Multimedia Streaming Technology in Broadband Networks
4 – Development of High-Quality Picture Encoding Technology MPEG4-ASP

Zhixoing Wu

In recent years, the expansion of broadband networks, such as xDSL, FTTH, high-speed wireless LAN, and so on, has progressed with astonishing speed. The numbers of broadband network subscribers have increased at a similarly dramatic rate, and we are starting to witness the real dawn of the broadband age. Many consider the existence of killer applications and killer contents to be crucial in this broadband age, and the expectations of multimedia applications, and in particular, high-quality moving pictures, are extremely high.

In order to respond to these expectations, we embarked at an early stage on the development of a video streaming server for broadband networks, the Oki MediaServer*1), and on a MPEG4 codec for supporting the operation of the server. In the early part of this year (2002), we released the Oki MediaServer V5 incorporating a MPEG4-ASP codec (described below). This article starts with an overview of MPEG, the international standard for moving picture encoding, and goes on to give a technological summary of Oki's high-quality picture MPEG4-ASP codec and its features.

Overview of MPEG

"MPEG" is an acronym for the Moving Picture Expert Group, which operates under the umbrella of the ISO (International Standard Organization) and has established a number of international standards relating to the encoding of moving pictures. These international standards are often denominated by appending a number to the group acronym, for example, MPEG1, which is used widely in VCD, karaoke, etc., MPEG2 used for DVD and digital broadcasting, and MP3, the world-famous standard for Internet music delivery.

MPEG4 is the latest international standard established by the MPEG group in relation to moving picture encoding. Work on MPEG4 started in 1994, but in the initial stages, it was aimed at achieving ultra-low bit rates or ultra-high compression rates. A typical application of this kind is mobile communication based on IMT2000, or the like. However, from 1998 onwards, the emergence of broadband has led to ever increasing demand for high-quality and multi-function moving pictures. In response to these disparate market needs, MPEG4 was adapted to specify a number of profiles (technology groups).[1)] In these, the Advanced Simple Profile (ASP) was established as a profile aimed at high-quality moving picture transmission over broadband networks.[2)] ASP received wide industry backing, and was adopted early on as a standard by the Internet Streaming Media Alliance (ISMA).

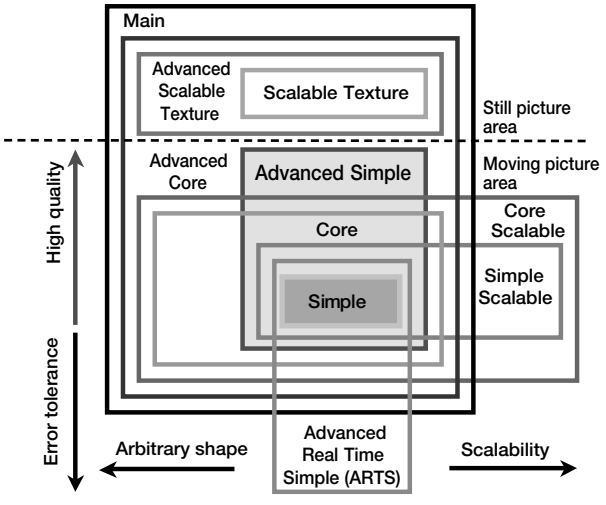

Fig. 1 Structure of MPEG4 standard

Compared to the first version of the MPEG4-SP (Simple Profile) used in next-generation portable terminals and narrowband Internet, etc., MPEG4-ASP has various extra tools which contribute to improved encoding efficiency (see Table 1). It also incorporates various measures for enhancing encoding efficiency compared to MPEG2 used in digital broadcasting, DVD, and the like (see Table 2). By means of these new tools, MPEG4-ASP has achieved greatly increased encoding efficiency, and in assessments both by the MPEG and other expert groups, it has been confirmed to give 1.5

Table 1 Features of MPEG4-ASP (vs. MPEG4-SP)

Name of tool	Function
MPEG Q	MPEG quantization method
Q-pel	Motion compensation with 1/4 picture element accuracy
B-VOP	Bi-directional prediction
Interlace	Compatibility with interlaced input pictures
GMC	Total picture area motion compensation

*1) OKI MediaServer is a registered trademark of a generic video delivery system made by Oki Electric Industry Co., Ltd.

Table 2 Features of MPEG4-ASP (vs. MPEG2)

Name of tool	Function
4MV	Prediction by 4 MVs for 1 MB
UMV	Allows MV search outside picture area
Q-pel	Motion compensation with 1/4 picture element accuracy
D-mode	Efficient use of MV in bi-directional prediction
AC Pred.	AC component prediction
3D VLC	Three-dimensional - variable length encoding
GMC	Total picture area motion compensation
Error Resilience	Packet segmentation, data segmentation, reversible variable length encoding

to 2 times the performance of MPEG4-SP and MPEG2.[3] Below, the technical features of the MPEG4-ASP are described, along with the features of Oki's MPEG4-ASP codec.

Technical features of the MPEG4-ASP

Fig. 2 shows the structure of an MPEG4-ASP encoder [1) 2)].

In this diagram, when a picture is input, firstly, the movement between that picture and a previously encoded and decoded reference picture is investigated (motion detection), to derive a (motion compensated) differential picture which compensates for this movement. This differential picture is then converted by discrete cosine transform (DCT), quantized, variable-length encoded together with other related side information, and output as a bit stream (compressed information). Simultaneously with this process, the quantized differential picture undergoes inverse quantization, and discrete cosine inversion, and is complemented by means of the motion compensated reference picture, and then stored in the memory as the next reference picture. By taking the differential between motion compensated frames, redundant information between pictures is removed, and hence the information volume is compressed. Moreover, redundancy

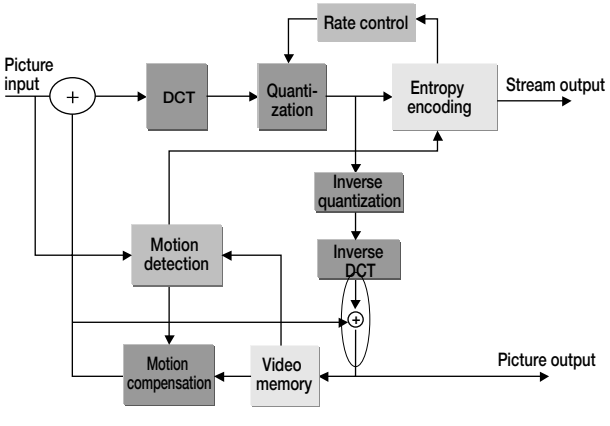

Fig. 2 Structure of MPEG4-ASP

(correlation) of information within the picture frame is eliminated by performing DCT on the differential picture, the information is compressed by quantization, and this information is then further compressed by variable length encoding. By implementing this series of information volume compression techniques both between frames and within frames, a very high rate of compression is achieved. MPEG4-ASP has the following particular advantages in terms of high compression and high quality. Firstly, it comprises various motion compensating techniques, one of which, the 1/4 picture element accuracy motion compensating technique, contributes hugely to improved accuracy in motion search and motion compensation in high-resolution pictures, and hence brings enhanced encoding efficiency. Other features, such as direct mode in bi-directional prediction, and frame/field adaptive prediction and frame/field adaptive DCT for interlaced images, also help to boost encoding efficiency and picture quality enormously. Furthermore, the adoption of MPEG quantization which takes account of human audio-visual characteristics, as well as the use of three-dimensional variable length encoding (3D VLC), make it possible to improve picture quality at the same compression ratio.

Features of the Oki MPEG4-ASP codec

Oki has been involved in standardizing the MPEG4 system right from the very start, and whilst working actively on the company's proposals and choices in this area, we have been building up technology and investigating its possible application to products. We have focussed on the following objectives in our development of a MPEG4 codec product.

(1) Pursuing high-speed operation

The power of the initial development PC did not readily permit real time encoding of moving pictures. In order to achieve software-based real time encoding, we looked at the possibility of increasing operational speed, from the viewpoints of computer architecture, encoding algorithms, and mounting technology. From the viewpoint of computer architecture, we attempted various measures, such as efficient utilization, pipe lining, and parallel operation of the cache, use of mmX*[2)], SSE, and the like. With respect to the encoding algorithm, we concentrated on the development of high-speed algorithms, in particular high-speed algorithms for motion searching, as well as parallel algorithms which allow mmX and SSE technology to be used to maximum effect. We approached mounting technology very carefully, looking to reduce memory copies and the frequency of operation as far as possible, substituting time-consuming calculations such as multiplication and division with other faster calculations, and preventing pipe line breakdowns by condition branches. By adopting these measures, we were able to achieve real-

*2) Pentium and mmX are registered trademarks or trademarks of Intel Corp.

Fig. 3 Delivery example using MPEG4-ASP (QHD = 960 x 540 x 30 fps @ 4 Mbps)

time encoding at an early stage, and we have maintained our excellence in terms of high-speed operation. At present, we have achieved real-time encoding of TV-quality (VGA@30fps) pictures and real-time decoding and playback of QHD (high vision 1/4 size and 960 x 540 pixel) images. Fig. 3 shows one frame of a QHD picture of average 4 Mbps@30fps. Next, we looked for high image quality.

(2) Pursuing high quality

At the same time as pursuing high speed operation, we also looked into achieving high picture 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 picture 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 pictures, and searching to the last to develop high-speed algorithms which maintain picture quality. In order to improve picture quality in a stable manner, we developed our own combined usage type rate control algorithm for feed forward and feedback, to improve picture quality in real time, with minimum encoding delay. We also looked into DCT/IDCT operation accuracy, to achieve extremely fast operation and accuracy far in excess of IEEE standards. In this way, we were able to suppress degradation in picture quality due to IDCT mismatching in high-quality pictures, to extremely low levels. Moreover, in order to guarantee stability in picture quality at low bit rates, a contour storage type pre-filter was developed. This initiative bore fruit, and the picture quality we have currently achieved is very highly rated.

(3) Pursuing low delay

When considering bi-directional communications, high-speed operation and high quality are not in themselves enough – another crucial factor is low delay. In tackling the issue of low delay, we sought to reduce buffering to a minimum, aiming at a balance between stable picture quality and low delay through the use of control technology. In current LAN environments, the overall delay from picture input, through transmission to the encoder/server and distribution from the server to the recipient, to final reception, decoding and display, is less than 200 msec. Windows Media*3), Real Media*4) and other applications used in Internet streaming require long buffering times before image reproduction starts and can only be used for one-way communications, whereas Oki's low-delay system provides a real competitive advantage in the field of bi-directional communications, and the like, which demand real-time performance.

(4) Pursuing expandability

*3) Windows Media, Windows and Windows CE are registered trademarks or trademarks of Microsoft Corp. *4) Real Media is a trademark of Real Networks.

As Fig. 1 indicates, with the host of functions that it comprises, MPEG4 is a very attractive encoding system for many different applications. Accordingly, there is a high requirement for functional expandability. What is more, time is passing very quickly in this field, and one generation of technology is rapidly replaced by the next. In order to respond quickly to this situation of functional expansion and rapid change, it is extremely important to have software flexibility. In the MPEG4 codec developed by Oki, we have paid careful attention to layering and modularization from the first design stages, and by incorporating maximum independency and shielding properties between layers and between modules, we have been able to improve software flexibility, as well as its capacity to be expanded and repaired. Thanks to this high level of shielding, it is possible to expand functions without hardly affecting the original source code, thus helping greatly to shorten the design stage and improve reliability.

(5) Multi-platform compatibility

The MPEG4-ASP codec belongs to the third-generation of technological development, but care has been taken to make it compatible with multiple types of platforms, so that it can incorporate object-oriented design from second-generation technology, as well as achieving maximum performance at all times with different types of CPU, and allowing easy implantation in different operating systems. In terms of CPU compatibility, the codec has been designed so that it always gets the best performance out of any CPU, from Non-mmX through to Pentium 2, 3, 4*[2], by dynamic operation. Adoption of object-oriented design means that the codec can easily support various OS, such as Windows CE*[3], Windows*[3], Linux*[5], and so on, and as well as being expanded to non-PC devices, such as PDAs, set-top boxes and special terminals.

Conclusion

The MPEG4-ASP is an international standard for encoding of high-quality moving pictures aimed at video streaming services for broadband networks, which are expanding at breakneck speed. MPEG4-ASP is predicted to be a focus of increasing interest in the future, and the MPEG4-ASP codec developed by Oki achieves picture quality superior to the already high quality provided by the standards, hand in hand with high-speed operation and reduced delay. We aim to continue development of the codec into a wide range of products, making full use of features such as its easy expandability and multi-platform compatibility. Finally, we plan to go on enhancing introduction of the codec into next-generation encoding systems, in order to maintain lasting excellence in this field. ◆◆

References

1) ISO/IEC JTC1/SC 29/WG11 N3904, Coding of audio-visual objects Part 2: visual, FDAM 4: Streaming video profile, Jan. 2001

2) ISO/IEC 14496-2:1999/Amd. 1:2000 (E), Coding of audio-visual objects Part 2: visual, AMD 1: Visual extensions, Jan. 2001

3) ISO/IEC JTC1/SC 29/WG11 M4698, Report of the ad hoc group on MPEG-4 video verification tests, Jul. 1999

Author

Zhixiong Wu Broadband Media Company, R & D Div.

TIPS Basic Glossary of Terms

mmX
"mmX" is a multimedia processor built into Intel's Pentium CPU. Using this processor, various high-speed and parallel operations can be performed, but a dedicated language is required for programming.

SSE
"SSE" is an extension version of mmX and supports a greater number of high-speed, parallel operations.

VGA
This indicates the screen display size (or resolution) on a PC, etc. The image size in picture elements is 640 (horizontal) x 480 (vertical).

QHD
This indicates an image size that is 1/4 the size of a high-resolution image (High Definition e.g: High-vision image = 1920 x 1080).

*5) Linux is a registered trademark or trademark of Mr. Linus Torvalds, in the United States and other countries.