate in wireless channel transmission. But too complex channel coding methods increase the complexity and cost of user terminals. 3GPP Standardization Organization recently selects Digital Fountain's FEC technology as the mandatory component of 3G multimedia broadcasting and multicast transmission business MBMS.

Reference [7] refers a video error resilience transmission strategy in combination with the typical characteristics of 3G wireless channel and based on data division redundancy package and FMO. The method can adjust the length of redundancy package to maximize the saving rate and protect the important data when the channel error rate changes. Reference [8] refers a method based on the application of data partitioning and unequal error protection. The influence factor of inter MB and γ curve are introduced. Reference [9] puts forward a RTP packetization algorithm for H. 264/AVC video which is called Hybrid Mode Packetization (HMP). HMP considers video stream content correlation and unequal protection of important information with H. 264/AVC NAL video features.

Reference [10] refers two improved methods: the first method changes the burst error in the network into statistically independent error, using convolution and interlaced technique on the packet encode; the second method divides video data into important data and unimportant data and checks the error in important data at receiving terminal, to control accordingly the retransmission mechanism at transmit side in the wireless transmission system.

Reference [11] first presents that mobile terminals support only part of error resilience tools of H. 264 /AVC because of the hardware restrictions. Then presents various error resilience tools and their usability in wireless environment, at the same time, some simple error resilience schemes for packet switched conversational services are proposed. Finally experiments data show that encoder with low complexity FMO method and appropriate intra block refreshing can get better error correction ability. Reference [3] refers taking full advantage of the results of feedback coming from the reverse channel when bidirectional video transmission. Encoder can accordingly change the reference frames based on the result of feedback under the circumstances of multi-frame reference in the coder end so as to enhance the ability of error resilience.

3.3 Rate control
Different from MPEG-2, MPEG-4 and H.263 encoding system, the rate control algorithm of H.264/AVC standards is much more complex. This is because the quantitative parameter QP attained from rate control not only processes rate control, but also uses as the basis of selecting macro block code mode in rate distortion optimization model (RDO). In order to resolve this contradiction that can not be avoided, H.264/AVC reference software adopts a number of methods based on the prediction to ensure the bits that each frame will take up and the MAD value, thus the quantitative parameter QP can be ensured.

The rate control algorithm of H.264/AVC has used HRD buffer reference model. MAD prediction model, second rate-distortion model theories etc, these are important to the prediction of bit rate parameters in each frame.

EBR rate control algorithm aims at the wireless channel in particular [12, 13]. And EBR is the abbreviation of Explicit Bit Rate, that is, the bit number after encoding is made certain in advance by encoder according to the rated channel bandwidth, the output of encoder can not exceed the scope so as to meet the wireless network transmission need better. EBR rate control algorithm is designed specifically for the wireless channel, EBR rate control algorithm has the following advantages:

Lower end-to-end delay: as EBR adopts the idea of Source adapting to channel, each frame must complete transfer in the given time, thus, not only the buffer size of decoder can effectively be reduced, but also the end-to-end delay can be reduced too.

Improve the system's error resilience: compared with VBR methods, EBR strictly controls the bit number of each encoding frame, minimizing the number of PDU in a given bandwidth so as to minimize the lost number. It can effectively improve error resilience in the wireless channel environment.

Improve compression efficiency: as EBR can minimize the number of PDU and RTP package in the transmission, the additional expenses in the unit head reduce to the least too. So the compression efficiency improves.

The key to EBR rate control algorithm is the realization of slicing. JVT reference software provides two methods: one is fixed macro blocks, the number of macro block in each slice is certain, as the number is specified by the user in encoding, the number of slice is certain after each frame is compressed. The other method is to fix the bytes of each slice, if the cumulative number of bytes when encoding exceeds a given value, a new slice is created.
Reference [3] refers a dynamically self-adaptive slicing strategy to realize EBR coding. It first makes use of a feedback coding method to determine the reasonable slicing position in the first I frame and P frame of GOP, and then regards the result of slicing as the reference of remaining P frame in GOP. This slicing method can efficiently improve the utilization rate of channel resources and enhance the quality of the video encoding in relation to the fixed slicing method, and ensure the encoding efficiency at the same time.

### 3.4 Mode decision

In recent years, many fast algorithms on mode decision are present. In reference [14], macro block is classified based on the price of rate-distortion, and each mode type is tested according to a certain sequence, if the coding price of a certain mode is less than the threshold, the test to other remaining mode will stop. This method does not consider the correlation between the adjacent macro block modes.

In reference [15], the encoded macro blocks modes of the current frame and the former frame are used to predict the current macro block mode, the adjacent macro block mode which costs the smallest price is chosen as the current. This method only chooses one macro block of the former frame among the time-adjacent macro blocks, in which case the prediction mode is not always optimal, which may reduce the video encoding quality for the wrong mode prediction. In reference [4], the high complex problem of mode decision and the algorithms in reference [14] and [15] are analyzed, a fast mode decision algorithms is present. The concept of mode reference aggregate decision algorithm is proposed based on region forecast. This method can significantly reduce the number of candidate mode, thus simplifying the procedure of mode decision.

### 3.5 Scalable extension

There are tow ways to realize the scalable extension of video encoding: one is the use of the scalable coding methods, such as embedded bit-plane coding; the other is the use of layered coding. The SVC extension of H.264/AVC combines layered coding and fine coding, realizes the complete flexibility of the space, the time and SNR.

Spatial scalability is realized through up/down sampling filter changing the spatial resolution. Temporal scalability is realized through the internal temporal scalability in the coding structure of motion compensation and motion prediction. SNR scalability is realized by use of the embedded bit-plane fine coding.

The scalable stream of video transmission has another advantage, that is, different level is given the different protection strategy. The basic level can adopt the stronger protection strategy, such as: FEC and ARQ, a more flexible strategy to the enhanced level, such as: weak FEC or discarding directly.

![Fig.3 PSNR Comparison Results of Wireless Error-prone Channel](image-url)
3.6 Experiment and analysis
Not adopting FMO plus row intra-coded mode is chosen as error resilience strategy, rate control adopts self-adaptive EBR algorithm, mode decision adopts the concept of mode reference aggregate. Scalable coding is not considered. JVT H.264/AVC reference software JM10.2 is used as encoding and decoding software, the simulated RTP/UDP/IP packet Packaging and wireless channel transmission are tested through VCEG-N80 reference software (MIP software). Test sequence is the foreman QCIF sequence sampling by 10frames/second, the transmission result is observed through the 2% error channel. The result is shown as Fig.3 and Table 1.

4 Conclusion
In this paper, after the study of the existing key strategies and technologies of H.264/AVC in the wireless channel application, a method of combination to compare and verify is proposed. Experimental results show that the method significantly reduces the rate and ensures the encoding quality at the same time in the higher error rate wireless channel, and the decoded average PSNR improves greatly. It is very valuable and meaningful for video services to apply in the wireless channel.

References: