The Development of MPEG4-AVC/ H.264, the Next Generation Moving Picture Coding Technology

Zhixiong Wu Takaaki Hatanaka
Masayuki Tokumitsu

Introduction

Popularization of broadband networks, such as xDSL, FTTH and high-speed wireless LAN, is proceeding at a high rate, resulting in a rapid increase in the number of broadband network subscribers. Further, the full-scale implementation of video communication services, such as internet broadcasting or video distribution services, videophone and videochat over broadband networks, are about to take place. First indications have emerged of the popularization of digital audio visual equipment for broadcasting digital HDTV, as well as the reception and recording of such broadcasts, with the Olympic Games triggering such a trend. Furthermore, digital mobile broadcasts for mobile telephones are also about to begin. We are indeed witnessing the advent of the fusion of telecommunications and broadcasting, as well as the arrival of a ubiquitous society.

Existing video coding technologies across the board are inadequate for the transmission of much clearer content with more images over a network with a limited band. For example, in order to broadcast video with a quality equivalent to that of HDTV transmission is possible with only a handful of channels at 100Mbit/s even with FTTH. The situation is such that it is difficult to receive even a single channel on ADSL. Therefore, expectations for the next generation moving picture coding technology with better compression rates are quite high.

In support of the broadband video distribution solution Oki Electric has been developing a total video broadcasting system, the OKI MediaServer*1). The software CODEC, which supports the MPEG4-AVC/ H.264, the latest international standard for the next generation moving picture coding technology, was developed by Oki Electric in response to the aforementioned expectations for a highly efficient moving picture coding technology. Our software CODEC realizes the capability of executing real-time compression and decompression of D1 class images on personal computers readily available on the market. This paper will first introduce a summary of the international standard for moving picture coding, MPEG4-AVC/H.264, followed by an introduction of the technical features of the high-quality MPEG4-AVC/H.264 CODEC we have developed.

Summary of the MPEG4-AVC/H.264

The MPEG4-AVC/H.264 is the latest moving picture coding method formulated jointly by the leading minds of the world in the International Standard Organization (ISO) and the International Telecommunication Union (ITU). The Moving Picture Experts Group (MPEG) from ISO participated in the formulation of the standard, which is recommended for standardization, 14496-10, the Advanced Video Coding (commonly known as the MPEG4-AVC), by ISO. Similarly, a group of moving picture experts, the Video Coding Experts Group (VCEG) representing ITU, participated in the formulation of a standard, which is recommended as H.264, by ITU. The standard of this technology is referred to as "MPEG4-AVC/H.264" in this paper, in a format that combines the two participating parties.

So far MPEG have been formulating a lot of international standardizations for moving picture coding technologies. Such international standards are often named with numbers added to the abbreviation of the group's name. For example, the MPEG1, often used for VCDs and Karaoke, the MPEG2, adopted for DVDs and digital broadcasting, as well as the MPEG4, used for mobile communications and internet broadcasting, are all well known international standards. VCEG, on the other hand, have similarly also been formulating many international standards for moving picture coding technologies. These international standards are denoted as the H.26x series standards. For example, the H.261 adopted for videophones, the H.262 adopted for digital broadcasting, the H.263/H.263+ adopted for videophones, teleconferencing and mobile videophones, are all well known. In the past a joint standardization by these groups was also achieved, with the standardization of the MPEG2/H.262. It would not be an exaggeration to call the latest joint operation for moving picture technology an effort that is truly an orchestration of leading minds from around the world.

Technical features of the MPEG4-AVC/H.264

The composition of the MPEG4-AVC/H.264[1)] is shown in Figure 1. The MPEG4-AVC/H.264 performs coding in units of blocks and along a time axis the redundancies between frames are eliminated through the extraction of predicted residuals by compensating for the movements of blocks, as shown in Figure 1. Along the spatial axis, quadrature conversion and variable length coding occurs for each block to eliminate redundancies within each frame. Technically, therefore, this standard can be considered an extension of the MPEG4 and the H.263. The similarity is quite evident when compared with the composition of the MPEG4-ASP that was described in the article listed as the second reference document for this paper. A large number of modifications have been implemented to the respective technical

*1) OKI MediaServer is a registered trademark of Oki Electric Industry Co., Ltd., for their universal video distribution system.

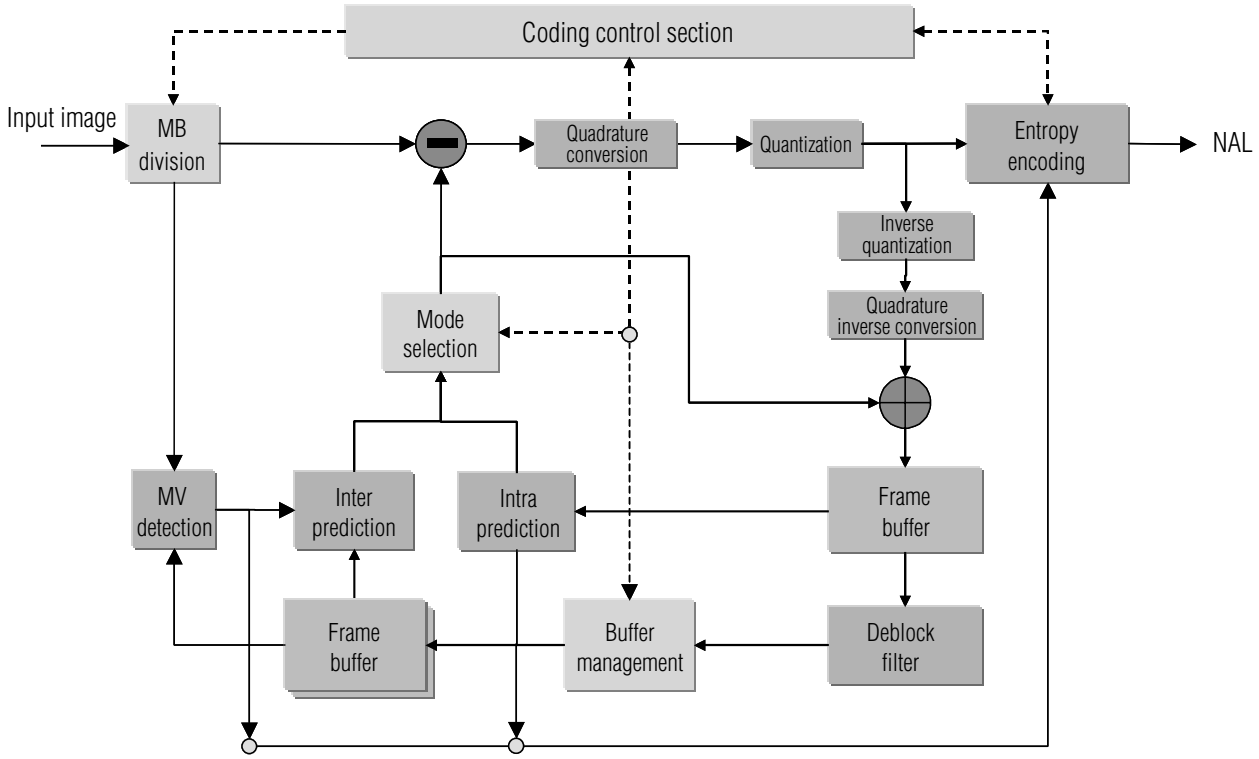

Fig. 1 Composition diagram of MPEG4-AVC/H.264

Table 1 Features of the MPEG4-AVC/H.264 (vs. MPEG4-ASP)

Technical item	MPEG4-ASP	MPEG4-AVC/H.264
Block division	16 x 16, 8 x 8	16 x 16, 16 x 8, 8 x 16, 8 x 8, 8 x 4, 4 x 8, 4 x 4
Inter prediction	P: 1 frame B: 2 frames (average only)	P: n frame to 1 frame B: n + m frames to 2 frames (Implicit, Explicit, Mean = default)
Intra prediction	DC + AC (partial)	Intra16 x 16: 4 types Intra 4 x 4 : 9 types
Quadrature conversion	8 x 8 DCT (with calculation errors)	4 x 4 pseudo-DCT (without calculation errors)
Quantization	Linear transformation	Exponential transformation
Variable length coding	VLC	CAVLC, CABAC
D-mode	Temporal only	Temporal, spatial
Interlace	Frame/Field prediction	MBAFF
Deblock filter	Non-standard	Standardized loop-filter

details, however, resulting in a vast overall improvement in the coding efficiency over the MPEG4 or the H.263.

The main technologies that have been improved for the MPEG4-AVC/H.264, over the MPEG4-ASP, are enumerated in Table 1. Although the overall tendencies have been inherited from the MPEG4 technologies of the past, quite a good number of ingenious devices have been implemented, when observed in detail, as shown in Table 1. Take the block division method, for example, only two types of 16x16 and 8x8 are available with the MPEG4-ASP. Seven block types and combinations of up to 259 types are available with the MPEG4-AVC/H.264, making it possible to select the most appropriate form to accommodate the characteristics of the images. This is expected to lead to a dramatic improvement in the predicting accuracy. There is a drawback, however, in that it requires an immense number of calculations in order to make a selection for the appropriate form of block. Further, up until now reference could be made of only one past frame, for predictions made between frames, however, with the MPEG4-AVC/H.264 it is now possible to select a single arbitrary frame from multiple frames for reference. This results in a marked improvement in the flexibility of the selections. Furthermore, a specific frame can be stored over a long period, making it possible to make a long-term prediction.

For bidirectional predictions, a marked improvement in accuracy can be expected, because the implicit weighted average that is proportional to time, other than the simple averaging of forward and backward reference images, as well as the explicit weighted average that can be specified externally, are made available. It is now possible to expect improvements in the coding efficiency particularly for the effects on images such as fading in, fading out and cross fading, for which effective processing was not possible with existing coding technologies.

Further, although most of the current moving picture coding technologies use the Discrete Cosine Transform (DCT), it was not possible to avoid the nightmarish image quality deterioration due to the accumulation of mismatches arising from calculation errors of the inverse DCT. By using a pseudo DCT for the MPEG4-AVC/H.264, which does not have any mismatches, it has become possible to eliminate this problem. This has a very significant meaning particularly for those who are involved in the development of LSIs.

In the past the simple Huffman Table was used for variable length coding. With the MPEG4-AVC/H.264, however, more efficient coding is realized with the adoption of the CAVLC that selects an appropriate Huffman Table depending on the context (positional context) of the various coding coefficients, as well as the binary arithmetic coding (CABAC), which is more efficient than the Huffman Table and replaces its use. Due to the adoption of more sophisticated and more complex coding systems, however, the amount of calculations to be performed increased, creating an even greater difficulty for the practical use of this method.

Features of Oki Electric's CODEC for the MPEG4-AVC/H.264

Oki Electric has been aggressively engaged in the standardization activities of MPEG4 and its product commercialization. This resulted in the development and sales of our MediaServer, an integrated video distribution platform based on the MPEG2/4 standard. By utilizing our technologies that have been accumulated over the years we have now developed a real-time software encoder/decoder that realizes high performance levels for both image quality and processing speeds. Prior to commencing our development of the CODEC for the MPGEC4-AVC/H.264, we thoroughly researched the technical features of the MPEG4-AVC/H.264, how these technologies brought about advantageous effects and what their issues were. The following targets were established for that development, based on the results of the aforementioned research.

1) Realization of real-time encoding and decoding of D1 class images using personal computers readily available on the market.
2) Minimization of image quality deterioration due to higher speeds and increases in the coding amounts.
3) Realization of stable transmission routes that are both real-time and with fewer incidences of delays.
4) Realization of further improved image quality by resolving various issues of the current reference software.

(1) Increasing speed

Since the initial design stage at the beginning of the development, we have been pursuing the potential to increase speeds from the aspects of computer architecture, encoding algorithms and packaging technology, with a realization of real-time encoding by the software as our target. From the aspect of computer architecture, we made an effort to contrive various devices, such as the effective use of cache, pipelining, parallelizing, as well as the use of MMX and SSE. Further, from the aspect of encoding algorithms, efforts were made for the development of high-speed algorithms, in particular the high-speed algorithm for searching for motion, as well as the development of a parallelizing algorithm for maximizing the effectiveness of MMX and SSE. Furthermore, for the packaging technology we tried with the greatest possible care to reduce the number of copies made on the memory, to reduce to a minimum the number of calculations to be performed to have time consuming calculations, such as multiplications and divisions, performed by other high-speed calculations and also to inhibit the occurrence of pipeline stalling at the conditional branching for packaging. Through these many devices, we succeeded in realizing the real-time encoding of images that are equivalent to the television image quality (D1@30fps), using high-performance personal computers readily available on the market.

(2) Improving quality

At the same time as pursuing high speed operation, we also looked into achieving high image quality. High speed and high quality are not readily compatible, and until recently, virtually the whole industry was looking to achieving high speeds at the expense of image quality. However, at Oki, we have also fixed our sights on high quality, refusing to make any speed improvements which involve a pay off in terms of poorer images, and searching to the last to develop high-speed algorithms which maintain image quality. One such solution was to thoroughly resolve the various issues surrounding the current methods, in particular with reference software. Methods for resolving and improving such issues were then discovered. As a result, we have been able to create a method, which makes it possible to inhibit the deterioration of image quality or to maintain the increasing coding to extremely small amounts, even when real-time high-speed calculations are executed.

Further, in order to raise the image quality in a stable manner, we developed a proprietary feed forward and feedback combined-type rate control algorithm that raises the quality of images while allowing real-time execution, with encoding delays inhibited as much as possible.

(3) Fewer delays and stabilized transmission rates

When considering bidirectional communications, it is inadequate to simply consider high speed and high quality. It is essential to consider a fewer incidence of delays with stabilized transmission rates. We addressed our concerns on lowering the incidence of delays as well. Buffering was reduced to an utmost minimum, while a reduction in the delay incidences, the stabilization of transmission rates and stabilization of image quality, were sought through a proprietary control technology. The current method used for streaming over the internet requires a long buffering time before images can be

reproduced. This makes only one direction available for communications. The responsiveness for operations, such as the switching of channels, was extremely poor as well. We are certain that our CODEC on the other hand, will cause fewer incidences of delays and offer stable transmission rates to provide superior performances for the applications of bidirectional communications, when real-time is essential for broadcasting and video on demand (VOD) distributions that require operation responsiveness.

Conclusion

The MPEG4-AVC/H.264 is an extremely important international standard not only for high-quality motion picture distribution services, on rapidly developing broadband networks, but also for mobile broadcasting and the next-generation DVDs. We believe that it will continue to draw a lot more attention in the future. We are certain that the CODEC for MPEG4-AVC/H.264, developed by Oki Electric, is an extremely attractive product that realizes an image quality exceeding the high image quality fundamentally available with this standard and offers both increased speeds and fewer incidences of delays. Further, we intend to apply the know-how accumulated through the development of this CODEC to DSP/LSI, to continue our efforts by developing an even more attractive MPEG4-AVC/H.264 product lineup.

References

1) ISO/IEC 14496-10:2004, Advanced Video Coding (Second Edition), March 2004.
2) Zhixiong Wu: "Multimedia Streaming Technology for Broadband Part 4: Development of MPEG4-ASP Coding Technology for High Quality Images", Oki Technical Review, pp. 60-63, Issue 192, Vol. 69, No. 4, October 2002.

Authors

Zhixiong Wu: Broadband Media Company, Engineering Development Div.

Takaaki Hatanaka: Broadband Media Company, Engineering Development Div., Team-1

Masayuki Tokumitsu: Corporate Research and Development Center, Ubiquitous Systems Laboratory

TIPS Basic Glossary

MPEG4-AVC/H.264
The MPEG4-AVC/H.264 is an international standard of a motion picture coding method that has been jointly formulated by ISO and ITU-T.

CODEC
CODEC is a generic designation for the encoder and decoder.

D1 image
The D1 image is an image equivalent in quality to the current television broadcast images, having a size of 720 (width) x 486 (height).

MPEG4-ASP
The MPEG4-ASP (Advanced Simple Profile) is one of the profiles provided by the MPEG4 moving picture coding method formulated by ISO.

MMX
The MMX is a processor for multimedia processing, which is built into Pentium CPUs manufactured by Intel. Although various higher speeds and parallel processing becomes possible using this processor, it will also be necessary to create programs at the assembly language level with dedicated command sets.

SSE
The SSE is an expanded version of the MMX, which supports much higher speeds and parallel processing.