Development of Ultra-low Latency Codec

IKEDA Toshiyuki, SHINBO Gouhei, KUROSAWA Naoki, AOKI Hirofumi, IURA Toshiyuki

Abstract

The codec used for transmitting video material is, in principle, subject to various operational limitations as a result of its encoding latency, so there was a need for lower latency encoding. In this paper, NEC introduces a newly developed codec device based on the latest H.264/MPEG-4 AVC video encoding standard to achieve minimum 10 ms codec latency while maintaining broadcast quality.

Keywords

ultra-low latency codec, high definition, H.264

1. Introduction

On July 24, 2011, Japan reached a milestone with the migration to digital terrestrial broadcasting. Digital broadcasts are characterized by the fact that they are digital and they offer high definition. This holds true not only for outgoing broadcasts, but also for transmission of video material, and is spreading at a rapid pace. In this environment where the amount of information has grown to levels unimaginable during the analog era, a codec offering efficient transmission of enormous amounts of data has become a critical component.

From the earliest days of compression technology, we at NEC have developed and introduced into the market an MPEG-2 format codec, and more recently, a codec for H.264/MPEG-4 AVC format which is said to offer compression efficiency surpassing that of MPEG-2.

This codec is a leading-edge technology that takes a 1.5 Gbps band HDTV (High Definition Television) signal and compresses it to approximately 1/25th to 1/250th the original data volume, without sacrificing original quality. At the same time, by nature of its processing method, latency is unavoidable. In particular, the codec used for video transmission was subject to various operational limitations due to its application.

In this paper, we will discuss our latency reducing technology, and introduce our codec device which successfully implements that technology to achieve an extremely low latency of 10 ms.

2. Achieving the Low Latency Codec

In the realm of video material transmission codecs, there is a strong demand for low latency in addition to high quality videos. There are numerous usage cases where lag due to codec performance cannot be tolerated, such as when communicating via live feed, or when switching between wired and wireless cameras while broadcasting live golf or other sporting events. What's more, codec lag caused low responsiveness when remotely operating field cameras which need to provide real-time response. Users are thus demanding usability similar to non-compressed transmission (**Fig. 1**). In response to such requests for lower latency, we totally reexamined everything from encoding hardware architecture to encoding method and embedded algorithm.

With conventional encoder hardware architecture, a single video frame is set as the unit for processing; cumulative and sequential type processing is applied. By its nature, a fixed lag of n x frames (n is positive integer) occurred. To rectify this, we developed a new encoding LSI in which on-the-fly processing is made possible within each frame by setting the minimum encoding unit level as the processing unit. As a result, the lag resulting inherently by the hardware architecture was reduced to the level of just a few ms.

In addition to the abovementioned latency, there is also latency dependent on the encoding standard of the codec. It is said that there is a trade-off relationship between this latency and the quality of video. That's because key factors for video quality within the same standard all contribute to latency, including motion prediction and encoding architecture, as well as encoding assignment optimization, because processing requirements can increase according to the precision and combination of factors, and depending on the selected mode, may necessitate the changing of frame order. And since encoded data is sent through the smoothing buffer as predetermined by the video encoding standard, the buffer transit time also contributes to the lag. The smoothing buffer is meant to absorb the time fluctuation of the generated signal, so the buffer transit

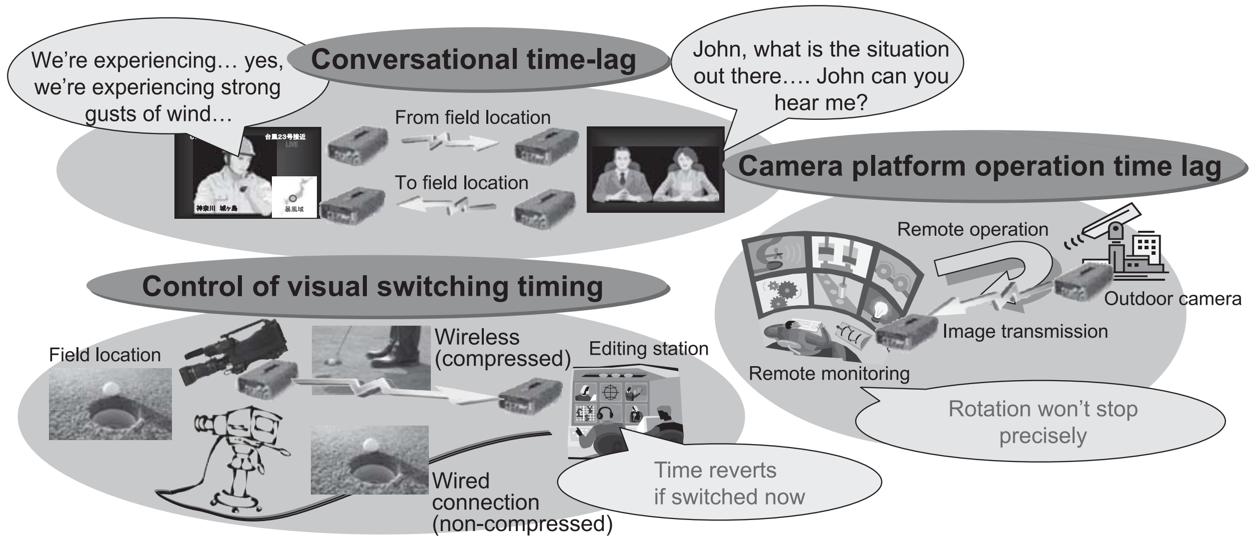


Fig. 1 Problems caused by codec latency.

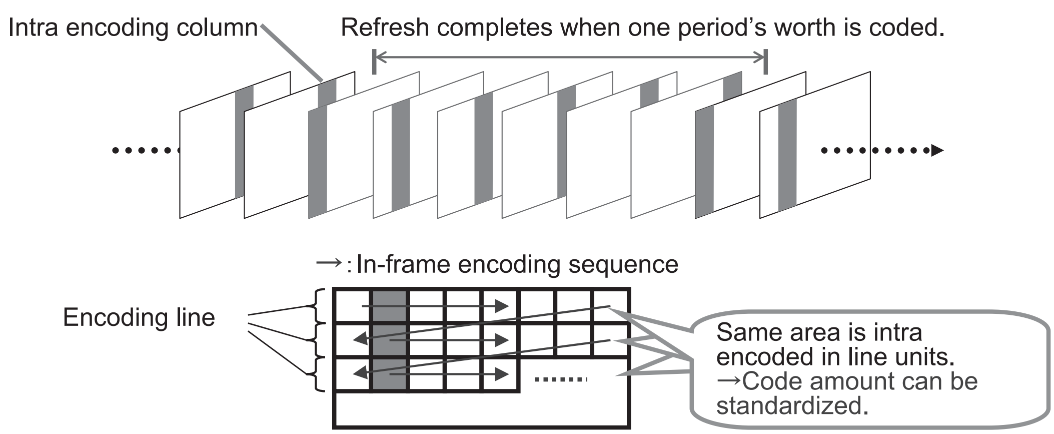


Fig. 2 Intra-column refresh method.

time is proportional to the fluctuation range. Our new development adopts H.264/MPEG-4 AVC, the latest video encoding method, so it supports not only 4:2:2 format (High 4:2:2 profile) which is typically required for video transmission, but also supports high bit rate (over 50 Mbps level). As for the encoding architecture, we adopted the intra column refresh method (**Fig. 2**) which allows equalization of the generated signal amount per line of encoding, from the standpoint of smoothing buffer latency control. These were the optimum choices within the international standards for encoding as of now, and although adoption of our own unique encoding method may have yielded low latency results, we decided against this as the compatibility disadvantages would far outweigh what we could achieve in reduced latency.

Video quality improvement, particularly the difficulty of achieving it while maintaining the low-latency performance

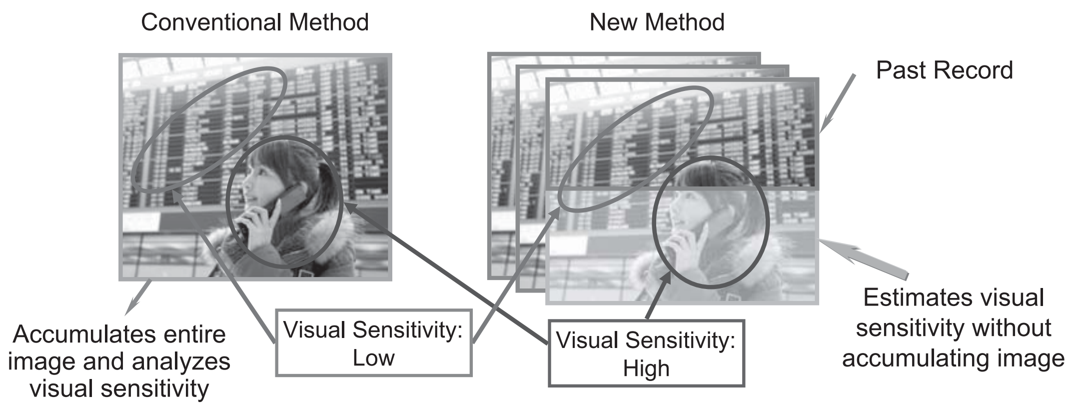


Fig. 3 Preferential video quality improvement for high visual sensitivity areas using original algorithm.

that is a prerequisite for our new development, is as we have already stated. In particular, preservation of video quality immediately before and after an abrupt scene change, is the biggest challenge. Typically with a scene change, the video frames before and after have no correlation with each other, making it necessary to allocate large amounts of code to the immediate vicinity. When code allocation is not adequately performed, the amount of code in the first half of the frame following the scene change becomes insufficient, causing a rapid deterioration in video quality. This method embodies an increase in generated code amount fluctuation over time, and as per its relationship with the smoothing buffer as explained earlier, it becomes impossible to reduce latency. In order to avoid this, it is necessary to allocate the right amount of code at the right time by detecting scene changes with high precision and minimum latency. To solve this problem, we have developed the highly precise complexity estimation feature based on NEC's original probability model. As a result, we were able to successfully suppress video quality deterioration by accurately detecting scene changes from just a few lines of video input data before encoding, and optimizing it by code amount control feedback at the encoding minimum unit level. In addition to this algorithm, we have also embedded a unique algorithm that preferentially improves video quality in high visual sensitivity areas to improve overall subjective video quality. This is done by estimating the spatial distribution of visual sensitivity depending on the local visual characteristics such as color, texture and movement (**Fig. 3**).

Through our efforts, we intend to develop an operation model that takes advantage of this codec's ability to boost efficiency of data transfer bandwidth, so that field locations transmitting videos can work without the stress felt with conventional codecs.

3. Outline of Ultra-low Latency Codec Device

The current NEC product lineup is as follows.

3.1 VC-7500/VD-7500 Board Codec

VC-7500 (**Photo1**) is a board encoder supporting H.264/MPEG-4 AVC, while VD-7500 (**Photo 2**) is a board decoder supporting both H.264/MPEG-4 AVC and MPEG-2. They provide all the basic functions necessary for HDTV (High Definition Television) and SDTV (Standard Definition Television) codec in a 180 mm × 120 mm sized circuit board. Power consumption is a low 15 W.

These board codecs were developed to be integrated into FPU (Field Pickup Unit) and such. Since the shape and interfaces offer compatibility with conventional MPEG-2 products VC-5510/VD-5510, any NEC FPU device with internal VC-5510/VD-5510 can be upgraded to H.264/MPEG-4 simply by replacing with these boards.

These board codecs include our original low latency function. This achieves a codec latency of 10 ms to 120 ms in ultra-low latency mode, and 300 ms to 700 ms in standard latency mode (not including latency of the transmission channel). The maximum video rate is 120 Mbps, making it possible to accommodate a wide range of video transmission requirements from high quality (high bit rate) to low bit rate. And from the

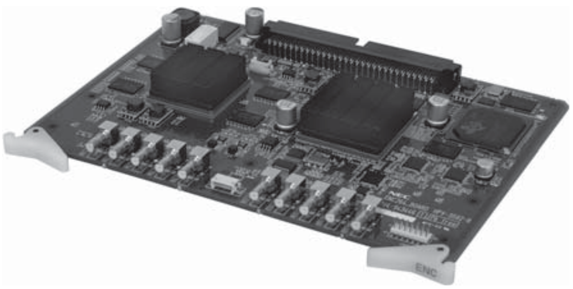


Photo 1 VC-7500 encoder board.

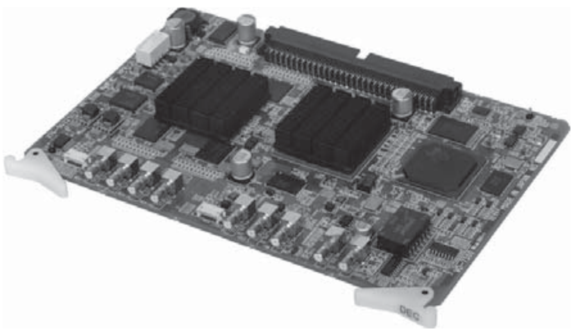


Photo 2 VD-7500 decoder board.

standpoint of video transmission, it also supports 4:2:2 format. As for audio encoding, MPEG-2 AAC, MPEG-1 Layer II, and non-compressed LPCM (SMPTE302M-2002) formats are selectable. Supplementary data transmission, including device ID monitoring data, time code, inter-station operation data, and text data, are also supported.

3.2 VC–7700/VD–7700 Half–Rack Codec

VC-7700 (**Photo 3**) is a 1U half-rack encoder supporting H.264/MPEG-4 AVC, while VD-7700 (**Photo 4**) is a 1U half-rack decoder supporting both H.264/MPEG-4 AVC and MPEG-2.

These can be used in a variety of applications including digital SNG, mobile units, field stations, and even for IP network transmission including FLET'S network and Internet. The codec function itself is identical to the VC-7500/VD-7500 board codecs we just introduced, but since the size and external connection ports offer compatibility with conventional MPEG-2 products VC-5700/VD-5700, system upgrading using these products can be done with superior cost efficiency. The IP interface uses a robust FEC for error correction,

Photo 3 VC-7700 encoder.

Photo 4 VD-7700 decoder.

so stable transmission is afforded even if the line's quality is less than ideal.

4. Conclusion

Through this paper, we introduced our H.264/MPEG-4 AVC ultra-low latency devices. We are confident that these devices will contribute to more efficient bandwidth usage and the provision of higher quality services, by their ability to deliver both high quality videos and an ultra-low latency of a minimum 10 mm, which were unattainable until now. However, since constant progress in technology will enable further improvements in functional performance, or since demands differ according to specific applications being undertaken, NEC will continue pursuing higher functional performance for its new and current products based on market demands.

In closing, we would like to express our gratitude to all those who contributed to the development of the codec device introduced in this paper.

*FLET'S is a registered trademark of Nippon Telegraph and Telephone East Corporation and Nippon Telegraph and Telephone West Corporation.

Authors' Profiles

IKEDA Toshiyuki
Manager
Broadcast and Video Equipment Division
Social Systems Operations Unit

SHINBO Gouhei
Engineering Expert
Broadcast and Video Equipment Division
Social Systems Operations Unit

KUROSAWA Naoki
Assistant Manager
Broadcast and Video Equipment Division
Social Systems Operations Unit

AOKI Hirofumi
Assistant Manager
Information and Media Processing Laboratories

IURA Toshiyuki
General Manager
Broadcast and Video Equipment Division
Social Systems Operations Unit

Information about the NEC Technical Journal

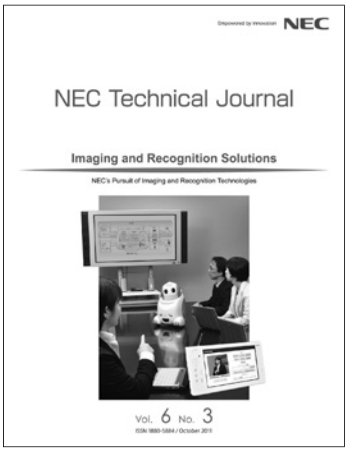
Thank you for reading the paper.
If you are interested in the NEC Technical Journal, you can also read other papers on our website.

Link to NEC Technical Journal website

Japanese

English

Vol.6 No.3 Imaging and Recognition Solutions



Vol.6 No.3

October, 2011

Special Issue TOP