# H.264/AVC Analysis of Quality in Wireless Channel

Alexander Chuykov

State University of Aerospace Instrumentation
Saint-Petersburg
Email: alex@chkv.net

**Abstract**

The article describes main conception how to use error protection codes to transmit digital video over wireless channels. The article describes experiments which obtain parameters for these codes, decoder behavior and some results. The objective metrics are described too.

## I. INTRODUCTION

The market of mobile devices goes up. There are lots of new devices with new features. They include mobile communication, mobile internet, photo- and video-cameras, GPS and etc. Software and customers services improve existing services and add new such as conference call, video call, multimedia libraries and so on.

Like any radio frequency transmission, wireless networking signals are subject to a wide variety of interference, as well as complex propagation effects that are beyond the control. Of course the video transmission over the mobile network requires the special error-cancelation algorithms [1].

The simplest error-cancelation algorithm is usage the error protection codes [2]. But this way rough. The significant of bit varies in the stream. Some bits are more important than other. Another way, special codes can be used for until compression e.g. RVLC (Reversible Variable Length Code, [3]). But they do not cover some important information like headers.

The model below describes how to choose optimal code for different parts of the stream with its own significant.

## II. CHANNEL MODEL

Let see the generic system of the video transmission over wire- and wireless- networks. It includes different kind of wire and wireless devices (cellular, PDA, cameras, netbooks, mediaservers, tablet PC and etc), different wire and wireless channels (Ethernet 10/100/1000, IEEE 802.11abgn, 3G, 4G and other).

### *A. Model for simulation*

This model is not useful for modeling. Look upon another model of communication channel. It consists of:

1) Video encoder. It creates the compressed video-sequence with known attributes (like frame sizes, frame rate and quality)
2) Video streams parser. It separates compressed video-sequence to (sub-) streams. Each stream has channel error model with a priori attributes. The parser simulates the channel coder, channel decoder and channel

3) Video decoder. It restores compressed frames
4) Statistics acquisition software

*B. Improved model for simulation*

It is easy to notice that video stream parser should merge with video decoder. Thus decoding process is not changed and data processing is simplified. This way has following advantages:

1) The source code of video encoder is not change. The compressed video-sequence can be restored by any compatible decoder
2) Only one copy of compressed video-sequence is required for all experiments
3) Video decoding is more faster than video encode

But this model has one limitation. The error positions vary on each run. It is hard to have sequences with fixed positions of errors. But this limitation can be avoided by setting the same seed of the random number generator for each model run.

The behavior of the decoder (channel model) describes by configuration file.

*C. Test equipment*

Two video-sequences were used for research. The first one is ”Claire“ (slow motion) and the second is ”Trueman“ (fast motion). These video-sequences are well-known and used for tests often.

The PSNR (peak signal-to-noise ratio) was used as objective measure of quality. The visual quality is very important, but it was measured for some custom clips only. The PSNR compares with reference video-sequence. This is the restored (compressed and decompressed) videosequence without any channel errors. Of course, other objective metrics more accuracy and taking into image structure. The section III below compares two objective metrics and chooses one.

## III. Metrics compare

*A. PSNR*

The PSNR is objective measure. It bases on mean-square error (1) (MSE).

(1)

Where $x_i$ are items of original sequence and are items of restored sequence. The sequences have same length N. The N is area of images. Thus (1) describes the mean power of error ($E_N = MSE$). The mean power E of signal can be described as (2).

(2)

So PSNR can be calculated by formula (3).

(3)

Obviously, the mean power of signal depends on color depth L (bit per pixel). So, $2^L-1 \geq E$ is true, and the (3) can simplify to (4).

(4)

We can note that the complexity is linear and depends on $N$. The integer computations perform for $E_N$ in common cases. Total number of computations is $2N-1$ for sum and $N$ for multiplication. Formally the complexity of PSNR is $\Theta(N)$. We should notice that MSE and PSNR do not recognize the different kinds of noise on image. It is clear that modern method of image comparison should have following features:

1) The image has geometry not sequence of pixels
2) The different kinds of noise have different effect on visual quality

*B. SSIM*

The structural similarity (SSIM) index [4] compares intensity group of pixels. The pixels are normalized by luminosity and contrast. The value of SSIM is normalized. The ”1“ means highest quality, ”0“ means worst. The restored image was specified as . Formula for SSIM is (5).

(5)

The image was processed by square window size M (block). Each block has its own metric’s value. The result SSIM computes as mean for all blocks. The parts of (5) are described by formulas (6) to (9). The acute accent means the formulas are correct for one block only.

(6)

(7)

(8)

(9)

The encoder block determines weight vector and window size $M$. Authors of SSIM recommend set window size $M=11$. The weight function is weight function of normal distribution for $\sigma_\mu = 1.5$ and normalized values .

*C. Complexity estimation*

The program was written to estimate complexity of SSIM. The program language was chosen to ”C“, the reference source was written in ”M“ (Matlab). The image has width W and height $H$, the area is $N=W\times H$. The formulas (6), (7) and (8) use bilateral convolution with $O(NM)$ complexity [5]. The number of convolutions is 5. The elementwise multiplication is required for $(W-M+1)\times(H-M+1)$ matrixes 17 times. The complexity of elementwise multiplication is $\Theta(N)$. Totally, the complexity of SSIM is $5O(MN)+17\Theta(N)$. This is more complex than PSNR. It is interesting to note that computations are float for SSIM. Let us identify the memory usage. There are 6 auxiliary arrays $(W-M+1)\times(H-M+1)$ and one array $N$ were required after optimization. The PSNR requires some extra memory cells for subproducts.

*D. The PSNR and SSIM usage*

The complexity of SSIM limits it. It is difficult to substitute PSNR/MSE for SSIM. Also, SSIM has no PSNR disadvantages. Thus SSIM is more accuracy and can avoid the expert judgments partly. Research shows the SSIM and PSNR behavior for different kinds of video sequences and bit rates. So, the results are:

1) For high bit rate SSIM and PSNR have the same behavior
2) For low bit rate SSIM has large proportional range of values than PSNR

Resume, for high bit rate PSNR is more preferable (it has less complexity). For low bit rate SSIM is better to increase codec adaptation accuracy. The research in this article has a high rate priority. So, the PSNR is major objective metrics for research.

## IV. Video codec

The H.264/AVC (ITU-T H.264 and ISO/IEC 14496 (MPEG-4) Part 10), reference 14.2 was used as video codec [6]. This codec has a lot of pre-installed profiles to encode video with different quality and resolution from mobile phones up to home theater [7]. The H.264 is best universal video-codec and it becomes most popular for mobile devices. Some options which are used for encode video-sequence:

1) The VLC (variable length codes) were used instead of CABAC (Context-based Adaptive Binary Arithmetic Coding) to decrease research complexity
2) The RTP (Real Time Protocol, [8]) was used as container for compressed data

## V. Research results

The research shows that H.264 (reference version) has not any error protection. Channel error halts decoder for probability $P_e > 10^{-6}$. All bits in compressed video-sequence were classified to obtain limiting characteristics. There are two classes of bits. The first one is fatal. Any error in fatal bit halts decoder immediately. The fatal error can locate in file header, macroblock type and etc. The second class is nonfatal bits. Any error in nonfatal bits corrupt image, but not halt the decoder. Subsequently, the compressed video-sequence was separated to sub-streams by size in bits. For low rate and fast-motion clip the motion vectors (MVD) have portion 30% of compressed video-sequence size. When rate increasing, the portion of the MVD decreases. For rate ≈ 1 MBit, the luminosity AC for inter frames (Luma AC inter) has portion more than 50%, the MVD has minor portion.

To choice error protection code, the quantity of fatal bits should be estimated. For these purposes the errors locate in MVD sub-stream only. There are three kinds of bits in the stream - fatal, nonfatal and others. Obviously the model is simple but it provides estimate portion of fatal bits for MVD and chooses the error protection code. The experiments show that the portion of fatal bit is high (≈ 90% of all MVD bits). It is easy to explain. The error corrupts the MVD codeword not vector.

The results of experiments show:

1) That is impossible to decode H.264 stream without any error protection for channel error probability $P_e > 10^{-6}$

2) The H.264 stream decoding is possible for channel error probability $P_e > 10^{-6}$. PSNR is lower than reference about 13 – 30 dB
3) The separation of the H.264 stream to sub-streams decreases the PSNR falling (up to 3 – 8 dB in comparison with reference). Each sub-stream has independent error protections options which depends on sub-stream significance

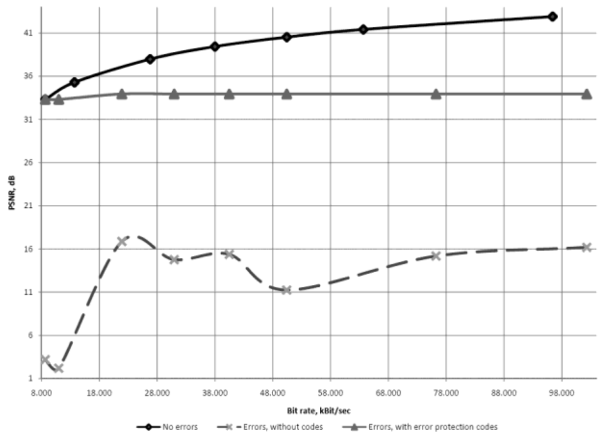


Figure 1. Simulation result for “Claire” clip

It should be noted that comparison of compressed video-sequences performs for same rate. It means that rate for sequence with error protections codes has more compression rate than the same sequence without any error protection code. But the result rate (file size) is the same.

Figures 1 and 2 show some rate-distortion curves for each video sequence for channel SNR =7dB. The first curve shows the rate-distortion for channel without errors. The second curve shows the rate-distortion for channel with errors but error protection codes are not used. The third curve shows the rate-distortion for channel with errors and error protection codes are used too.

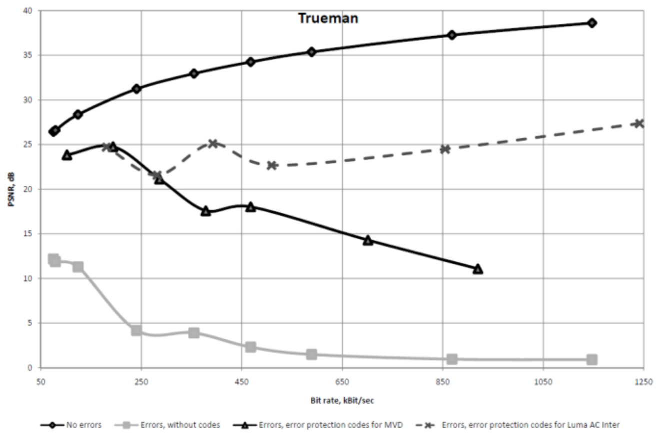


Figure 2. Simulation results for “Trueman” clip

## VI. Conclusion

H.264 is one of the best codec for video. But it is quite sensitive for channel errors. The error protection codes with separation for sub-streams increases objective quality of video. The future research will improve:

1) The error protection codes and algorithms
2) The model of channel to model existing networks

## References


[1] T. Stockhammer, M. M. Hannuksela, and T. Wiegand, “H.264/avc in wireless environments,” *IEEE Trans. Circuits Syst.*, vol. 13, no. 7, Jul. 2003.

[2] B. Sklar, *Digital communications. Fundamentals and Applications*. Prentice Hall PTR, 2004.

[3] Y. Takishima, M. Wada, and H. Murakami, “Reversible variable length codes,” *IEEE Trans. Commun.*, vol. 43, pp.158–162, 1995.

[4] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[5] T. H. Cormen, C. E. Leiserson, and R. L. Rivest, *Introduction to Algorithms*. The MIT Press, McGraw-Hill Book Company, 1999.

[6] I. E. G. Richardson, *H.264 and MPEG-4 Video Compression. Video coding for Next-generation Multimedia*. John Wiley & Sons, 2003.

[7] A. M. Tourapis, G. Sullivan, K. S¨uhring, T. Oelbaum, and A. Leontaris, *H.264/14496-10 AVC Reference Software Manual*, 2009.

[8] S. Wenger, M. M. Hannuksela, T. Stockhammer, M. Westerlund, and D. Singer, “RTP payload format for H.264 video,” RFC 3984, Feb. 2005.