A Whole Rate-Distortion Optimize Pronounce Based on Wireless Sensor Network

Qingsheng YU [1]
Foshan Polytechnic, Foshan, Guangdong,528137, China
fsgarden@163.com

Jian ZHANG [2]， *Jinxiang PENG* [2]
Guangdong Vocational Technology Institute
Foshan, Guangdong, 528041,China
zj3319896@163.com

***Abstract*: Based on the Joint Video Team (JVT) of the ITU-T Video Coding Experts Group VCEG and the ISO/IEC Moving Picture Experts Group MPEG, an RD optimal Macro Block mode decision scheme for Internet error channel streaming is introduced. The scheme employs the luminance Rate Distortion (RD) optimal mode decision scheme so as to take the effects of video encoding distortion and the channel error propagation to get higher error robustness for error transmission. Based on the Wireless Sensor Network, this paper analyzes the data distortion problem when transmitting H.264 coded video stream over error-prone channel. And the authors also have discussed a widely accepted technique that introduces more intra-coded information on macro block basis. Additionally, this paper introduces a simple loss and multiplication factor estimation method, the rate-distortion optimized assessing strategy over the whole situation.**

Keywords*: *statistics, global rate distortion optimality, H.26L video compression, tagrangian

I. INTRODUCTION

H.264 has been accepted as the unified video coding standard by the new generation of broadcast, communication and storage media (CD DVD), which uses simple effective blocks to acquire relatively high coding efficiency greatly superior to the current standards. On the basis of mature techniques such as block transformation, motion estimation/compensation, quantified entropy coding, etc. employed by current video standards, H.264 also uses multiple new techniques, such as the unified VLC symbol coding, the high precision/multi-mode displacement estimation, the integral transformation based on 4*4 blocks, the hierarchical coding grammar, etc. that increase the error resilience ability and ensure effective coding. The highly efficiency of video coding and solid network adaptability of H.264 enable it to become the most possible standard for broadband interaction, network video communication and new wireless video communication media compression. Because H.264 is efficient coding, it has relatively high requirement on systems like video terminals, gatekeepers, gateways, MCU, etc. applied in video bitstreams in different channels. It is very sensitive to the channel error rate for even a single primary error may cause dramatic decrease of video quality after recovery. In addition, actual wireless and IP channels may also cause error codes. Therefore, it is necessary to use a video error resilience method in the H.264 standard to ensure the quality of images.

II. VIDEO TRANSMISSION MODELS

Video streams may cause channel distortion, including channel errors and packet loss, as illustrated in Fig I. It is a video transmission model under a network video communication error channel, in which Model Yn is the encoding end quantified distortion introduced to Xn, the error concealment distortion introduced due to channel distortion and the error propagation distortion introduced due to reference frame errors.

The decoding end reconstruction value Yn is a random variable based on Xn. When the video frames received by the decoder end have errors, the decoder will apply the error concealment technique to reconstruct the error part, but there is certainly video information loss in the reconstructed frames compared with the original frames at the encoding end. If the lost information part is used as reference frame information of motion compensation and estimation, it will cause error propagation that not only decreases the quality of recovered images of the error frames but also brings unrecoverable loss to following frames.

This can ensure fast recovery of the reconstructed video frames from errors at the decoding end and the encoding end can retransmit complete I frame data, but transmission of excessive I frame data will swiftly increase the instant data volume transmitted in channels and is very easy to cause network obstruction. For this, by full use of Intra coding macroblocks, error codes can be prevented from inter-frame propagation and the decoder must be aware of the positions and ranges of the distorted segments of reconstructed frames at the decoder end. As illustrated in Table I, four methods of macroblock refresh strategies are given.

III. MACROBLOCK REFRESH METHODS

In the H.264 standard technology, the Intra macroblock refresh method is provided, i.e., recoding the error macroblocks at the encoding end in the Intra model, and the encoding end only sends Intra coding macroblock data.

It can be known from the analysis in Table I that on the basis of mathematical statistics theory, a simple loss and multiplication factor estimation method, the rate-distortion optimized assessing strategy over the whole situation, is proposed.

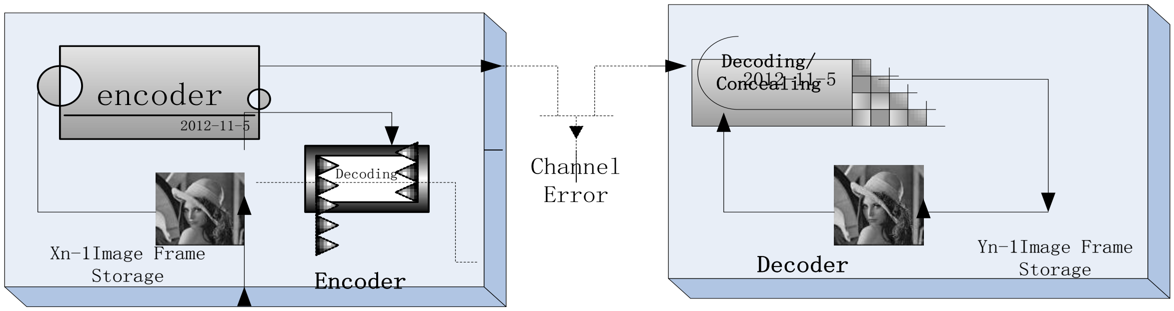

Fig. I Video Transmission Model Nnder Error Channel

Table Ⅰ

Method	SEI Domain Feedback Mechanism	Random Refresh Strategy	Macroblock Refresh Strategy	Rate-distortion Optimized Assessment Strategy
Features	The decoder is very easy to report the positions and ranges of error macroblocks to the encoder. The decoder at the encoding end can reconstruct the current frame based on such feedback information, assess the error macroblocks and conduct Intra coding to these macroblocks, thus terminating error propagation. However, The time delay of channels in real transmission and the error propagation within this period should be considered if using this method. Its efficiency is not satisfactory and it can only be used where very few transmission errors exist.	It randomly selects an appointed number of macroblocks to have Intra coding and conduct macroblock refresh to the decoding end by real-time transmission. This method is simple and easy, but due to the refreshed macroblocks being randomly selected, some error ones may not be refreshed for a long time. Therefore its effect is not satisfactory.	One strategy is only refreshing an appointed number of macroblocks once and finally refreshing all macroblocks consecutively. For instance, it first refreshes Macroblocks 1 to n, later n+1 to n+n and so on, until all macroblocks are refreshed. The other strategy is taking into account the content of images during Intra coding, which has higher refresh frequency to the macroblocks containing motion vectors or motion compensation information, but ceases refreshing by use of Intra macroblocks the static segments, such as the background of images, in those images.	Targeting at those situations with relatively high rate-distortion requirement, when considering coding distortion, the error concealment distortion and error propagation distortion caused by effective estimation of channel distortion should also be considered. It can take care of both the coding efficiency of video streams and the error recovery ability in the selection of macroblock coding models.

IV. ALGORITHM OF RATE-DISTORTION MODE ASSESSING OVER THE WHOLE SITUATION

H.264, as a new generation video coding standard, applies the rate-distortion assessing strategy to motion vectors, reference frame selection, intra-frame estimation mode judgment and macroblock coding mode judgment in a highly complicated mode. The decoder defines the macroblock coding mode corresponding to the minimum cost function to be the rate-distortion optimized macroblock coding mode through computing the rate-distortion assessing functional expression.

Definition 1: Assume the reference frame set is R (Reference), only an I frame is transmitted in the serial video frames at the beginning, and other video frames in this series are all P frames. If Mp is the macroblock coding model set of the prediction codes related to the time-space domain and MI is the macroblock coding model set of any

prediction code, the minimum cost functional expression of Macroblock m of Frame n in this series should be:

$$OPTION_{n,m}=\mathrm{Min}\left(D_{n,m}(MODE)+\lambda R_{n,m}(MODE)\right) \quad (1)$$

In the above term, Dn, m (MODE) is the coding distortion of Frame n Macroblock m, i.e., error

concealment distortion introduced when the channel is distorted and the error propagation distortion brought by correct receiving of motion prediction; λ is the Lagrangian multiplication factor; Rn, m (MODE) is the macroblock coding bit rate; MODE is the coding mode of the current macroblock.

It can be learnt from Formula (1) that to the specific coding mode MODE, the encoding end is very easy to obtain the macroblock coding bit rate Rn, m (MODE). Then in the error channel transmission environment, how Dn, m (MODE) and the multiplication factor can be computed?

A. Estimation Algorithm of Dn,m(MODE) Based on Law of Large Numbers

Definition 2: Assume the original value of Pixel i in Frame n Macroblock m is Vn, m, i, and K is the channel error information feedback from the real-time storage SEI domain, then according to such channel error information and the original pixel values, in the decoder at the encoding end, the pixel value V'n, m, i (i) and the statistical value V'n, m, i after K errors can be reconstructed. According to the law of large numbers, if K pieces of error information are independently and identically distributed, when K is large enough, it will be:

$$\frac{1}{K}\sum_{k=1}^{K}\left|V_{n,m,i}-V'_{n,m,i}(k)\right|^2=E\left|V_{n,m,i}-V'_{n,m,i}\right|^2=d_{n,m,i} \quad (2)$$

In the above term, dn, m, i is the mathematical expectation of the distortion statistical value that represents the distortion statistical value of Pixel i. Here, in a situation where the requirement on precision is relatively high, the value of K can be larger, but the computing at the encoding end will be comparatively higher. Nevertheless, when the value of K is relatively small, this statistical method also has good convergence and can achieve a fairly good result.

It can be learnt from Formula (2) that if following this estimation method, all distortion statistical values of the pixels within the macroblock can be known at the encoding end to obtain all pixel distortions within it and compute the macroblock distortion.

Definition 3: Assume there are T pixels in a macroblock, then the sum of the three factors in Dn, m (MODE) distortion should be:

$$D_{n,m}(MODE)=\frac{1}{T}\sum_{i=1}^{T} d_{n,m,i}(MODE) \quad (3)$$

In the above term, dn, m, i (MODE) is the distortion statistical value when the macroblock coding mode of Frame n Macroblock m Pixel i.

It can be learnt from Formula (3) that based on statistics and the estimation algorithm of the law of large numbers, if the distortion of each pixel in the macroblock is known, through averaging all pixel distortions, the macroblock distortion can be easily computed.

B. Computing of Lagrangian Multiplication Factor $\boldsymbol{\lambda}$

It can be learnt from the discussion on Formula (1) that the macroblock coding bit rate R is relevant to the macroblock coding mode, for D (MODE) is a function of the macroblock coding mode MODE, and D (MODE) is also a function of R, recorded as D(R). It can be learnt according to the H.264 coding and decoding standard paper [1] and Formula (1) that the computing method of the multiplication factor $\boldsymbol{\lambda}$ that meets the limits under the Lagrangian condition should be:

$$\lambda=-\frac{dD(R)}{dR} \quad (4)$$

Namely the optimal $\boldsymbol{\lambda}$ is a negative derivative of the distortion function D(R) and the macroblock coding bit rate R.

According to the above discussion, D (MODE) is the sum of the three factors of error propagation distortion brought by reference frame errors. Here, the coding distortion under an error-free environment is set to be DS (Source); the error concealment distortion at macroblock loss is set to be DC (Concealment); the error propagation distortion is set to be DP (Propagation); the probability of channel distortion (including packet loss and error introduction) is set to be p; the probability of errors in the reference frame motion prediction part is set to be q, then the total D(R) can be expressed as:

$$D(R)=(1-p)(1-q)D_S+pD_C+(1-p)qD_P \quad (5)$$

The macroblock coding distortion DS is relevant to the macroblock coding mode under the error-free environment, a function of the coding bit rate R; DC is the distortion caused by channel errors, not a function of R; DP is mainly affected by channel distortion, hardly affected by coding distortion and can be omitted. Then from Formulas (4) and (5), it should be:

$$\lambda=-\frac{dD(R)}{dR}=-(1-p)(1-q)\frac{dD_S}{dR}=(1-p)(1-q)\lambda_0 \quad (6)$$

In the above term, $\lambda_0=-\frac{dD_S}{dR}$ indicates the value of $\boldsymbol{\lambda}$ taken under the error-free environment. The value under the error-free environment in the H.264 draft is $\lambda_0=0.85*2^{\frac{Q}{3}}$, where Q is the macroblock quantification parameter.

In Formula (6), the values of p and q are similarly obtained by use of the statistical method and the approximation to the maximum entropy principle based on the feedback information from the decoding end.

C. Approximation to the Maximum Entropy Principle

From the perspective of channel distortion, using the approximation method to compute the values of p and q is equivalent to computing and approximately describing the channel distortion distribution cluster under the error-free environment P0(qn), to replace the probability distribution cluster of the reference frame motion prediction part unable to be computed in channel distortion P(qn).

Definition 4: Assume P (qn) is the channel distortion cluster, H0 (qn; α) is the Hamilton's value of the reference frame motion prediction cluster, according to Paper [1], it can be known that:

$$P(q_N;\alpha)dq_n=\frac{e^{-\beta H(q_N;\alpha)+\int d^3r\lambda(r)\hat{n}(r)}}{Z}dq_N \quad (7)$$

$$P_0(q_N;\alpha)dq_n=\frac{e^{-\beta H_0(q_N;\alpha)+\int d^3r\lambda(r)\hat{n}(r)}}{Z_0}dq_N \quad (8)$$

In the above term, α represents the parameter of different reference frame motion prediction clusters, which can replace complicated macroblock coding distortion; $\boldsymbol{\lambda}$(r) is the Lagrangian factor to restrict the restriction formula related to the expected error-free environment of any position in the space (6).

From the maximum entropy principle, by use of the reference frame motion prediction cluster P0 (qN) to replace the channel distortion cluster P (qN), the probability distribution of this cluster can be described as:

$$S[P_0|P]=-k_b \int dq_N P_0(q_N) \log \frac{P_0(q_N)}{P(q_N)} \quad (9)$$

Put (7) and (8) into (9), and by use of the variational principle, the maximum entropy can be obtained, as

$$S[P_0|P]=\beta[\Omega-\Omega_0-\langle H-H_0\rangle_0] \quad (10)$$

In the above term, $\langle H - H_0\rangle_0$ represents the computing of the expected value based on the reference frame motion prediction cluster. Because $S[P_0|P] \leq 0$, the maximized entropy equals to all the minimized adjustable parameters α covering the reference frame motion prediction cluster,

$$\Omega_U=\Omega_0+\langle H-H_0\rangle_0, \text{ or: } \bar{\Omega}_U=\min\Omega_U \quad (11)$$

By this, the best parameter α can be found to replace P (qN) by P0 (qN).

V. Comparison of Error Resilience Abilities

Here, based on the JM75C software, the error resilience abilities of the random refresh strategy, the strategy of macroblock refresh in lines and the strategy of rate-distortion mode assessing over the whole situation are compared. The test sequence is carphone.yuv, the number of macroblocks refreshed each time in the random refresh strategy and the strategy of refresh in lines is 10, the number of decoders in the rate-distortion mode assessing strategy is 30 and the packet loss ratio is 10%. The numbers of reference frames are all 5, while other parameter settings are the same. The comparison of results obtained from reference frame motion prediction is shown in Table II.

It can be seen from the above reference frame motion prediction data that the error resilience ability in the strategy of rate-distortion mode assessing over the whole situation is apparently better than other two strategies. However, from the perspective of time consumed, even if in the situation that K is not big (e.g., K=30), its computing load and complexity are far bigger than other two strategies.

VI. Conclusion

H.264, as a new generation video compression standard, has absorbed the advantages of past coding programs and achieved great improvement in aspects like video compression coding performance, robustness and network friendliness, to lift the motive image compression technology to a higher level, capable to transmit high-quality images by relatively low bandwidth. This paper integrates the video transmission model under error channels, based on analyses of the intra-frame and inter-frame coding distortions introduced during the transmission of H.264 video coding data, the concealment distortion caused by channel errors and the distortion arisen from error propagation, discusses the macroblock coding mode rate-distortion assessing strategy of H.264, and combined the statistics and probability theory knowledge, by use of the approximation to the maximum entropy principle, discusses the computing methods of the distortion sum D(Mode) and the Lagrangian multiplication factor $\boldsymbol{\lambda}$ respectively. This computing method is simple and feasible, having fairly good flexibility and robustness. The reference frame motion prediction proves that this coding mode assessing strategy can effectively increase the error resilience ability.

TABLE II. COMPARISON OF PERFORMANCES IN THREE STRATEGIES

Mode	Parameter	Number of Reference Frames	QP	SNR(Y)	Bit-rate(dB)	Time Consumed (s)
Random Refresh	Number of Refreshed Macroblocks 10	5	28	36.93	323.20	7.351
Refresh in Lines	Number of Refreshed Macroblocks 10	5	28	36.95	331.36	7.160
RD	K=30, p=0.1	5	28	32.85	243.20	20.189

References

[1] ITU-T VCEG-N18,Performance evaluation of H.26L,TML8 vs.H.263++ and MPEG-4,VCEG(SG16/Q6), Fifteenth Meeting, Santa Barbara, CA, September 2001.

[2] ITU-T Q15-J-22,Video coding requirements for network robustness, VCEG(SG16/Q15), Tenth meeting, Osaka, 16-18 May, 2000.

[3] Gisle Bjontegaard, H.26L test model long term number(TML-8) draft0, VCEG (SG16/Q6), Fourteenth Meeting, Santa Barbara, CA, September 2001.

[4] ITU-T VCEG-N52, H.26L over IP and H.324 framework, VCEG(SG16/Q6), Fifteenth Meeting, Santa Barbara, CA, USA, September 2001:21-24.

[5] H.26L Test Model 8.0 software packet, ftp.pictel.com/Videosite/H.26L/,2001.