Ultra Low Delay and Resilient Multi-channel Video Encoder for Efficient Remote Construction

Yasuyuki Nakajima[1], Yasukazu Hontama[1], Kazuhiko Shibairi[1], Yasuyuki Jitsuta[2], Daisuke Endo[3], Genki Yamauchi[3], Takeshi Hashimoto[3], and Hitoshi Itoh[4]

[1]Hytec Inter Co. Ltd., Japan
[2]Jitsuta Chugoku Co. Ltd., Japan
[3]Public Works Research Institute, Japan
[4]Chudenko Corporation, Japan

y-nakajima@hytec.co.jp, hontama@hytec.co.jp, shibairi@hytec.co.jp, y.jitsuta@jtch.co.jp,
endou-d177cl@pwri.go.jp, yamauchi-g573bs@pwri.go.jp, t-hashimoto@pwri.go.jp, h.itoh@chudenko.co.jp

Abstract –

Efficient remote operation of construction machines equipped with multiple video cameras requires smooth and real-time video transmission to simulate the experience of operating machinery in person. Due to the diverse network environments at construction sites, including satellite, Wi-Fi, and both public and private LTE/5G networks, an ultra-low delay video codec with adaptive bitrate control is crucial to maintaining low end-to-end latency in these environments. In this paper, the development of a multi-channel video encoder for the above requirements is described demonstrating the codec's performance across five types of wireless networks at a remote construction site.

Keywords –

Unmanned construction, Remote operation, Low delay video transmission, Network resilience

1 Introduction

In 2024, Japan's Ministry of Land, Infrastructure, Transport and Tourism (MILT) introduced the 'i-Construction 2.0' action plan, aiming to automate construction sites by 2040 [1]. In this plan, it is reported that Japan's working-age population is expected to be 20% smaller in 2040 than in 2023. Therefore, the plan focuses on labor reduction via increase of productivity through automation and autonomy in construction. For this purpose, remote construction technology is expected to help to solve labor issues as a step toward automated and autonomous construction.

In the remote construction, it is necessary to observe the situation of the remote site via the captured video, but there is a possibility of decrease of construction efficiency due to transmission delay and packet loss when using wireless networks. In this paper, firstly issues related to the remote construction using wireless networks such as private 5G and LEO (Low Earth Orbit) satellite networks are discussed. Next, we detail the development of a multi-channel, low-delay video codec to address these challenges.

Through experiments conducted across five types of wireless networks, we demonstrated that the ultra-low delay codec supports adaptable remote construction architectures tailored to each network.

2 Issues and countermeasures for remote construction

Figure 1 shows a typical system configuration of a remote construction and its issues (#1-#4) for efficient operation. In the following sections, we describe each issue and its countermeasures.

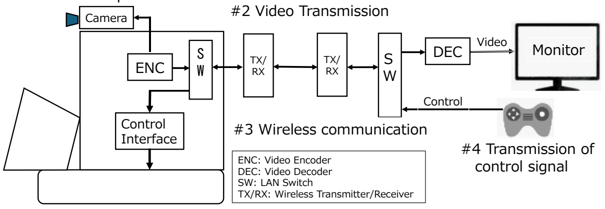

Figure 1. System configuration of remote construction

2.1 Issues and countermeasures for implementation

First of all, the issues with remote, automated, and autonomous operations of construction machines lay in the fact that it requires high expertise for installation and setup in such fields as video transmission, machine control, and communication networks since these factors are involved for the total systems. For example, four video cameras facing forward, backward, left, and right would be installed on a construction machine to grasp the field situation and spot any danger. Since each camera requires a separate video encoder, a multi-channel encoder is preferable to reduce both cost and space. In addition, all the setup processes for the total system should be as simple as possible.

2.2 Issues and countermeasures for video transmission

As for the issues in remote operation using video, (1) end-to-end delay, (2) coding bitrate fluctuation, and (3) communication network quality should be pointed out. For example, if the advanced video coding architecture with temporally separated frames is employed to enhance the prediction performance, the coding efficiency can be improved. However, this might affect coding delays since the encoder has to wait for the following frames to encode the current frame. In the following subsections, each issue and countermeasures are described.

2.2.1 End-to-End delay

In this subsection, methods for end-to-end low delay architecture and target delay are discussed.

Remote construction efficiency would be directly affected by end-to-end delays [2]. As shown in Figure 1, each component such as video camera, video encoder, or communication network involves a certain amount of delay. Therefore, it is required to decrease each delay as much as possible. For example, 60fps video camera should be used instead of 30fps since the delay becomes 16.6msec instead of 33msec. On the other hand, delays of video codec and communication network may vary from tens to hundreds of msec, and suitable components should be chosen to realize very low end-to-end delay.

As for the human sensitivity on delay, it is reported that a delay of 100msec or over can be perceived [3]. For the electric rear-view mirrors for automobiles, the UN regulated the end-to-end delay not to take more than 200msec [4]. As for the remote construction, it is reported that the work efficiency was almost the same as the onsite remote operation when the end-to-end delay was less than 300msec [2]. Therefore, this figure can be recognized as one of the targets for the efficient remote construction.

2.2.2 Video coding delay and bitrate fluctuation

Here, delay and coding bit rate fluctuation mechanism caused by video coding scheme and countermeasures are addressed.

The international standard for video coding, MPEG (ranging from MPEG-1 and MPEG-2 to recent standards HEVC and VVC), combines intraframe coding (I frames) with interframe coding (P frames and B frames) to achieve a high compression ratio [5]. The intraframe coding reduces redundant information within the frame, whereas the interframe coding reduces redundant information between frames located temporally earlier and/or later. Here, P frames only use the earlier frames as a forward prediction and B frames can use both the earlier and the later frames as a bi-directional prediction.

In terms of coding efficiency, I frames are usually inserted every 0.5sec or 1sec to be used as random-access points, and the other frames are coded by interframe coding (P or B frames). Normally, considering the high correlation of frames in the temporal domain, P frames are inserted every few frames and the rest of the frames are coded with B frames to enhance the coding efficiency as shown in Figure 2(a).

	#1	#2	#3	#4	#5	#6….				
(a)	I	B	B	P	B	B	P…	I	B	B P…
(b)	I	P	P	P	P	P	P…	I	P	P…
(c)	I	P	P	P	P	P	P…	P	P	P…

Figure 2. Examples of coding frame layouts

Although this IBP frame structure can achieve a high compression ratio, there are two issues in terms of delay and coding bit amount for remote construction using network.

The first issue arises from variations in coding delay based on the video codec architecture. As stated earlier, MPEG uses interframe coding predicted from the earlier and/or later frames. Therefore, both encoder and decoder must wait for the later frames to encode/decode the current frame, which results in additional delay. In the case of Figure 2(a), for example, the frame #2 is coded as a B frame, but it can be coded after the frame #4 is coded as a P frame.

The second issue evolves from the fact that coding bit amount varies depending on the coding frame types. In the IBP coding structure, I frames have the largest bit amounts and B frames have the smallest bit amounts. For example, there is a case where the ratio of bit amounts for B:P:I frames becomes 1:10:25 [6]. Therefore, there is a possibility where the coding bit amounts of I frames are very large and it may take some time to send I frame data for limited network bandwidth. Then in the decoder side, the decoding timing of I frames becomes late for playback timing which may result in a frame drop.

In order to achieve low delay coding, the use of only earlier frames (i.e. IP frame structure) for the interframe coding as shown in Figure 2(b) is preferable in exchange for lower coding efficiency than that of IBP frame structure. As for the coding bit amount fluctuation, the introduction of intra-slice architecture within an IP frame structure as shown in Figure 2(c) which is further described in Section 3 may alleviate the bit amount variation.

2.3 Issues and countermeasures for communication networks

2.3.1 Communication network quality

In both wired and wireless networks, it is necessary to pay attention to four types of quality degradation causes (transmission delay, jitter, packet loss, and bandwidth fluctuation) for video transmission over IP networks. In this section, these issues and countermeasures are described.

As for the delay, low delay transmission can be realized when local networks such as Wi-Fi and private 4G(LTE)/5G. However, when public networks such as 4G(LTE) and 5G are employed for remote operation, a few dozen msec is usually involved due to transmission over the backbone and the access networks. In the case of satellite networks, GEO (Geostationary Earth Orbit) satellite communication has a few hundred msec roundtrip delay, whereas less than 100msec delay can be expected in the LEO satellite communication. To choose the desired network, network delay should be investigated before designing the remote construction system.

Regarding the jitter, since the video decoder is required to output the decoded frames to a video monitor within the playback timing, there is a possibility of frame drops when the packet data arrival is delayed due to the jitter. To solve this issue, the decoder buffer size should be designed to absorb the maximum jitter which can be determined by the actual amount of jitter.

As for the packet loss, obstacles and/or noise may cause transmission losses, especially in a wireless environment. The former packet loss is caused when the connection between a construction machine and a base station is interrupted by the backhoe movement, for example. The latter packet loss is caused by heavy rain and/or interference where packet data cannot be recovered due to noisy signals, for example. To solve the former packet loss issue, ARQ (Automatic Repeat reQuest) technique should be used to retransmit the lost packets while it involves additional delay. To solve the latter issue, FEC(Forward Error Correction) should be used, although it requires additional overhead data for error correction.

With regard to the bandwidth(throughput), the upload bandwidth from a construction machine is a very important parameter which determines transmission video quality. In the case of public networks such as 4G and 5G, more than a few Mbps throughputs can be expected for upload link. Recently, satellite networks using LEO became widely available and users can get up to 10Mbps for upload via Starlink [7], whereas only a few hundred kbps bandwidth can be used in satellite networks such as Inmarsat on the GEO. However, there is an issue of bandwidth variation of public networks since it may change dynamically according to the amount of network usage. To solve this, the video encoder should adapt to the available bandwidth.

2.3.2 Wireless networks

Although the history of remote construction in Japan is about 30years, the wireless network environment has been greatly advanced in this period. The unmanned construction using video cameras installed on the construction machines first adopted and applied for the whole disaster recovery work after the eruption of Mt.Unzen from 2km distance under wireless and relay environment at 2Mbps in 1993 [8].

In the damaged Fukushima nuclear power plant after the East Japan Great Earthquake in 2011 in Japan, ten construction machines (two 600-ton class large crawler tower cranes and eight construction machines) with video cameras were remotely operated from 500m distance using five mesh Wi-Fi base stations at 5GHz [8]. Recently, high speed wireless networks such as 5G network and LEO satellite network became widely available and a wide choice of wireless networks including Wi-Fi became possible. However, as stated earlier it is necessary to transmit low delay, stable, and high-quality video over the wireless network to realize an efficient remote construction.

In addition, resilient wireless networks should be constructed using mesh type network and/or error protection techniques such as ARQ for such cases as disconnection due to movements of construction machines, for example. It should be noted that preparation of a backup network would also be considered for network resiliency and safe remote operation. Another issue we should consider is that adaptive bitrate control for stable transmission is also very important since the bandwidth of wireless networks may fluctuate due to network congestion and/or interference, for example. In section 5, we show the solutions toward the above network related issues employing five types of broadband wireless networks and ARQ to realize efficient remote operations of construction machines.

2.4 Control signal transmission

In order to remotely control construction machines, control signals should be transmitted between the construction machines and the control sites. The amount of coded video data will become very large when the construction machines move or rotate since a large amount of coded video data is expected due to large motion or scene changes. Therefore, there are many cases where video is transmitted over Wi-Fi, while control signal is transmitted over the other network such as specified low power radio as in the case of Fukushima [8], which may result in complicated system configuration. In order to transmit both the video and the control data in the same network, it is necessary to strictly control the total amount of coded video data below the bandwidth limit under any circumstances.

3 Development of multi-channel ultra low delay video codec

In this section, the development of video codec to solve the issues of low delay optimization, bit rate control mechanism and multi-channel support is described. Table 1 summarizes the main issues and solutions.

Table 1. Main issues and solutions

Issues	Solutions
Codec implementation	4ch HD encoder
Coding delays	IP frame structure
I-frame bursts	Intra slice coding
Bandwidth fluctuations	Strict rate control

For situation awareness and safe operations, a typical construction machine is equipped with four video cameras positioned at the front, rear, and both sides. In addition, at least HD quality video cameras would be required for the efficient remote operation [9]. Considering the above resolution requirement and overall network bandwidth, an all-in-one type four channel 2K low delay video encoder(LVRC-4000) based on the 4K HEVC video encoder (it can only support IP frame structure) [10] was developed, which can be installed on a construction machine (Figure 3 and Figure 4).

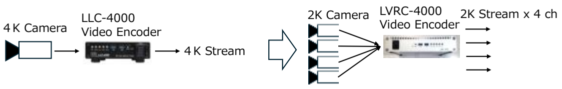

Figure 3. Development of 4ch HD video encoder

Figure 4. Ultra-low latency 4ch HD video encoder

As for the video coding structure, we use IP frame structure shown in Figure 2(b). Unlike the IBP structure, which uses both forward and backward predictions, the IP structure relies solely on forward prediction. While this reduces coding efficiency, it significantly lowers codec delay by eliminating the need to wait for future frames during encoding and decoding.

To address bandwidth fluctuations, the codec employs strict bitrate control, ensuring that the data rate remains within the available bandwidth at all times. In the case of IP frame structure shown in Figure 2(b), periodic insertion of I frames can cause bitrate spikes that exceed available bandwidth, leading to potential transmission delays or frame drops. Therefore, as an alternative structure, we have adopted an intra slice scheme within IP frame structure, which can also be used in HEVC coding. In this case, as shown in Figure 2(c), the first frame is encoded as I frame, but all the subsequent frames are coded by P frames where a slice of P frame is intra coded periodically. This intra slice coding can improve bitrate stability by periodically encoding small portions of each frame independently, preventing sudden spikes in data transmission.

For example, as shown in Figure 5, shadowed blocks (blocks on the third line in the figure) which construct a slice of blocks are intra coded. The position of the intra slice is moved at each frame. Then, in the next P frame, the slice of blocks below the current slice is intra coded. Since each P frame contains both intra- and inter-coded blocks, maintaining consistent picture quality and frame-to-frame bitrate stability is crucial. We implemented a strict rate control mechanism that regulates the bitrate of each frame, taking into account the differing data requirements of inter- and intra-coded blocks.

This approach ensures stable picture quality and bitrate while minimizing coding delay.

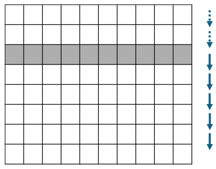

Figure 5. Intra slice coding

4 Video encoder performance

The video encoder performance was evaluated based on two key metrics: coding delay and bitrate fluctuation. Tests were conducted with four asynchronous video inputs, including high-motion scenes, to assess encoder stability under varying conditions. All the measurements in Section 4 and Section 5 were carried out for 3min. As shown in Figure 6, this experiment was carried out indoors since video codec performance test without wireless network can only be done in the indoor system.

Figure 6. Example of high-motion video input used for encoder evaluation

Figure 7 and Figure 8 show coding bitrate variations of IP frame structure for one and four channels, respectively. Table 2 shows statistics (average bitrate, fluctuation range, standard deviations and 95% confidence intervals) of coding bitrate variations. Here, the coding bitrate is set to 3Mbps and the intra frame insertion period is set to 0.5sec which is generally used. In the case of one channel encoding, the coding bitrate changes rapidly according to motion and texture change, and periodical peaks at every 0.5sec can be seen in Figure 7. Since relatively large fluctuation is observed as shown in the table, it is necessary to set the coding bitrate to much lower than the available bandwidth in order to correspond to the high peak bit rate, which may result in lower video quality than expected.

Table 2 Statistics of coding bitrate variations

Coding Mode, #ch	AVG(kbps)	Fluctuation Range	SD(kbps)	95% CI(Mbps)
IP structure, 1ch	2681	+28% to -38%	217	2.71 to 2.65
IP structure, 4ch	11209	+3.5% to -14%	366	11.26 to 11.16
Intra slice, 1ch	2896	+1.5% to -3%	27	2.90 to 2.89
Intra slice, 4ch	10462	+3% to -8%	160	10.49 to 10.43

On the other hand, in the case of four channels, although the periodic peaks at every 0.5sec can also be observed in Figure 8, the bitrate fluctuation range becomes smaller than that of one channel as shown in the table. This would be caused by the statistical multiplexing effect for four asynchronous video channel encoding, which absorbs bit rate fluctuation of each channel to some extent. In statistical multiplexing, it distributes bitrate peaks across multiple asynchronous streams, reducing overall fluctuation.

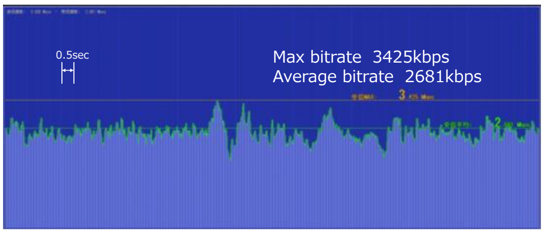

Figure 7. Bitrate variation for one-channel video encoding with IP structure (3Mbps x 1ch, Intra@0.5sec)

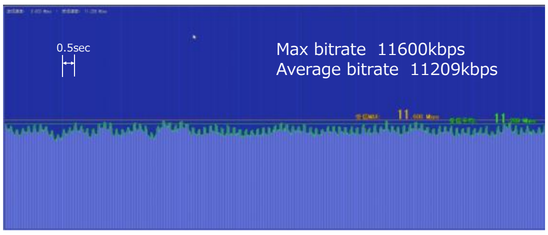

Figure 8. Bitrate variation for four-channel video encoding with IP structure (3Mbps x 4ch, Intra@0.5sec)

Figure 9 and Figure 10 show the bitrate variations for one and four channels, respectively when intra slice structure is employed. In this case, the bitrate fluctuations to the average are greatly reduced when compared with IP structure, which can also be observed in Table 2. As for the average bitrate, it was 2681kbps for the IP structure, whereas it was 2896kbps for the intra slice structure. Since the bitrate was set to 3Mbps for both structures, it can be said that the latter structure can use the available bandwidth more efficiently than the former structure.

When the amount of the bitrate fluctuation of four channels and one channel in intra slice coding is compared, the former is larger than the latter. Since the fluctuation range for each channel is very small due to intra slice coding, each small amount of fluctuation may accumulate when four channel bit streams are integrated.

As for the codec (from the encoder input to the decoded monitor) latency, 114msec delay was observed for the IP structure, whereas 66msec delay was obtained for the intra slice structure. This result means that the proposed intra slice structure can achieve both ultra-low latency while strictly keeping the coding bitrate and making the most of available bandwidth.

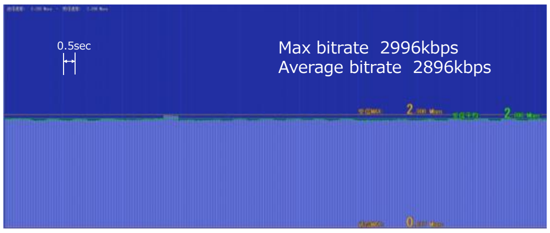

Figure 9. Bitrate variation for one-channel video encoding with intra-slice structure (3Mbps x 1ch)

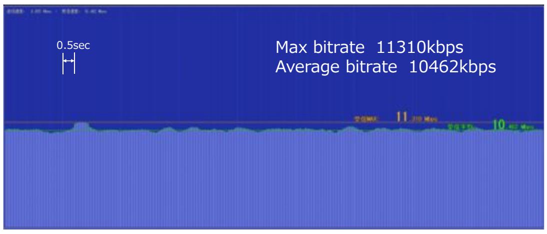

Figure 10. Bitrate variation for four-channel video encoding with intra-slice structure (3Mbps x 4ch)

5 Video transmission experiment

5.1 Experimental setup

We have conducted an end-to-end video transmission experiment using video coding setup with intra slice structure described in Section 4. The experiment aimed to evaluate transmission performance across various wireless networks using a remote construction machine setup, as shown in Figure 11. Four video cameras (OMRON SENTECH video cameras (STC-HD853HDMI) for full HD@60fps) and the newly developed encoder (LVRC-4000) were installed on the construction machine (Hitachi Construction Machinery ZX-120).

Regarding the wireless networks, we used the following commercially available networks to evaluate various types of wireless environments.

1) Wireless for robots (5.7GHz, Doodle Labs RM-5700, point to point, Antenna 2.4m high)
2) Mesh Wi-Fi (Rajant ES1-2450R@const. machine, FE1-2450x2+ES1-2450R@AP)
3) Public LTE (au, Hytec Inter HW5G-3200x2)
4) LEO satellite network (Starlink)
5) Private 5G (Fujitsu@AP, Hytec Inter HW5G-3200 x 2)

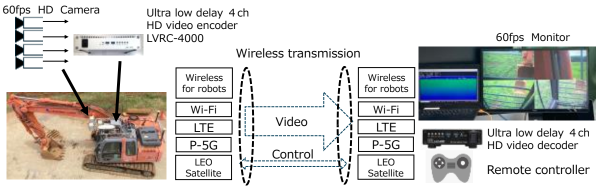

Figure 11. Experimental setup

As for 1) wireless for robots, it became available under license for image communications of unmanned systems such as robots and drones at 5.7GHz and 2.4GHz in Japan in 2016 [11]. As for 4) LEO satellite network, considering the situation of public mobile and optical fiber network failure due to a large natural disaster, we had an experiment connecting the construction machine to the LEO satellite network via mesh Wi-Fi, and video data was received at the remote control center via the Internet and public mobile and optical fiber network.

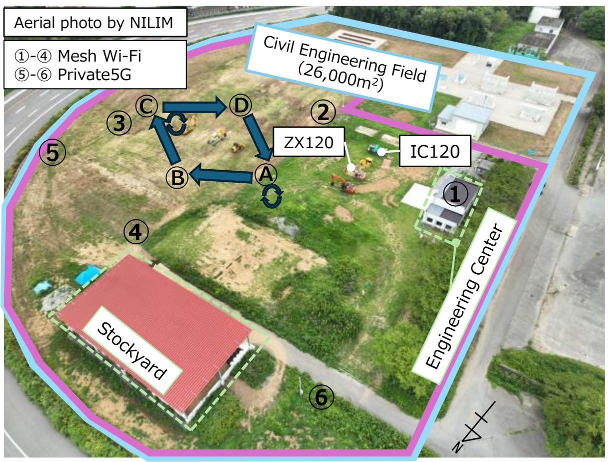

Figure 12. Experimental field

Figure 12 shows an experimental field. Four video cameras, video encoder, and network equipment were installed on the construction machine (ZX120). A 60fps video monitor (Iiyama XUB2390HS-B5), a video decoder (Hytec Inter LLC-4000), a remote controller (Hitachi S-RS-A), and network equipment are installed in the engineering center. In the experimental field, the construction machine moved around from point A to point D via points B and C, and rotated at points A and C.

Transmission experiments were conducted using an open platform called OPERA (Open Platform for Earthwork with Robotics and Autonomy) [12] facilitated for the efficient development and easy dissemination of autonomous construction technology. Table 3 shows the basic transmission performance for various wireless networks measured for 3min. For example, wireless for robots achieved 27Mbps uplink and 26Mbps downlink throughput with stable video transmission for both one channel and four channels. Similar throughputs are also obtained in public LTE and private 5G. Mesh Wi-Fi achieved 50Mbps for both uplink and downlink whereas LEO achieved 15Mbps level throughput.

In addition, the required bandwidth for the control signal was very small(<1kbps). Therefore, the control signal was transmitted in the same wireless network as that of the video data in the following experiments.

Figure 13 shows an example of one channel video transmission over the LEO satellite network at 3Mbps. As can be seen in the figure, stable video coding and transmission were realized through effective bitrate control.

Table 3. Throughputs, video coding bitrate and control signal bitrate for each network

System config.	Stay/ Move	Trhoughput Mbps		Coding bitrate per ch	Coding bitrate per 4ch	Datarate for control	Total bitrate
		To BS	From BS				
Wireless for robots	Move	27	26	3Mbps	12Mbps	Under 1kbs	Under 15Mbps
Mesh Wi-Fi	Move	52	50	3Mbps	12Mbps	Under 1kbs	Under 15Mbps
Public LTE	Move	30	26	3Mbps	12Mbps	Under 1kbs	Under 15Mbps
Private 5G	Stay	26	26	3Mbps	12Mbps	Under 1kbs	Under 15Mbps
LEO sat + fiber	Stay	15	14	1Mbps	4Mbps	Under 1kbs	Up to 6Mbps

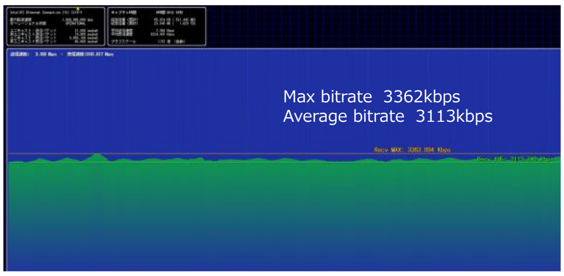

Figure 13. Transmission example of one channel video at 3Mbps over LEO satellite network

5.2 Network characteristics when the construction machine was stationary

While stationary, the networks achieved stable throughput and low packet loss, with latency varying by network type. Wireless for robots and mesh Wi-Fi networks demonstrated minimal delay, while public LTE and private 5G showed slightly higher average delays. Table 4 shows the transmission performance when the construction machine was stationary. Due to functional limitation of measurement equipment, detailed statics such as standard deviation of jitter were not obtained.

Since less than 15Mbps throughput is obtained in LEO network, we set 1Mbps per channel for this network. Similarly, we set 3Mbps per channel for the other networks since more than 25Mbps throughputs are obtained.

As for the network latency, 4-5msec average delay was observed in the wireless for robots and mesh Wi-Fi networks, whereas there was 30-40msec average delay for the public LTE and private 5G networks. In the case of the LEO satellite network, there were 84msec average delay via the public LTE and 60msec delay via the optical fiber networks, which are much smaller latency when compared with that of the conventional GEO satellite networks.

As for the delay jitter, due to network configuration limitation for the measuring equipment, we could only have results on wireless for robots and mech Wi-Fi. Although a small jitter of +/- 7msec was observed in the former, about 50msec jitter was obtained in the latter. Although packet loss rate was only available for the wireless for robots, there was no packet losses in the experiment.

Table 4. Throughputs, packet loss, network delay, and jitter for each network when move or stay still

System config.	Stay/ Move	Throughput(Mbp		Packet loss(%)	Delay(ms)			Jitter(ms)	
		To BS	From BS		Min	Avg	Max	Min	Max
Wireless for robots	Stay	33	28	0%	2.77	3.87	8.25	-6.02	6.98
same as above	Move	27	26	0.80%	2.86	24.34	401.8	-448.5	283.7
Mesh Wi-Fi	Stay	39	44	-	2.96	5.40	30.00	-52.66	52.81
same as above	Move	52	50	-	3.47	11.82	142	-274.0	275.1
Public LTE	Stay	29	27	-	22	33	54	-	-
same as above	Move	30	26	-	22	46	175	-	-
Private 5G	Stay	26	26	-	31	44	56	-	-
LEO sat + LTE	Stay	9	10	-	55	84	203	-	-
LEO sat + fiber NW	Stay	15	14	-	33	60	290	-	-

5.3 Network characteristics when the construction machine was in motion

Table 4 also shows the transmission performance when the construction machine was moving and rotating. In this case, we compared the results of wireless for robots, mesh Wi-Fi and public LTE with stationary situations. While the throughput remained stable during movement, delays increased significantly. For instance, the wireless for robots network experienced an increase from 4msec to 24msec average delay, with jitter rising to 300-450msec due to such motion as rotation of the construction machine. On the contrary, in the case of mech Wi-Fi, although the min and max delay jitter was about 300msec level, the average delay was about 12msec due to the use of multiple routing even when the construction machine was rotating.

5.4 End-to-End delay from video camera to monitor

Table 5 and Table 6 show the end-to-end delay from video camera input to monitor output when ARQ is off and on, respectively. As can be seen in Table 3, the end-to-end delay was 115-163msec and very low delay transmission was realized. When the ARQ was on (installed in the video codec), there was an additional 100msec for this function and the total delay of less than 300msec was achieved as shown in Table 6, which satisfies the delay requirement of efficient remote construction [2].

Table 5. End-to-End delay when ARQ is off

Network	Network delay(ms)	Delay expt. NW(ms)	E-E delay(ms)	Bandwidth(Mbps)
Wireles for robot	24	103	127	27
Mesh Wi-Fi	12	103	115	52
Public LTE	46	103	149	30
Private 5G	44	103	147	26
LEO + fiber	60	103	163	15

Table 6. End-to-End delay when ARQ is on

Network	Network delay(ms)	Delay expt. NW (ms)	E-E delay(ms)	Bandwidth(Mbps)
Wireles for robot	24	201	225	27
Mesh Wi-Fi	12	201	213	52
Public LTE	46	201	247	30
Private 5G	44	201	245	26
LEO + fiber	60	201	261	15

5.5 Improvement of transmission video quality when ARQ was on

When the remote operator rotated the construction machine, there was a network disconnection between the base stations and the construction machine when wireless for robots or private 5G was used due to the limited number of base stations (one or two). In the case of ARQ=off, picture quality degradation and/or frame drop were observed due to packet loss. When the mesh Wi-Fi or the public LTE was used, the frequency of instantaneous interruption was low since the possibilities of either one of the base stations within the line of sight was high. Especially in the case of the public LTE, very stable connection was observed. However, it should be noted that the situation would vary depending on the area coverage or network congestion.

(a) ARQ is off (b) ARQ is on

Figure 14. Comparison of picture quality when ARQ is off(left) and on(right)

When ARQ was on, stable video transmission was realized at every wireless network. For example, in

Figure 14(a), the shape of the building in the right-hand side was degraded due to packet loss when ARQ was off. On the contrary, when ARQ was on, lost packet was successfully retransmitted as shown in Figure 15 and error-free decoding was realized as shown in Figure 14(b).

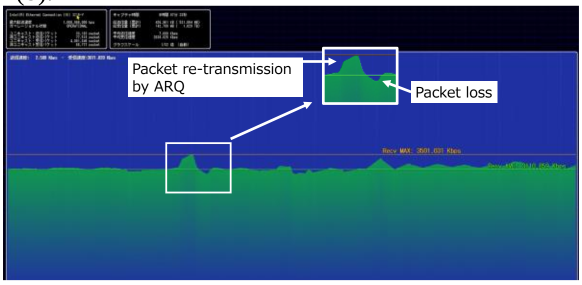

Figure 15. Packet loss and recovery via ARQ

6 Conclusions

This study focused on the development and evaluation of a multi-channel ultra-low delay video encoder to enhance remote construction efficiency. Network performance and video quality were assessed across five broadband wireless networks, including wireless for robots, mesh Wi-Fi, public LTE, private 5G, and LEO satellite. Enabling ARQ increased end-to-end delay by 100msec but significantly enhanced video quality, particularly during high-motion events such as machine rotation.

Throughout the experiments, less than 300msec of total delay was realized by the proposed system using commercially available wireless networks, meeting the delay requirements for efficient remote construction operations. While the system demonstrated high performance, simplifying the setup process remains a challenge. To support adoption in diverse construction environments, future work should focus on developing user-friendly interfaces and automated configuration tools that minimize the need for technical expertise particularly for operators with limited technical experience.

Acknowledgments

This study was carried out through the joint research work with Public Works Research Institute, Jitsuta Chugoku Co. Ltd., and Chudenko Corporation on productivity improvement of mechanical earthwork using OPERA (Open Platform for Earthwork with Robotics and Autonomy). The authors would like to thank all for discussions and support of the experiments.

The authors are also grateful to National Institute for Land and Infrastructure Management (NLIM) for the use of private 5G system.

References

[1] Japanese Infrastructure Ministry Unveils Action Plan “ i-Construction 2.0 ” to Achieve Construction Site Automation by 2040. On-line: https://www.decn.co.jp/?p=163015, Accessed: 31/03/2025.

[2] Ito S., Fujino K. and Ando H. Evaluation of work efficiency in the unmanned construction with communication delays. *The 16th Symposium on Construction Robotics in Japan*，No.06-1，2016.

[3] Miyasato T. and Kishino F. Subjective evaluation of perception of delay time between visual information and tactile information. *The Journal of the Institute of Television Engineers of Japan*, Vol.49，No.10，pp.1353-1356，1995.

[4] Proposal for Supplement 4 to the 04 series of amendments to Regulation No. 46 (Devices for indirect vision) 2016. On-line: https://unece.org/fileadmin/DAM/trans/doc/2016/wp29/ECE-TRANS-WP29-2016-089e.pdf, Accessed 31/03/2025.

[5] Sullivan G., et al. Overview of the High Efficiency Video Coding (HEVC) standard. *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649-1668, 2012.

[6] Si J., Ma S. and Gao W. Efficient bit allocation and CTU level rate control for High Efficiency Video Coding. *2013 Picture Coding Symposium (PCS)*, pp. 89-92, 2013.

[7] Starlink specifications. On-line: https://www.starlink.com/, Accessed:08/04/2025.

[8] Chayama K., Fujioka A., Kawashima K., Yamamoto H., Nitta Y., Ueki C., Yamashita A., and Asama H. Technology of Unmanned Construction System in Japan. *Journal of Robotics and Mechanics*, Vo.26 NO.4: 403-417, 2014.

[9] Xiang Y., Xu B., Su T., Brach C., Mao S.S., and Geimer M., 5G Meets Construction Machines: Towards a Smart Working Site. *2021 International Conference on Computing and Communications Applications and Technologies (I3CAT)*, 28-35, 2021.

[10] LLC-4000 -4K Video Encoder/Decoder optimized for 5G/LTE. On-line: https://hytec.co.jp/eng/prod/16271.html, Accessed: 31/03/2025.

[11] License-required Radio Equipment for Unmanned Aircraft (UA). On-line: https://www.tele.soumu.go.jp/e/sys/others/drone/, Accessed: 31/03/2025.

[12] Yamauchi G. et.al. Proposal of an Open Platform for Autonomous Construction Machinery Development. *Proc. of the 40th ISARC*, 2023.