World's Highest-performance HEVC Software Coding Engine

Kazuyuki Iso, Masaki Kitahara, Ryuichi Tanida, Tokinobu Mitasaki, Naoki Ono, and Atsushi Shimizu

Abstract

In recent years, worldwide attention has been focused on the distribution of 4K and 8K high-definition video. Video services for mobile terminals have also been expanding rapidly, and higher compressive power in video coding is expected. In response to this situation, the international standard HEVC (High Efficiency Video Coding, otherwise known as H.265/MPEG-H) was established in April 2013. We introduce here a software coding engine developed by the NTT Media Intelligence Laboratories that conforms to HEVC and features the world's highest coding performance.

Keywords: H.265/HEVC, video coding, video distribution

1. Introduction

Attention has been focused worldwide in recent years on high-speed distribution of 4K and 8K high-definition (HD) video. Transmission of the huge amount of data associated with 4K and 8K video while maintaining sufficient quality requires the use of video coding techniques that have high compression performance. Video services for smartphones have gained in popularity in mobile communication environments in which capacity has expanded due to the increased use of LTE (Long Term Evolution) or other high-speed communication technology. Many video distribution service providers are implementing mobile services, and telecommunications carriers have strong expectations for higher compression of motion picture and video content in order to cope with the rapidly increasing communication traffic. In response to this situation, the NTT Media Intelligence Laboratories developed a software video coding engine that has the world's highest level of video compression and conforms to the Main/Main10 Profile of HEVC (High Efficiency Video Coding). HEVC is said to double the video data compression performance of the previous international standard, H.264/MPEG-4 AVC (Advanced Video Coding), also called H.264. Our software coding engine improves com-pression performance by a factor of up to 2.5, which enables a reduction in video data of up to 60% compared to H.264 with equivalent image quality.

2. HEVC software coding engine

NTT Media Intelligence Laboratories has developed a technique that takes image features into account to efficiently distribute the coding rate over the image, and in doing so, has achieved the world's highest performance HEVC software coding engine. This technique may eventually be applied to Internet protocol television (IPTV) and other such services. Therefore, we also developed a bit rate control technique that achieves stable video distribution, and a fast technique to compress video in a practical amount of time. To enable viewers to have a pleasant experience when they view video provided by IPTV and similar services, a function for adjusting the amount of data flowing through the network is necessary to achieve efficient video transmission. However, no such function is included in the international standards. We therefore developed a proprietary technique for controlling the coding rate in order to allocate the coding rate efficiently on the basis of statistical data.

Without fast compression techniques, HEVC

Fig. 1.   Relation of image properties to subjective image quality.

involves trading high computational load and long processing time for compression performance. The HEVC software coding engine that we developed uses a proprietary fast compression technique to compress full HD video in about five times the playback time; in other words, if the playback time is 1 hour, this technique can compress the full HD video in 5 hours. This factor of five is generally considered a target for a practical processing time when working with actual content, so this software coding engine features both high compression performance and fast processing. With a steady eye on future expansion of the market for video distribution services, we are also targeting the Main 10 Profile for 4K and 8K high definition video. In the following sections, we describe the techniques developed by the NTT Media Intelligence Laboratories in more detail.

3.   Local quantization parameter (QP) variation processing for high compression performance

The allocation of a limited coding rate is important for improving image quality. A known characteristic of human vision is that distortion is difficult to detect in regions of video that contain complex patterns. However, there are no complex patterns in the region of a human face in a single image, as shown in the example in **Fig. 1**, and distortion introduced by coding is therefore noticeable in such regions. On the contrary, the background region in the example image contains foliage that has complex patterns due to the many color variations in the tree trunks and leaves, so there is a tendency for coding distortion to go unnoticed. Consequently, techniques that make use of this tendency are often applied to measure the pattern complexity within an image. Then the coding rate is reduced in regions that include complex patterns in order to improve the overall coding efficiency.

We also focused on the relation between regions with complex patterns, peripheral areas, and temporal changes and discovered that the loss of subjective image quality is barely detectable in peripheral areas that have both complex patterns and large temporal changes, even if the coding rate is further reduced. We therefore developed a proprietary local QP variation processing technique in which the coding rate is adjusted for each region in an image. Specifically, the image is analyzed to detect regions that contain complex patterns before the coding process is executed. Controlling the coding rate according to the relation between peripheral areas and the degree of temporal change in regions where the patterns are complex improves the overall video compression performance.

4.   Bit rate control for stable video transmission

IPTV and other video distribution services require a control for bit rate stability as well as high compression performance. Bit rate control is a fine control
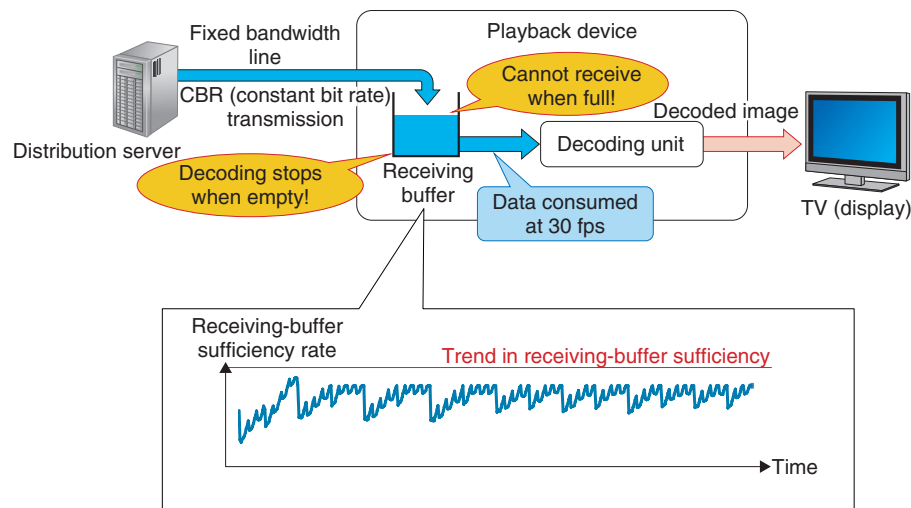
Fig. 2.   Coding rate control.

that allocates a bit rate to each video frame and each region of a frame so as to distribute video at a constant bit rate (**Fig. 2**). NTT Media Intelligence Laboratories has developed this technique over a long period of time while developing the H.264 coding engine and other such technologies. Specifically, the coding rate of a previously coded image is divided into a rate for header data and for video data to achieve separate statistical processing. The results of the statistical processing can be used to estimate with high accuracy the coding rate for the subsequent image. HEVC uses a complex hierarchical reference structure that involves random access of more frames than in H.264, so the compression performance is higher. Therefore, performing hierarchical statistical processing can be used to control the coding rate for HEVC as well.

5.   Video coding in practical time

In video coding, each image is divided into square blocks. Higher compression performance can be achieved in HEVC than in H.264 by selecting the block size from among four candidates. Additionally, many compression modes are prepared. To select the optimum block size and compression mode for each block, it is necessary to compare all combinations of block sizes and compression modes. However, simply comparing all the combinations requires a huge amount of computation time, so we developed a method that achieves fast processing while maintaining a higher compression rate by analyzing regions of
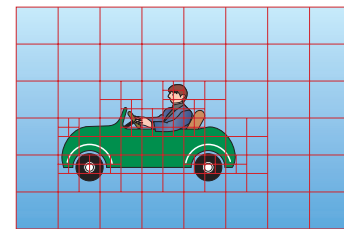

Fig. 3.   Division of image into blocks.

high similarity and regions of local complexity in images.

A high compression rate can be achieved in HEVC by selecting a small block size for regions with complex video patterns and a large block size for regions with simple patterns (**Fig. 3**). Analysis of the results of block partitioning for different types of video reveals that the reduction in compression rate is slight even when the block sizes are fairly similar between adjacent blocks, because image features are similar for consecutive regions on the screen, even when the image is partitioned. We developed a method in which that information can be used to shorten the computation time by considering the block sizes of adjacent regions that have already been compressed (with compression being done sequentially from the top left of the screen to the bottom right) to reduce the number of block size candidates (**Fig. 4**).

Also, one important compression mode for improving compression performance is the skip mode. The
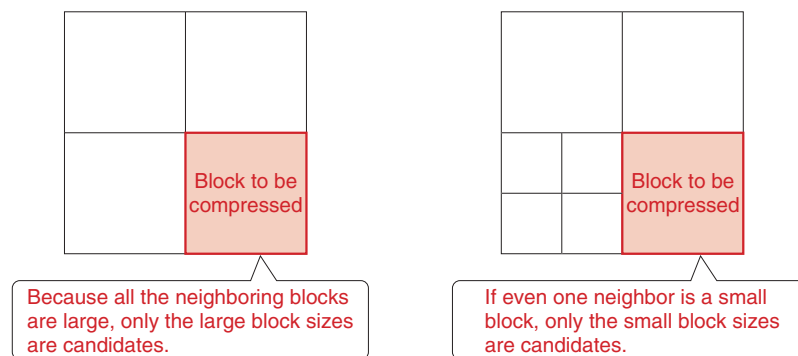
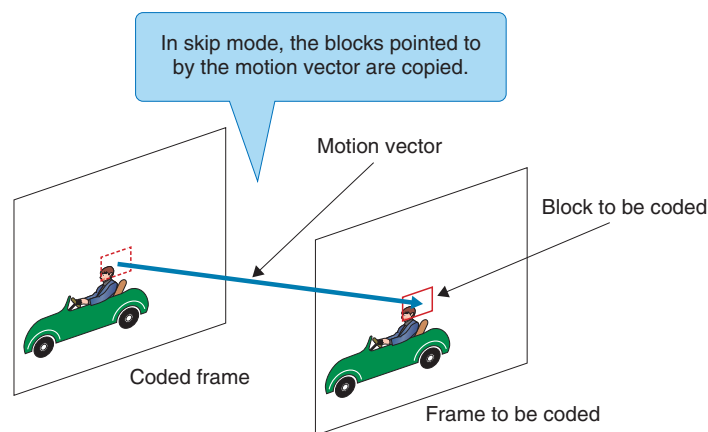Fig. 4.   Selection of candidates for block subdivision.

Fig. 5.   Skip mode of video compression.

concept of the skip mode is illustrated in **Fig. 5**. In the skip mode, image blocks that are selected according to a motion vector that represents the movement of the subject within the image are coded by copying previous blocks. The skip mode has a wide application range in HEVC, so compression performance can be increased by compressing most of the image using this mode. Therefore, if we can determine that the skip mode is appropriate by estimating the image quality when using the skip mode, we can eliminate the coding processing for trials of other compression modes and thus greatly reduce the computation time. Specifically, we compare the motion vector of the block to be coded with the motion vector of the adjacent block that has already been coded to determine if the two blocks contain the same subject. If they do, we compare the image quality obtained by using the skip mode to the image quality of the adjacent block.

If it is possible to achieve the same image quality, the skip mode is judged to be appropriate. That makes it unnecessary to try multiple compression modes and greatly reduces the processing time (**Fig. 6**).

Next, for compression modes other than the skip mode, a linear transform* is applied to the difference between this image and the previous image (predicted difference) to convert it to data that are easily compressed, and then the coding is performed. Many linear transformation methods are possible in HEVC, and comparing all of them would take a huge amount of computation time. We therefore conducted experiments using various types of video to determine the

---

*   Linear transformation: A process applied to the difference data for the block to be compressed and the previous image block to reduce the amount of data involved in video compression. A typical conversion process is the discrete cosine transform.
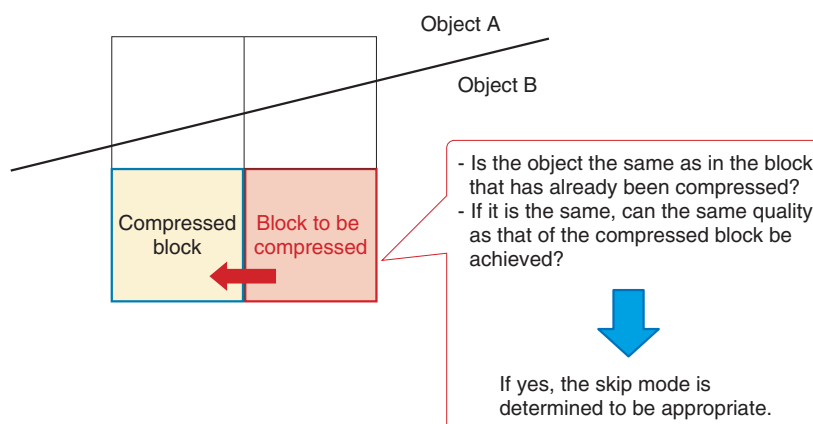
Fig. 6.   Skip mode decision.

optimum linear transformation method. The results indicated there was a constant correlation of the image complexity of the regions to be coded and the block size distribution, so we can reduce the computation time by selecting the linear transformation method according to the image complexity and the block partitioning situation.

6. Future development

Our software coding engine has been commercialized by NTT Advanced Technologies and is being sold as a codec software development kit that includes the encoder and decoder as a set. The decoding engine for playing a stream generated by this coding engine was developed independently by NTT Advanced Technologies.

We intend to continue developing high image quality video distribution services that require even higher video compression performance. The NTT Media Intelligence Laboratories will push forward with research and development to contribute to the further expansion of the 4K and 8K video distribution market.

Kazuyuki Iso

Senior Research Engineer, Visual Media Coding Group, Visual Media Project, NTT Media Intelligence Laboratories.

He received the B.E. degree in electronics from Gunma National College of Technology in 1998, and the M.S. degree in knowledge science from Japan Advanced Institute of Science and Technology, Ishikawa, in 2000. He joined NTT in 2000 and engaged in research on visual communication systems. He is a member of the Institute of Electrical and Electronics Engineers (IEEE), the Virtual Reality Society of Japan (VRSJ), and the Information Processing Society of Japan (IPSJ).

Masaki Kitahara

Senior Research Engineer, Visual Media Coding Group, Visual Media Project, NTT Media Intelligence Laboratories.

He received the B.E. and M.E. degrees in industrial and management engineering from Waseda University, Tokyo, in 1999 and 2001, respectively. After joining NTT in 2001, he engaged in the development of compression and rendering algorithms for 3D video. He moved to NTT Advanced Technologies (NTT-AT) and managed various software development projects. In 2009, he returned to NTT and has been researching and developing HEVC encoding technologies. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).

Ryuichi Tanida

Research Engineer, Visual Media Coding Group, Visual Media Project, NTT Media Intelligence Laboratories.

He received the B.E. and M.E. degrees in electrical engineering from Kyoto University in 2001 and 2003, respectively. He joined NTT to study video coding for developing a real-time H.264 software encoder application. He moved to NTT-AT in 2009 to manage various software development projects related to video compression, distribution, and TV phones. He has been at NTT Media Intelligence Laboratories since 2012, as a research engineer in the visual media coding group. His current research focuses on the development of HEVC software encoders. He is a member of IEICE.

Tokinobu Mitasaki

Research Engineer, Visual Media Coding Group, Visual Media Project, NTT Media Intelligence Laboratories.

He received the B.E. and M.E. degrees in electrical engineering from Hokkaido University in 2001 and 2003, respectively. After joining NTT in 2003, he engaged in the development of holographic memory and video compression. He moved to NTT Communications in 2008 to develop Digital TV relay service. He has been at NTT Media Intelligence Laboratories since 2012, as a research engineer in the visual media coding group. His current research focuses on the development of HEVC software encoders. He is a member of IEICE.

Naoki Ono

Senior Research Engineer, Visual Media Coding Group, Visual Media Project, NTT Media Intelligence Laboratories.

He received the B.E. and M.E. degrees in information engineering from Niigata University in 1989 and 1991 respectively. He joined NTT in 1991 and has been engaged in video compression algorithm and software development. He is a member of IEICE and the Institute of Image Electronics Engineers of Japan (IIEEJ).

Atsushi Shimizu

Senior Research Engineer, Supervisor, Visual Media Coding Group, Visual Media Project, NTT Media Intelligence Laboratories.

He received the B.E. and M.E. degrees in electronic engineering from Nihon University, Tokyo, in 1990 and 1992, respectively. Since joining NTT in 1992, he has been engaged in video compression algorithm and software development. He is a member of IEICE, IIEEJ, and the Institute of Image Information and Television Engineers.