

The Application of H.264 Video Coding Technology in the Digital Monitoring Systems

Wang Yuncheng

Dongchang College of Liaocheng University, No. 105 Huanghe Road, the Economic Development Zone of Liaocheng, Shandong Province, China

Abstract: This study presents an application of H.264 video coding technology in the digital monitor system, in which H.264 was used as a new method. And now the majority of monitor system use MPEG1/MPEG2 for coding the video while consumes large disk and net load cost. For getting the goal of real time data compression, we proposed three sides to enhance the computing speed: the first is optimization of selection of inter-frame mode, the second is tune the computation structure of interpolation, the last is assembly optimization with MMX and SSE.

Keywords: Digital monitoring system, H.264, video coding technology

INTRODUCTION

As technology continues to evolve, the video surveillance system has been widely used in intelligent transportation, smart buildings, shopping malls, supermarkets, hospitals, campuses, corporate production and living quarters and other range, has penetrated into all areas of production and life. Video surveillance is much more intuitive, convenience and rich content, increasing attention has been paid. The total market value of the national image monitoring system is rapid growth in recent years (Zhou and Deng, 2004). China's video surveillance industry is facing good opportunities for development.

In the domestic and overseas markets, there are two types of products digitally controlled analog video surveillance and digital video surveillance. The former technology development has been very mature, stable performance and has been widely used in practical engineering, especially in large and medium-sized video surveillance project is particularly extensive. The latter is a newly emerging video to computer technology and image compression as the core of a new video surveillance system, the system to solve the drawbacks of the analog part of the system, but still needs further improvement and development (Zhou and Deng, 2004; Kaup and Mooshofer, 1999; Panusopone *et al.*, 2000).

At present general remote video monitoring system, the video data compression and decompression completely based on the PC video card to be finished, so video front (such as CCD video signal collection, compression and transmission) is more complex, reliability and stability is not high and the price is more expensive, occupied space is bigger also (Kaup and Mooshofer, 1999). Now there is a new type of network remote video monitoring system, which is based on embedded video server technology of remote video

monitoring. Embedded video server front end joint camera, video the server code, through the Internet transmits to the client, the client through the corresponding software to watch video (Panusopone *et al.*, 2000).

Based on the embedded video server technology of remote network monitoring system not only require a high performance multimedia terminal support, but also have a multimedia information transmission network of communication as the foundation (Dong and Tao, 2010). This monitoring system to network, supported by digital video compression, transmission and decoding play as the core and the intelligent practical image analysis is characteristic, promote the monitoring industry of new developments in technology have received great attention of academia, industry.

At present most of the digital video monitoring system used MPEG1/MPEG2 video coding standard, resulting in large amount of video data, video data storage cost is high, for the integrated network video monitoring, it is difficult to through the network real-time playback, H.264 video coding has better compression performance and adapt to a variety of channel capacity, the code flow structure network adaptability and its application prospect is very broad, has been the industry's attention. The H.264 video coding technology of video monitoring system in obtaining good video playback quality at the same time, greatly improve the video compression encoding efficiency, effectively save digital monitoring system of storage cost (Dong and Tao, 2010).

The Objective of the study is to propose three sides improvement H.264 video coding technology in the digital monitor system by enhancing the computing speed. The first is optimization of selection of inter-frame mode, the second is tune the computation structure of interpolation, the last is assembly optimization with MMX and SSE.

H.264 VIDEO CODING

H.264 concept: Figure 1 shows the encoding and decoding diagram of the standard. The standard includes the use of block-based motion compensation, DCT-like residual coding, scalar quantization, zigzag scanning and entropy coding. Furthermore, the standard includes some novel features to offer better compression efficiency (Panusopone *et al.*, 2000).

First, the motion compensation model used in H.264 is more flexible and efficient than those in the early standards. Multiple reference frames for prediction is supported in the standard and a much larger number of different motion compensation block sizes are used for each macro-block (16×16 , 16×8 , 8×16 , 8×8 , 8×4 , 4×8 , 4×4). In addition, de-blocking filter within the motion compensation loop is used to reduce visual artifact. The new methods provide a more precise model for motion compensation, which can dramatically minimize the impact of the difference of predicted blocks and yield a much better perceptual quality for the decoded video stream.

Another uniqueness of H.264 is that the DCT transform is replaced by a DCT-like integer transform

in the H.264 standard. These transforms can be implemented with additions and shifts. Therefore they are so-called ‘multiplication-free’, which is believed to have a significant computational advantage compared with DCT transforms. Also the small size helps to reduce the blocking and ringing artifacts. Moreover, the precise integer transform eliminates any mismatch between the encoder and the decoder.

H.264 video coding features: H.264 used many different advanced technology which don’t used in the previous standard, such as the VLC symbols encoded, integer transform, with high-precision, multi-mode displacement estimation, layered coding syntax. These make the algorithm of H.264 has a high coding efficiency and in the same reconstructed image quality, better than H.263 to save about 50% of the bit rate, than H.263 better compression performance and ability to adapt to a variety of channels. H.264 stream structure has good network adaptability and error recovery capability.

Similar to earlier standards, the H.264/MPEG-4 AVC standard specify the syntax for a compliant bit stream as well as a set of decoding semantics that

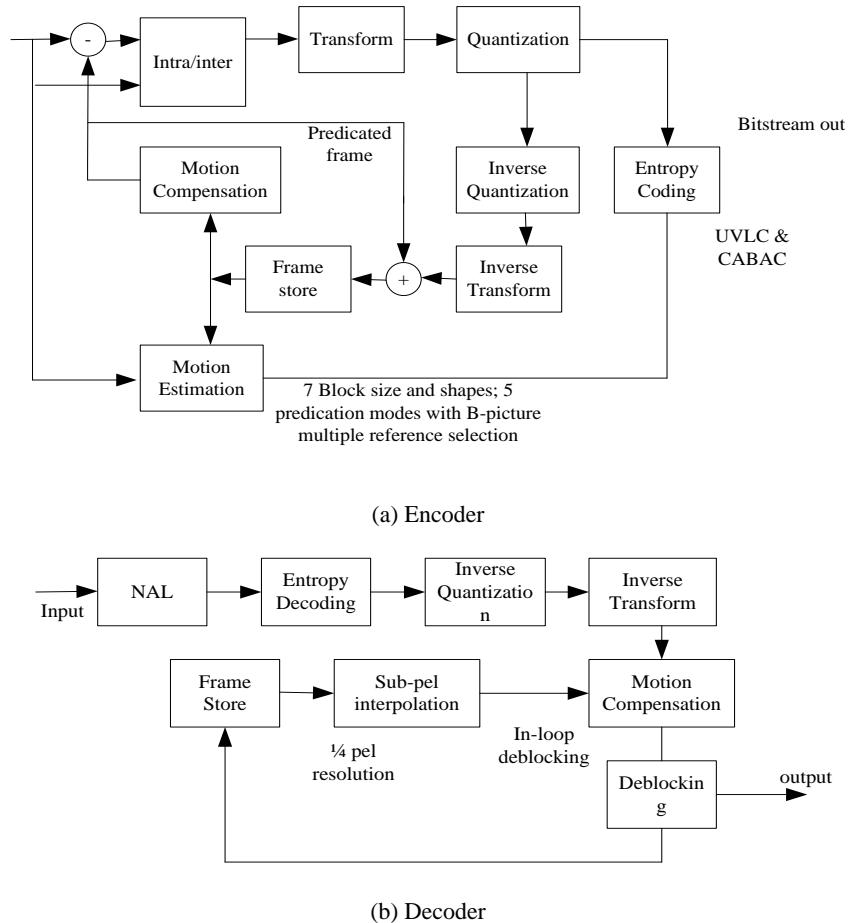


Fig. 1: Encoder and decoder diagram of H.264

describe how to interpret the syntax elements in the bit stream to produce decoded pictures. Thus, the decoding operations and thus the decoder are fully specified but there is considerable flexibility in design of the encoder. However, as in earlier standards, even encoder must follow a standard set of operations for the resulting picture quality to be good, although algorithmic shortcuts are often taken at the encoder to tradeoff picture quality performance for speed. In fact, due to a plethora of prediction modes in H.264/MPEG-4 AVC, considerable effort was put by the JVT committee in demonstrating good coding quality (although optimize decoding is not standardized) to show value of many of the available prediction modes.

H.264 video compression system components: H.264 standard compression system is composed of two parts by the video coding layer (VCL) and network abstraction layer (Network Abstraction Layer, NAL). VCL, including the VCL encoder and VCL decoder, the main function is to video data compression encoding and decoding, which comprises a motion compensation, transform coding, entropy coding, such as compression unit. NAL is used to provide a uniform interface with a network independent for VCL and it is responsible for the encapsulation of it in the network transmission of the video data, using a unified data format and includes a single byte in the header information, a plurality of byte video data and a set of frames, the logical channel signaling, timing information, the sequence end signal. The header contains the stored sign and types of signs. The storage flag is used to indicate the current data does not belong to the reference frame.

High-precision, multi-mode motion estimation: H.264 supports a motion vector of 1/4 or 1/8 pixel precision. 1/4 pixel accuracy, using the 6-tap filter to reduce high-frequency noise, for the 118-pixel precision motion vector may be an 8-tap filter using a more complex filter. Performing motion estimation, the encoder can also choose to enhance interpolation filter to improve the prediction results.

In H.264 of the motion prediction, a Macro Block (MB) can be divided into different sub-blocks, forming seven different modes of the block size. Flexible and meticulous division of this multi-mode more attuned to image the actual shape of the moving objects, greatly improving the accuracy of the motion estimation. In this manner, each macro block can contain 1, 2, 4, 8, or 16 motion vectors.

In H.264, allows the encoder to use more than one of the previous frames is used for motion estimation, which is called the multi-frame reference technique. For example, 2 or 3 just coding a good reference frame, the encoder will select the macro block to get a much better prediction frame for each target and for each macro block indicates which frame is used to predict.

4X4 block integer transform: H.264 is similar to the previous criteria, using block-based transform coding for residual, but the transform is an integer operation instead of the real number operation, substantially similar to the process and the DCT. The advantage of this method is that: in the encoder and decoder allows the same precision transform and inverse transform, easy to use a simple fixed-point arithmetic. In other words, there is no "anti-conversion error. Transform unit 4X4 block, rather than the conventional commonly used in 8x8 blocks. Due to the reduced size for the transformation blocks, the division of the moving object is more accurate.

H.264 in order to improve the ability to rate control, the quantization step change in amplitude control of about 12.5% instead of the constant increase changes. The normalized amplitude of the transform coefficients on the inverse quantization process is processed to reduce the complexity of the calculation. In order to emphasize the color fidelity of the chrominance coefficients, using a smaller quantization step size.

Unified VLC: For H.264 entropy coding there are two methods, one is to be encoded symbol unified VLC (UVLC: Universal VLC), another is the use of the content-adaptive binary arithmetic coding (CABAC: Context-Adaptive Binary Arithmetic Coding). CABAC is optional, its coding performance is slightly better, but also has high computational complexity. UVLC using a length of the infinite set of a code word, design structure is very regular and is encoded using the same code table can be a different object. This method is very easy to generate a code word, while the decoder is also very easy to identify the code word prefix. UVLC bit error can be quickly obtained to re-synchronization.

Intra prediction technology based on airspace: With the help of removing the image space and time correlation video coding to achieve the purpose of compression. Spatial correlation through effective transformation to remove, such as DCT transform, H.264 integer transform; time correlation is removed by inter-frame prediction. Here said transform to remove spatial correlation, is confined within the transform block, such as 8 X 8 or 4 X 4 not between blocks processing. H.263+and MPEG-4 introduce the intra-frame prediction techniques, in the transform domain is based on adjacent blocks do some of the coefficients of the current block prediction. H.264 is in the airspace, the use of the adjacent pixels of the current block directly to each coefficient prediction, more effectively to remove the correlation between adjacent blocks, which greatly improves the efficiency of the intra-coded.

MATERIALS AND METHODS

The main framework of H.264 in video monitor system: The video monitor system structure showed in

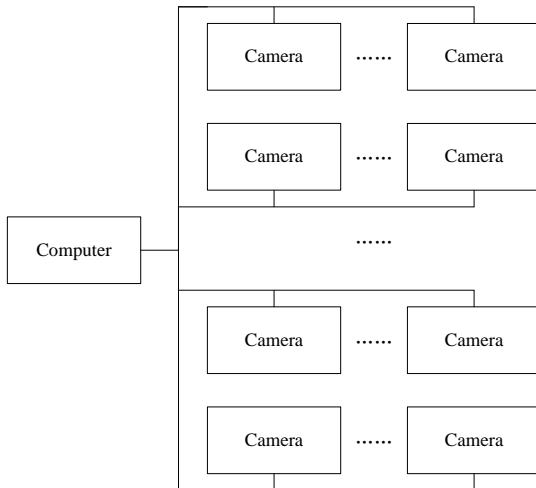


Fig. 2: Video monitor system structure

Fig. 2. Camera shots will send the video signal to video monitor calculation center, host encoding compression processing of these signals, the encoded video data streams are stored directly on the hard disk inside and then the pick each camera image is displayed on a display, to achieve recording, playback, display and transcription duplex synchronous operation. Currently, most of the video surveillance system using MPEG-1/MPEG-2 of video coding standards, video media files requires a lot of storage space, 352*288 video resolution, 30 frames per second image, single-camera video bit-rate about 250M byte/h. Table 1 shows the experimentally measured to the storage space occupied hour video files in different compression (Mo *et al.*, 2008).

The advantages of H.264 based monitor system: Digital surveillance video required a higher quality of video playback, real-time processing of digital video signals and these video data stream is stored in the computer's hard drive, to achieve real-time playback. The current monitoring system, most of the video coding standard used MPEG-1/MPEG-2. The encoded video file requires a large amount of storage medium space. The study shows, H. 264 encoded video stream and H.263 or MPEG-4 Simple profile encoded video streams compared to an average savings of 30% to 50% of the bit rate (which vary with the type and size of compressed video sequences). Table 2 shows certain types of video sequences of several common encoding bit rate savings comparison results.

The disadvantages of H.264 based monitor system Because H.264 with up to 13 kinds of the intra prediction mode (4×4 size 9, 16X 16 size 4) and 8 sizes, category 5 of the inter-frame prediction mode, therefore, it will consume a large amount of calculation, while using a CABAC coded manner, which will

Table 1: Video size comparison of different types of coding

Types of compression	Static	Normal motion	Dramatic motion
MJPEG	535	625	950
MPEG-1	240	240	240
MPEG-2	205	205	205
H.264	60	86	146

Table 2: Save code rate compassions of several types of coding ways

Coding methods	MPEG-4ASP	H263HLP	MPEG-2
H264	37.44	47.58	63.57
MPEG-4ASP		16.65	42.95
H263HLP			30.61

increase the implementation complexity. H.264 with respect to the MPEG-2 implementation, the complexity of the encoder has increased about 5-fold, while the complexity of the decoder is increased by about 3 times. In addition, H.264 also uses a fourth interpolation algorithm, leading to expanded 16-fold in the temporary storage space occupied by the same number of reference frames.

RESULTS AND DISCUSSION

In order to realize the goal of real-time encoding compression, from several aspects to improve its operation speed:

- By analyzing the statistical distribution of the different inter-frame macro-block mode and the transition probability of the intra-frame, to reduce the number of search and calculate the different modes, while significantly improve processing speed, compression efficiency is unchanged and only slightly lower compression quality
- Adjust encoder operation structure
- Mmx, sse, assembler optimization

Optimize inter mode selection: Motion compensation inter-frame coding, H.264 supports seven different sizes and shapes of the macro-block (16×16) and the sub-macro-block (8×8) is divided, respectively: 16×16, 16×8, 8×16, 8×8, 8×4, 4×8, 4×4 (Fig. 3). By RDO (Rate Distortion Optimization) to select a different block size.

The original program of the JM calculated for each mode distortion D and encoded bit rate R, according to RD optimized results to choose a model with the smallest RD cost as the final coding mode. In fact, the probability distribution of the various modes is not the same and between different model calculations with a lot of repetition. The actual monitoring system in the mode of statistics found that the model 16×16 accounted for the largest proportion, more than half and the second is model 16 × 8, then 8×16.

According to the above findings, we improve the process of program. The first model is 16×16

Table 3: Time cost of coding before and after mode selection

Picture	Original code time cost/us	Optimized code time cost/us	Optimization ratio/%
Foreman	13 281 790.16	2 815 739.51	21.20
Hall	13 250 412.79	2 477 827.19	18.70
Mobile	13 462 989.61	2 329 097.20	17.30
Paris	13 315 486.56	2 729 674.74	20.50
Bus	12 815 486.56	2 511 835.37	19.60
Akiyo	13 115 486.56	2 506 706.49	19.10

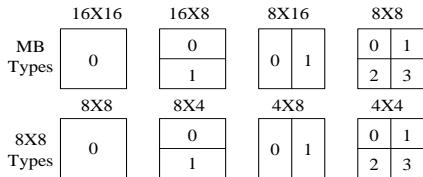


Fig. 3: Different split mode of motion search in H.264

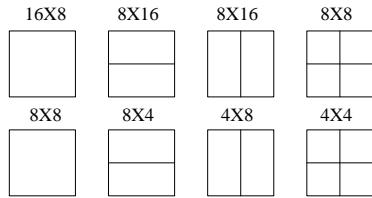


Fig. 4: Macro and sub-macro split

calculation, if it less than a certain threshold, code this mode and skip the other mode comparison. Table 3 shows the performance of optimization before and after the comparison, to be fairly, we adopted some common video sequence (Lin, 2009).

Operation structure of adjusting interpolation algorithm: H. 264 motion vector displacement accuracy for the luminance component is a quarter pixels, for the chrominance components is 1/8 pixel. 112 pixel sample used 6 tap filter (1,-5, 20, 20,-5, 1) which get from the whole pixel sample. To use linear interpolation for 1/4 pixel sample, with eight tap filter achieve 1/8 pixel accuracy. Each pixel accurate position was shown in Fig. 4.

A quarter pixel accuracy for 1/2 pixel accuracy under the condition of coding efficiency, enhance obviously. But 1/8 pixel accuracy relative to the quarter pixel accuracy of coding efficiency in addition to high bit rate circumstances and no obvious enhancement and 1/8 pixel interpolation formula is more complicated, the actual application is not strong. Therefore, in the

practical application of h. 264 only commonly adopted in the quarter pixel accuracy.

After interpolation image expand 16 times, in the original JM program into a single piece, so that, when inserted operation, on the one hand, cause inconvenience, cannot read and write memory for reading and writing. According to the above shortcomings, readjust the interpolation of image memory allocation. After interpolation, with different location to form 16 pictures which has the same size. So, speaking, reading and writing memory is fluent, motion estimation is convenient. Especially, during the MMX or SSE optimization, it does not need additional costs.

The MMX, SSE assembly optimization: Improve the encoder algorithm speed mainly depend on single instruction multiple data (SIMD) parallel operation, Intel X86 system of the SSE and MMX instruction set, to realize SIMD operation become possible. SIMD optimization based on block image difference, based on block transformation and SAD value calculation part see Table 4 and 5. Table 4 and 5 show that, through the above measures, in the video quality of H.264 coding speed has improved greatly (Chen and Zhou, 2006; Wu and Wu, 2004; Youngsub *et al.*, 2009).

The optimization results: Through the inter-frame macro block model selection method improvement, encoder operation structure adjustment and using MMX, SSE assembly instruction optimization, make the image coding speed to get bigger improve. Table 6 is to optimize the time before and after the contrast of image coding table, the measured after optimization to improve the encoding time about one order of magnitude.

Due to the above three kinds of optimization method is to improve the original JM operation method, optimization while image storage space remain unchanged. Inter-frame mode selection optimization, if the threshold value is undeserved, can affect the quality of the images or coding speed, but through the experiment statistical measure appropriate threshold, skip the other mode comparison operation, can be doubled the encoding speed, at the same time doesn't make the image quality is significantly reduced, in general occasions (image motion not acuteness), the human eye is difficult to distinguish the subtle changes. While the other two kinds of optimization method

Table 4: Time cost of coding before and after assembly optimization

Picture	Original code time cost DCT + Quantization/us	Optimized code time cost DCT + Quantization /us	Optimization ratio/%
Foreman	2 889 934.42	238 690/3	8.26
Hall	2 884 860.92	150 988.38	5.23
Mobile	2 966 516.84	296 742.23	10.00
Paris	2 886 868.87	176 126.53	6.10
Bus	4 034 638.472	290 493.97	7.2
Akiyo	1 478 054.197	142 632.23	9.65

Table 5: Time cost of interpolation algorithm before and after assembly optimization

Picture	Original code time cost/us	Optimized code time cost /us	Optimization ratio/%
Foreman	10 391 846.74	135 381.36	1.30
Hall	10 365 551.87	133 844.75	1.29
Mobile	10 496 472.77	134 158.16	1.28
Paris	10 428 617.69	138 163.39	1.32
Bus	10 587 220.47	134 457.70	1.27
Akiyo	10 234 676.34	134 074.26	1.31

Table 6: Coding time cost comparison before and after optimization

Picture	Original code time cost/us	Optimized code time cost /us	Optimization ratio/%
Foreman	49 806 713.10	5 345 824.25	10.73
Hall	49 689 047.96	4 571 240.18	9.20
Mobile	50 486 211.04	4 786 346.87	9.48
Paris	49 933 074.60	5 037 381.87	10.09
Bus	50 316 040.08	5 042 608.07	10.02
Akiyo	47 429 629.57	4 590 179.21	9.68

is image lossless optimization, won't cause a drop in the quality of image, using the above three methods for optimization and ensure the quality of the image, make the image coding computation to obtain the very good optimization, greatly shorten the operation time of coding, it is very important on the multiple digital monitoring system.

CONCLUSION

Because the storage cost and network bandwidth is high of digital video monitoring system based on the current MPEG1/MPEG2, the study proposes the H.264 frequency coding technology solutions. A new generation of multimedia coding technology H.264 algorithm has high coding efficiency, in the same reconstruction image quality, than to H.263 save 50% of code rate. H. 264 code flow structure network adaptability, increased the error recovery ability and by using previous mature coding technology, in the pursuit of high coding efficiency and concise expression, but also provides a very good video quality, is the most efficient video compression method. In this study the video monitoring H.264 coding of the specific application, it put forward to enhance the speed of operation optimization method. Through the analysis at different frame between macro block model of statistical distribution and frame, frame in between the transition probability and put forward effective, rapid

variable-size block model selection method, reduce the search and calculating different model number, significantly improve the operation speed and at the same time compression efficiency unchanged, compression quality is slightly lower. In addition, by adjusting the encoder operation structure and using MMX, SSE assembly optimization, greatly enhance the video coding speed. The algorithm proposed in this study can satisfy the general digital video monitoring system requirements.

REFERENCES

- Chen, Z.B. and P. Zhou, 2006. Fast Integer Pel and Fractional Pel Motion Estimation for JVT. Awaji Island, JVT of ISO/IEC MPEG and ITU-T VCEG, pp: 264-290.
- Dong, P. and L. Tao, 2010. Re lose free intra frame coding based on H.264. J. Electron. Inform. Technol., 32(1): 117-122.
- Kaup, A. and H. Mooshofer, 1999. Performance and complexity analysis of rate constrained motion estimation in MPEG-4. Proceedings of the SPIE Multimedia Systems and Applications II. Boston, 0020 Mass, September 20-22, SPIE 3845: 202-211.
- Lin, T., 2009. Achieving re-loss-free video coding. IEEE T. Signal Proces. Lett., 16(4): 323-326.
- Mo, L., J. Bu and C. Chen, 2008. Fast predictive mode decision algorithm in H.264. J. Zhejiang Univ., 42(4): 612-617.
- Panusopone, K., X. Chen and R. Eifrig, 2000. Coding tools in MPEG-4 interlaced video. IEEE T. Circ. Syst. Vid., 10(5): 755-766.
- Wu, D. and S. Wu, 2004. Block inter mode decision for fast encoding of H.264. Proceedings of International Conference on Acoustics, pp: 181-184.
- Youngsub, J., J. You, W. Kim and J. Jeong, 2009. Fast mode decision algorithm using efficient block skip techniques for H.264 P slices. Proceeding of 1st International Conference on Advances in Multimedia, pp: 92-97.
- Zhou, Y. and Z. Deng, 2004. Research and application of H.264 in video coding. Comput. Inf. Technol., 2004(8): 111-118.