Transmission and Functionality Test of MMT-Based Next-Generation Satellite Broadcasting System over “KIZUNA (WINDS)” Satellite

By Yuki KAWAMURA, Masaaki KOJIMA, Yoichi SUZUKI, Kazuhiro OTSUKI, Naoyoshi NAKAMURA, Takeshi KIMURA and Shoji TANAKA

NHK Science & Technology Research Laboratories, Tokyo, Japan

(Received July 31st, 2015)

In preparation for launching 8K (7,680 by 4,320 pixels) Super Hi-Vision (SHV) test broadcasting in 2016, a SHV satellite broadcasting system has been standardized by the Association of Radio Industries and Businesses (ARIB) in Japan. Because of the small roll-off factor of 0.03, 16APSK modulation scheme, and an LDPC coding rate of 7/9, the standard for the physical layer can achieve a transmission capacity of approximately 100 Mbps within the occupied bandwidth of a 34.5-MHz satellite transponder. The standard for the transport layer introduces MPEG-H MMT to realize harmonization and hybrid channel usage of broadcasting and broadband for content delivery. To evaluate the transmission performance of the physical layer and the functionality of the transport layer, we conduct a transmission test over “KIZUNA (WINDS)” satellite. The experimental results show that synchronized audio and video signals can be presented successfully even though they are delivered via asynchronous channels.

Key Words: Satellite Broadcasting, MPEG Media Transport, KIZUNA (WINDS)

Nomenclature

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>WINDS</td>
<td>Wideband InterNetworking engineering test and Demonstration Satellite</td>
</tr>
<tr>
<td>VSAT</td>
<td>Very Small Aperture Terminal</td>
</tr>
<tr>
<td>LET</td>
<td>Large Earth Terminal</td>
</tr>
<tr>
<td>UHDTV</td>
<td>Ultra High Definition TeleVision</td>
</tr>
<tr>
<td>HEVC</td>
<td>High Efficiency Video Coding</td>
</tr>
<tr>
<td>LDPC</td>
<td>Low Density Parity Check</td>
</tr>
<tr>
<td>APSK</td>
<td>Amplitude and Phase-Shift Keying</td>
</tr>
<tr>
<td>TS</td>
<td>Transport Stream</td>
</tr>
<tr>
<td>MMT</td>
<td>MPEG Media Transport</td>
</tr>
<tr>
<td>UTC</td>
<td>Coordinated Universal Time</td>
</tr>
</tbody>
</table>

1. Introduction

In preparation for launching 8K (7,680 by 4,320 pixels) Super Hi-Vision (SHV) test broadcasting in 2016, a SHV satellite broadcasting system has been standardized by the Association of Radio Industries and Businesses (ARIB) in Japan. Because of the small roll-off factor of 0.03, 16APSK modulation scheme, and an LDPC coding rate of 7/9, the standard for the physical layer can achieve a transmission capacity of approximately 100 Mbps within the occupied bandwidth of a 34.5-MHz satellite transponder. The standard for the transport layer introduces MPEG-H MMT to realize harmonization and hybrid channel usage of broadcasting and broadband for content delivery. To evaluate the transmission performance of the physical layer and the functionality of the transport layer, we conduct a transmission test over “KIZUNA (WINDS)” satellite using the prototype modulator, demodulator, MMT streamer, and MMT receiver.

2. Overview of the Standards

2.1. Physical layer

In the current satellite broadcasting system in Japan, a single 34.5-MHz bandwidth transponder has a capacity of approximately 52 Mbps. This conventional system is based on Trellis-Coded 8 Phase-Shift Keying (TC8PSK) with a coding rate of 2/3, and the modulated signal has a symbol rate of 28.86 Mbaud with a roll-off factor of 0.35. It is specified to deliver two HDTV services with one transponder. The current system cannot deliver even one 8K SHV service. The channel capacity required for 8K SHV broadcasting is a range of 80 to 100 Mbps according to the evaluation test results for HEVC, which is the most appropriate video coding system for UHDTV service delivery. We developed a new transmission system to improve transmission capacity considering the target channel capacity for 8K SHV broadcasting. This system complies with the allocated bandwidth of 34.5 MHz by the Radio Law of Japan by applying a smaller roll-off factor of 0.03 and symbol rate of 33.7561 Mbaud. The transmission capacity can reach approximately 100 Mbps in a satellite transponder using the 16APSK and an LDPC coding rate of 7/9. 32APSK schemes, whose transmission capacity is increased by 25% over the 16APSK, are also applicable. However, a larger receiving dish antenna (exceeding 45 cm in diameter) is required for stable reception of the 32APSK.

2.2. Transport layer

MPEG-2 TS is widely adopted in current digital broadcasting systems and was standardized in the early 1990s, at which broadband Internet Protocol (IP) networks were
uncommon. The TS was designed for video communication using Asynchronous Transfer Mode (ATM) networks, which are homogeneous and quality certified. However, broadband IP networks such as Fiber-To-the-Home (FTTH) and Long-Term Evolution (LTE) networks have become readily available and affordable. As a result, the potential for a hybrid delivery system that can deliver customized content using both broadband networks and broadcast channels is growing.

To realize such a system, we have proposed a new broadcasting system that employs MPEG-H MMT, which has been in development by ISO/IEC since 2009. The first version of MMT was approved in March 2014. MMT is designed for transport of media data in heterogeneous environments. The MMT standards specify encapsulation format for media data, IP delivery protocol, and signaling information for media delivery. In the MMT process, a presentation timestamp for media data is provided in UTC form so that a receiver can display various media data in a synchronized manner even if the delivery delays differ.

The ARIB standards include both constraints and extensions of the MMT international standards. Fig. 1 shows the layer model of the MMT-based broadcasting system. Audio, video, closed captions, and applications are coded by each coding system and encapsulated into transport packets. The encapsulation scheme of media data is shown in Fig. 2. A block of media data is encapsulated into a Media Fragment Unit (MFU), which is the minimum unit for delivery. To reduce the encapsulation latency, each MFU encapsulates one Network Abstraction Layer (NAL) unit for video coding, one Access Unit (AU) for audio coding, and one file for caption/application coding. Essentially, an MFU is carried in one MMT Protocol (MMTP) packet. If the size of an MFU is greater than the Maximum Transmission Unit (MTU) of the target network, the MFU is fragmented into multiple MMTP packets. The MMTP header includes a packet Identifier (ID) field for the separation of media data.

A Media Processing Unit (MPU) is formed as a logical group of MFUs considering the coding process. The first MPU in an MPU is a random access point of media data, which is the first Intra (I) picture in a Group of Pictures (GOP) structure for video coding. For audio coding, all AUs are random access points. However, it is reasonable for the boundaries of audio MPUs to match with those of video MPUs in the timeline. The UTC-based presentation timestamp is given for each MPU by an MMT Package Table (MPT), which provides list information for the media data of a service. The MPT also has location information for the media data, such as the IPv6 source/destination address, User Datagram Protocol (UDP) destination port number, and MMTP packet ID.

To achieve smooth migration from conventional TS-based broadcasting systems, we introduced a signaling information structure named M2 Section Message that can convey MPEG-2 Section format tables, such as Event Information Table (EIT) and Application Information Table (AIT).

2.3. Interface between layers

MMTP packets are variable-length while conventional TS packets are fixed-length. Therefore, a scheme to convey variable-length packets in broadcast channels is required. Type Length Value (TLV) multiplexing scheme is used as an interface signal format to store IP packets in modulation frames effectively. This scheme is specified in Recommendation ITU-R BT.1869.7) IP headers can be compressed for efficient transport because they are common to all IP packets in a single IP data flow and verbose in broadcast channels. Lost fields, such as a UDP checksum, can be recalculated by the receiver if necessary.
3. Prototype Equipments

3.1. Modulator and demodulator
We developed a modulator and demodulator that conform to the technical specifications listed in Table 1. External views of these units are shown in Fig. 3. The central frequency of the modulated carrier is 140 MHz. The demodulator receives a 1-GHz band intermediate frequency (IF) signal. The input signal format for the modulator is TLV/IPv4, which indicates an IPv4 packet conveys a TLV data structure. We use Unshielded Twist Pair (UTP) cables with RJ-45 connectors as the signal I/O interface rather than coaxial cables with BCN connectors as in the case of TS.

3.2. MMT streamer and receiver
We developed an MMT streamer and receiver software that runs on a general-purpose computer with an x86-64 CPU. The MMT streamer generates MMTP/UDP/IPv6 packets that contain media data as the payload. The payload data is offline-encoded by HEVC for video or Advanced Audio Coding (AAC) for audio. The MMT streamer supports both unicast and multicast. As this system is designed for hybrid delivery, the number of output ports is greater than one. The MMT streamer can generate MMTP/UDP/IPv6 packets for the general network switch or MMTP/UDP/IPv6/TLV/IPv4 packets for the modulator.

The MMT receiver can receive one or more MMT flows from one or more delivery channels including both broadband networks and broadcast channels. It can decode up to single 4K (3,840 by 2,160 pixels) or double 2K (1,920 by 1,080 pixels) resolution videos in real time and can display decoded video frames based on the UTC-based presentation timestamp.

4. Experiment

4.1. Experimental system
Fig. 4 shows an outline of the experimental system. We used two satellite channels over WINDS to evaluate the performance of the physical layer and synchronization functionality of the transport layer using MMT. For the two up-link paths to WINDS, two satellite terminals, which are the LET and VSAT, were utilized. As shown in Fig. 4, we named the path from the LET to LET via WINDS “LET line” and also named the path from the VSAT to LET via WINDS “VSAT line”. Fig. 5 shows the LET’s and VSAT vehicles. Both the two terminals were placed in the same location but there were no common reference signals.

On the transmitter side, the central frequency intervals between the two carriers were chosen to be 40 MHz based on the channel spacing in the Ku band communication satellite operating in Japan. The 1-GHz IF signals from two modulators were individually up-converted to 3-GHz IF signals in order to connect the satellite terminals. The one signal was transmitted from the VSAT and the other was transmitted from the LET. In other words, the two up-link carriers were transmitted in a Frequency Division Multiplexing (FDM) manner.

On the receiver side, two down-link carriers were received by the LET. The 3-GHz IF signal from the LET that contained the two carriers was down-converted to a 1-GHz IF signal by a frequency translator. Fig. 6 shows the received power spectrum. The down-converted 1-GHz IF signal was distributed to two demodulators and a spectrum analyzer. The demodulators were tuned to the different channels corresponding to the two transmitted carriers.

The MMT streamer on the transmitter side provided MMT flows to the modulators. On the receiver side, two MMT receivers received the MMT flows from the demodulators.

---

Fig. 4. Experimental system.
Fig. 5. LET’s and VSAT vehicles.
Fig. 6. Received power spectrum (1-GHz IF).
MMT receiver #1 was connected to both demodulators #1 and #2. Therefore, it could decode and display multiple videos from the two MMT flows on a display. On the other hand, MMT receiver #2 was connected to only demodulator #2 and could decode and display a single video. The MMT receivers were connected to a Network Time Protocol (NTP) server that provided the UTC as the reference clock for media presentation.

4.2. Physical layer performance

We transmitted pseudo-random bit sequence (PRBS) to measure the Bit Error Rate (BER) of the physical layer. Fig. 7 shows the C/N vs. BER characteristics of the WINDS transponder. The following modulation schemes were used for the test.

1. 16APSK (3/4*)
2. 16APSK (7/9*)
3. 32APSK (4/5*) *inner coding rate

The modulation schemes for the two lines (LET line and VSAT line) were the same on each test. Additive White Gaussian Noise (AWGN) was added to the received 1-GHz IF signal, and the C/N values were measured by the spectrum analyzer. Here, the required C/N is defined as the smallest value that yields a BER of $1 \times 10^{-11}$. As shown in Fig. 7, the required C/N values for each modulation were less than 16 dB.

In the case of the 32APSK, there was a larger deterioration of the required C/N from the 3-GHz IF local loopback to the satellite loopback than in the 16APSK case. The 32APSK is more vulnerable to the nonlinear distortion of the satellite transponder than the 16APSK. One of the reasons is that instantaneous power for the 32APSK is greater than that for the 16APSK as shown in Fig. 8. This factor causes greater amplitude compression and differential phase rotation over satellite.
4.3. Transport layer functionality

We also tested the synchronization functionality of MMT. The MMT streamer encapsulated two streams of HEVC videos that comprised a multi-view program. There were two patterns.

(1) The main view was at 4K resolution, and the sub view was at 2K resolution. The 4K main video and the 2K sub video were delivered on the LET line and VSAT line, respectively. MMT receiver #1 received the MMT flow from the LET line, and MMT receiver #2 received the MMT flow from the VSAT line. Decoded video frames were shown on a 4K display by MMT receiver #1 and on a 2K display by MMT receiver #2.

(2) Both the main view and sub view were at 2K resolution. The 2K main video and 2K sub video were delivered on the LET line and VSAT line, respectively. MMT receiver #1 received the MMT flows from both the LET line and VSAT line. The decoded video frames were shown on a 4K display by MMT receiver #1 in a picture-in-picture manner.

Fig. 9 is a photograph of the displays for pattern (1) on the receiver side. The largest display is the 4K display, and the smaller display at the center bottom is the 2K display. Due to the UTC-based presentation timestamps provided by the MMT signaling information, the receivers could decode video streams in a synchronized manner with one-to-one frame accuracy.

The difference between the delivery delays of the two lines (LET line and VSAT line) was very small because most parameters were the same, with the exception of carrier frequency and some parts of the transmitter side. We tested the feasibility of synchronous presentation by varying delays of the transmitted packets in the VSAT line intentionally at the output of the MMT streamer. Table 2 shows the results of delay alignment test for synchronous presentation on MMT receiver displays with various delay values for the VSAT line. In Table 2, “OK” means that both the main and sub video frames were presented successfully in a synchronized manner. “NG” means that video frames of the sub video from the VSAT line were dropped or delayed relative to those of the main video. In the current implementation, the MMT receivers had fixed size input buffer memories for 2 seconds and displayed decoded video frames with UTC-based presentation timestamps that included an added 2-second offset value. Thus, differences in delivery delay less than 2 seconds could be aligned by the MMT receivers.

This test was performed using the three modulation modes listed in Section 4.2, and the results for all modes did not differ from those shown in Table 2. Thus, it is evident that the modulation parameter does not affect on the synchronization functionality of MMT.

4.4. Packet delivery delay measurement

There is a delivery timestamp field in the MMTP header, in which UTC data (16 bits for an integer and 16 bits for a fraction) are assigned when the packet is generated. Using the value in this field, the receiver can assume end-to-end packet delivery delay under the condition that the system clock in the receiver is synchronized with that of the streamer. We measured the average delivery delays of MMTP packets using our prototype software in the following two cases.

(A) Translator loopback
(B) Satellite loopback (LET line)

The translator is almost the last part of the up-link terminal; thus, the value of (B) - (A) was expected to be nearly equal to the round-trip time of radio wave traveling between the earth to WINDS.

Table 3 shows the actual measurement results, and Table 4 shows the estimated round-trip time. Our current software-based implementation on general-purpose computers

<table>
<thead>
<tr>
<th>Modulation</th>
<th>Translator Loopback</th>
<th>Satellite loopback</th>
<th>Round-trip time</th>
</tr>
</thead>
<tbody>
<tr>
<td>16APSK (3/4)</td>
<td>0.12320 s</td>
<td>0.37127 s</td>
<td>0.24807 s</td>
</tr>
<tr>
<td>16APSK (7/9)</td>
<td>0.12272 s</td>
<td>0.37029 s</td>
<td>0.24757 s</td>
</tr>
<tr>
<td>32APSK (4/5)</td>
<td>0.11888 s</td>
<td>0.36676 s</td>
<td>0.24787 s</td>
</tr>
<tr>
<td>Average</td>
<td>0.12160 s</td>
<td>0.36944 s</td>
<td>0.24784 s</td>
</tr>
</tbody>
</table>

Table 4. Estimation of round-trip time.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satellite orbit radius (m)</td>
<td>42,164,000</td>
</tr>
<tr>
<td>Equatorial radius (m)</td>
<td>6,378,137</td>
</tr>
<tr>
<td>WINDS/south latitude (degree)</td>
<td>0.01</td>
</tr>
<tr>
<td>WINDS/east longitude (degree)</td>
<td>143.02</td>
</tr>
<tr>
<td>NICT Kashima/north latitude (degree)</td>
<td>35.95</td>
</tr>
<tr>
<td>NICT Kashima/east longitude (degree)</td>
<td>140.66</td>
</tr>
<tr>
<td>Calculated straight distance from NICT Kashima to WINDS (m)</td>
<td>37,195,401.77</td>
</tr>
<tr>
<td>Light speed (m/s)</td>
<td>299,792,458</td>
</tr>
<tr>
<td>Round-trip time (s)</td>
<td>0.248141011</td>
</tr>
</tbody>
</table>

Fig. 9. Synchronous video display on receiver side.
achieved precise measurement of the delivery delay time. The expected measurement precision may differ depending on the implementation of the MMT streamer and receiver.

5. Possibility of Hybrid Service with MMT

In satellite broadcasting services, it is assumed that programs, promotions, and advertisements are common for all audiences across a wide service area, such as all of Japan. Under hybrid delivery mechanisms with MMT, broadcast services are expected to be more customized than that of conventional ones.

One of the possible service examples is to provide videos from multi-view angles in the same program, which was tested in our experiment. The audience can receive videos from alternative angles or other related information via additional channels on demand. The additional channels can be obtained from satellite broadcasting, terrestrial broadcasting, broadband networks, and so on.

Another example is targeted promotions and advertising. A broadcaster can provide customized contents to specific audience groups according to age, gender, location, and preferences. This customization increases the effectiveness of promotions and advertisements than that for unspecified general audiences.

Another possibility is the delivery of scalable video coding. With a scalable video coding system, such as Scalable HEVC (SHVC), MMT can provide synchronous mechanisms for two or more coding layers transported over two or more channels with different error tolerance. For example, a 4K is transmitted on the base layer with higher noise tolerance and less channel capacity. An 8K is on the enhanced layer having more channel capacity than base layer but weaker noise tolerance. In addition, temporal scalability for frame rate upgrade, such as 60p (60 frames per second progressive) to 120p, is also considered as one of the possible use cases. Hybrid content delivery with MMT may have greater possibilities beyond these examples.

6. Conclusion

Our experimental results indicate that the ARIB standards for next-generation satellite broadcasting are feasible using the WINDS transponder. Furthermore, MMT’s synchronous mechanisms enable synchronized media presentation even if the up-link systems for each channel are asynchronous in operation.

In our experiment, we used two satellite channels in the same radio frequency band with a common modulation system. The MMT-based next-generation broadcasting system may also be utilized by combinations of different radio frequency band, different modulation schemes, and different types of networks, such as satellite with terrestrial, terrestrial with LTE, and so on.

This experiment represents the first delivery of MMTP packets over actual satellite channels. We plan to continue development of equipment and conduct transmission trials for 8K SHV broadcasting and future hybrid services with MMT. We believe that MMT will facilitate effective harmonization and service evolution of broadcasting and broadband systems.

Acknowledgments

We would like to thank the National Institute of Information and Communications Technology (NICT) for providing us the opportunity to perform our satellite transmission experiment using the “KIZUNA (WINDS)” satellite.

References