

# Effectiveness of a Fast Switching Function of Continuous Video Signals on Unmanned Vehicle Controls

By Hisaya HADAMA, Masaki OGURA and Ryohei NAKAMURA

Reprinted from the Memoirs of the National Defense Academy  
Vol.57, No.1

September 2017

---

YOKOSUKA, JAPAN

## Effectiveness of a Fast Switching Function of Continuous Video Signals on Unmanned Vehicle Controls

By Hisaya HADAMA\*, Masaki OGURA\*\* and Ryohei NAKAMURA\*\*\*

(Received: March 31, 2017; Accepted for publication: June 9, 2017)

### Abstract

This research aims at the development of an affordable unmanned vehicle (UV) that can be controlled through cellular networks. Currently, given the instability of the packet transmission qualities of the cellular network, communication is likely to be interrupted occasionally. The interruption of the control signal would impede the precise real-time control of a UV. However, we believe that it might be possible to decrease the occurrence of these interruptions by transmitting communication over multiple wireless networks, given their availability in numerous areas. Accordingly, in this paper, we propose a fast switching function that can be used to facilitate video signal transmission and assist effective UV control. We begin by elucidating the availability and packet transmission qualities of commercial networks, then provide a detailed explanation of our proposal, and finally conclude with our evaluation results.

**Keywords:** Fast Switching, Video Signal, Cellular Network, UV Control

### 1. Introduction

Recently, unmanned vehicles (UVs) have been utilized in various roles in diverse fields, such as transportation, disaster recovery, agriculture, surveillance of architectural structures, and reconnaissance. Many studies have been performed<sup>1-4)</sup> that discuss other potential applications of UVs.

We believe that cost-effective UV systems can be realized using affordable and reliable wireless networks. However, to realize real-time remote control through a network, a reliable and real-time signal transmission channel is required. This channel must have a certain level of packet transmission quality, regardless of whether the UV is controlled autonomously or manually.

For example, the US military has built a dedicated global communication network with less than 0.5 s latency<sup>5)</sup>. However, this type of dedicated network is not available at affordable prices for many applications.

Meanwhile, many researchers have been

---

\*Professor, Department of Communications Engineering, National Defense Academy

\*\*JMSDF Repair and Supply Facility Yokosuka

\*\*\*Research Associate, Department of Communications Engineering, National Defense Academy

working to apply widely available mobile cellular networks to remote controls of UVs and other devices<sup>6, 7)</sup>. Currently, cellular networks, such as 3G and 4G long-term evolution (LTE), are widely available and can be accessed and used by almost anyone without difficulties. Accordingly, we believe that if these commercially available mobile networks could be utilized for the control of UV signal transmission channels, it would be possible to operate them over wide areas at affordable costs.

The aim of this research is the development of an affordable UV that is controlled through a cellular network. The primary difficulty that needs to be overcome to accomplish this objective is the continuous and small-delay packet transmission capability that is necessary to achieve accurate real-time control. However, as shown in the following section, adequate transmission qualities are not always available in a cellular network, and degradation of packet transmission qualities (referred as “delay spike”) can sometimes be observed. This degradation can cause the video to freeze and interfere with real-time remote control.

The cause of the delay spike in the mobile network is considered to be the processing time of ARQ (Automatic Repeat reQuest) or HARQ (Hybrid ARQ)<sup>8, 9)</sup>. Many studies on video and VoIP (Voice over Internet Protocol) have been conducted to suppress the degradation in voice quality caused by the interference of the delay spike in real-time communication.

To keep the transmission delay at a minimal value, reference 6) sets the transmission frame rate of the video to a

fixed small value. In this case, the video quality is restricted. Reference 7) treats specific traffic preferentially, and introduces control to keep the transmission delay at minimum, which is realized using additional scheduling functions that are required for all equipment in the communication network. To improve the quality of voice communication, it is effective to control the jitter buffer length on the receiving terminal side in consideration of the spike delay, thereby absorbing the discontinuity of the continuous signal due to the spike delay<sup>10, 11)</sup>. However, increasing the jitter buffer length leads to an increase in the delay time, which is not suitable for real-time control communication.

To counteract this problem, we believe that the use of multiple networks would be effective<sup>12)</sup>. More specifically, by applying a fast switching function that would allow the best quality network to be selected for the control signals, it is possible to reduce video freeze frequencies. Accordingly, in this paper, we propose a fast switching technique that can be applied to real-time video signals and describe the results of numerical evaluations that show its effectiveness.

The remainder of this paper is organized as follows: in Section 2, we show the results of transmission quality measurements involving the current commercial LTE network and discuss its applicability and problems that are related to UV control. Then, in Section 3, we provide a detailed overview of our proposal for a fast switching function that can be applied to a video signal of a UV control system.

The evaluation results of our proposal, including simulation and an LTE network experiment, are given in Section 4. These results show that continuous transmission of real-time video signals can be achieved efficiently through cellular networks.

The conclusions drawn from these results are described in Section 5.

## 2. Transmission quality

We will begin by examining the transmission quality of current commercial LTE networks to clarify whether they could be appropriately applied to real-time UV operations. Additionally, we will investigate the degree to which video quality is influenced by the transmission quality on these networks.

### 2.1 Transmission quality of cellular networks

To confirm the applicability of cellular networks for UV control, we first investigated the transmission qualities of LTE networks belonging to each of the three major mobile network carriers in Japan. To accomplish this task, we selected the packet loss ratio (PLR) and round trip time (RTT) parameters as transmission quality evaluation metrics and adopted 32 Android smartphones and one Web server for use in our test.

Each smartphone was carried by a student of our academy as he or she moved around our campus in Yokosuka City, Japan during a two-week evaluation period in December 2014. The server was connected to the Internet via a broadband optical access network service located in Kawasaki City, which is approximately 50

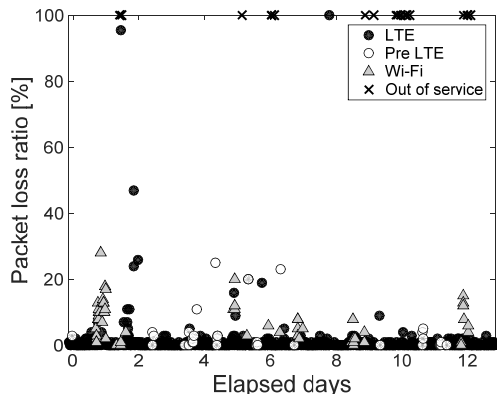
km away from the campus.

Each phone was equipped with in-house-developed software that automatically executed PLR and RTT measurements between each smartphone and the server. For one measurement, 100 ping packets with an interval of 200 ms were used. This measurement was repeatedly executed every 20 s during the evaluation period. The size of the ping packet was 32 B.

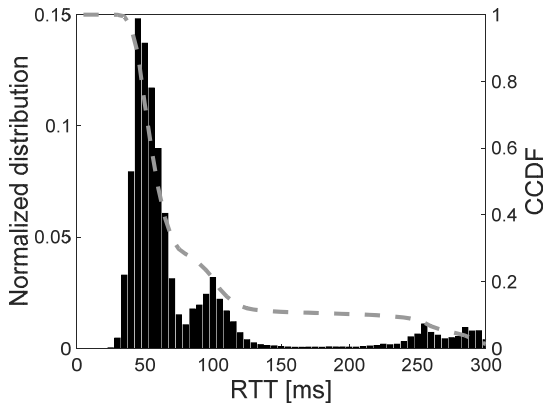
PLR measurement results for the entire measurement period are shown in Fig. 1, where the horizontal axis represents the elapsed days, and the vertical axis represents the PLR. In this figure, black circles, white circles, and triangles show the PLR of an LTE network, older generation cellular networks, and Wi-Fi, respectively. The cross marks indicate measurement times when no wireless connection could be established. The average value of PLR measurements of all smartphones throughout the measurement period was less than 1%, which means that we experienced no packet loss in most cases. However, UV control would have been lost at times when a large PLR resulted in degraded packet transmission quality, which is a problem that must be solved to achieve a reliable signal channel on a cellular network.

RTT measurement results are shown in Fig. 2, where the horizontal axis indicates RTT and the bin width is 5 ms. The left vertical axis indicates normalized RTT distribution. The dashed line indicates the complementary cumulative distribution function (CCDF) on the right vertical axis. In this figure, most of the RTT (approximately 75%) is distributed around

the 50 ms peak. Given our observation that less than 10% of the ping packets recorded RTT values of more than 200 ms, it would be difficult to apply a signal channel that requires RTT values of less than 200 ms to real-time UV controls.



**Fig. 1** PLR over the entire measurement period. Pre-LTE: older generation cellular networks.



**Fig. 2** Distribution of RTT over LTE at bin width = 5 ms.

Throughout the measurement period, we also confirmed more than 200 ms RTT periods that were observed somewhat less frequently in comparison to times when the packet-sending interval was set to 50 ms. This observation, which indicates that RTT characteristics depend on the interval

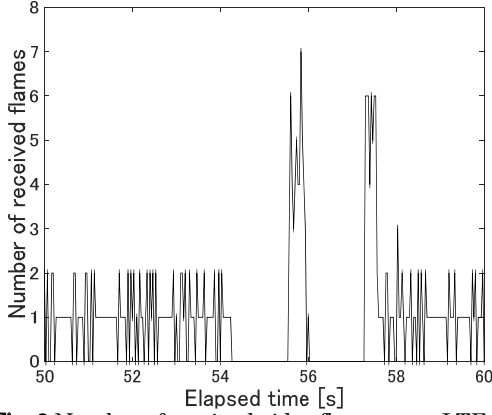
length of continuous packet arrivals, will be left for further study.

Given the knowledge that large packet transmission delays can cause sudden interruptions to continuous video streaming and degrade video quality, we hypothesized that this degradation could be minimized by using multiple networks simultaneously and adopting an adaptive channel switching function.

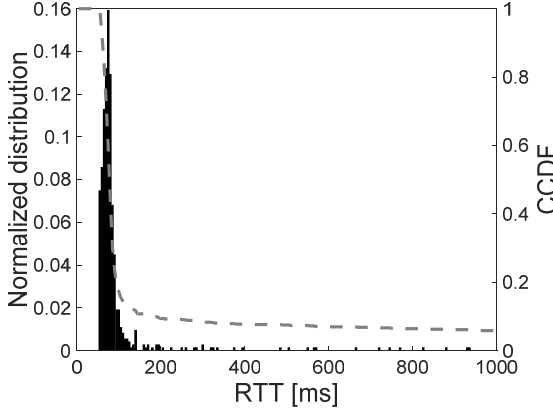
## 2.2 Video quality for UV controls

In this section, we show the evaluation results obtained during actual video frame transmission over a commercial LTE. Here, we adopted the Motion-JPEG (MJPEG) as the video coding method. Additionally, we set the resolution to quarter video graphics array (QVGA) and frame rate to 30 fps. The average bit rate was approximately 200 kbps. MJPEG uses the JPEG still-image compression algorithm and presents motion video by continuously displaying JPEG frames in the sequence. We determined that MJPEG would be most suitable for real-time UV control because of its small encoding and decoding times.

Figure 3 shows a typical example of unstable video signal packet arrivals. The graph indicates the number of incoming frames arriving every 33.3 ms. Here, despite most of the frames arriving smoothly, sometimes, a long interval would occur, after which a large number of frames will arrive at nearly the same time. We call this long interval (which can exceed 1 s) “delay spike.”



**Fig. 3** Number of received video frames over LTE.



**Fig. 4** Distribution of RTT while receiving video signals over LTE at bin width = 5 ms.

Figure 4 shows the RTT normalized distribution, which was measured with ping packets while transmitting an MJPEG signal. Here, significant proportion of the RTT is distributed over a period of less than 100 ms, but we can also confirm that seriously large RTT has occurred at a significant frequency. While these long delays in video signal transmission would critically degrade real-time UV operations, it should also be noted that, throughout these measurements, the measured packet loss ratio was very small (0.091%).

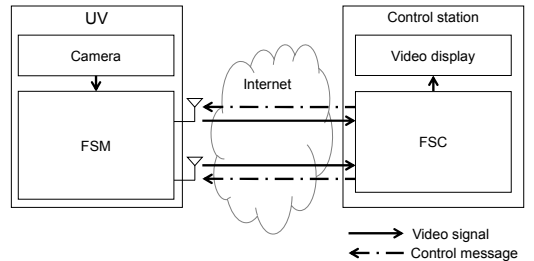
Based on the above evaluations, we

concluded that the delay spike must be overcome before it would be possible to realize reliable and real-time video signal channels.

### 3. Fast video signal switching

To decrease the delay spike, we propose the following fast switching technique, which works on multiple cellular networks. Currently, numerous devices have multiple wireless network interfaces. For example, smartphones can use several communication networks, such as LTE, 3G, Wi-Fi, and Bluetooth. Accordingly, given that it would not be difficult to equip a UV with interfaces for multiple wireless networks, we can improve the transmission quality of video signals by using multiple channels simultaneously. To accomplish this task, we developed a fast video signal switching function that adaptively selects the best available network.

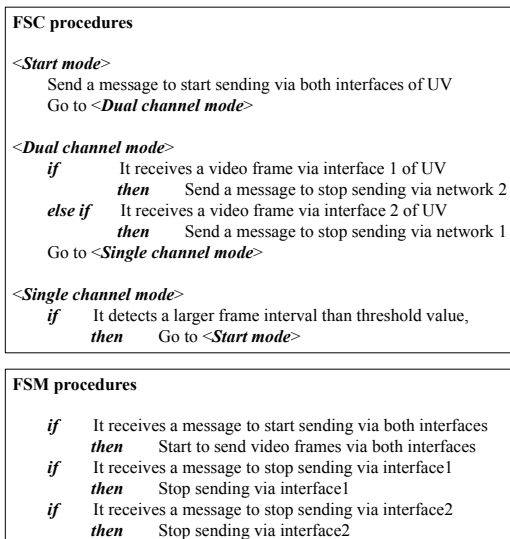
The configuration of our proposed scheme is shown in Fig. 5. In this scenario, a UV sends a video signal to a control station. Between the UV and a control station, we have set two independent cellular networks that are connected to the Internet. MJPEG frames are constantly transmitted, in sequence, from the UV to the control station through one of the two networks.



**Fig. 5** Conceptual diagram and video frames flow of the proposed technique.

The fast switching module (FSM) and the fast switching controller (FSC) functions are installed on the UV and control station, respectively. When the FSC detects a large interval in the arriving MJPEG frames, it immediately sends a channel switch instruction message to the FSM.

As described in Fig. 6, the FSC has three modes: Start, Dual channel, and Single channel. At the start of the operation, the FSC works in the Start mode and sends a message to the FSM instructing it to begin sending a video signal to both networks, after which it switches to the Dual channel mode. When the FSC is in the Dual channel mode and receives a video frame via one of the networks, it sends a message to the FSM instructing it to stop sending video to the other network and then switches to the Single channel mode. When the FSC is in the Single channel mode and detects an interval that is larger than the pre-determined threshold value, it returns to the Start mode.



**Fig. 6** FSC and FSM procedures.

To facilitate the detection of out of sequence frames, a sequence number is attached to each frame. If the FSC receives a frame out of sequence, that frame is discarded. FSM procedures are simple. It simply starts or stops video signal transmission based on the message orders received from the FSC. These procedures can be implemented as a running program in an application layer.

In our approach, we presumed that two independent active networks would always be available. Usually, cellular networks and mobile terminals are equipped with handover functions in the data link layer. Determining whether the proposed method would work properly via these handover functions is another important part of our future studies.

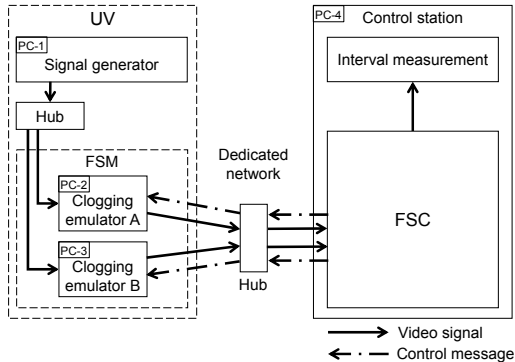
## 4. Evaluations

We first evaluated the potential effectiveness of our fast switch function via computer simulations. Then, to confirm the feasibility and applicability of our proposed scheme to an actual UV system, we conducted experiments using an approximately 1/10-scale model car and two Android (Nexus 7 2013) tablets that were equipped with the fast switching function.

### 4.1 Effectiveness of fast switch function

The simulation configuration is depicted in Fig. 7. The UV is modeled using three personal computers (PCs). PC-1 is dedicated to generating video signals from pre-recorded video images, which are then sent to PC-2 and -3 in parallel. PC-2 and -3, which must work together to achieve the

switch function, are also equipped with clogged network emulation programs that can be used to change the intervals between frames for the received video signals.



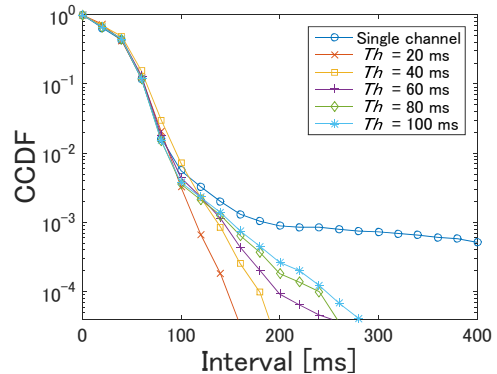
**Fig. 7** Simulation system setup.

A fourth PC (PC-4), which is equipped with the FSC function, was used to model the control station. We also measured the video frame arrival intervals from PC-4. All PCs were equipped with the Windows 8.1 operating system (OS), and communication was achieved via interconnected hubs and Gigabit Ethernet LAN cables. This simulation system was set up in our laboratory.

During the simulation, this LAN setup was used solely for our evaluation. Using this configuration, we measured all intervals between sequentially arriving video frames and confirmed that the RTT between each PC was always less than 1 ms.

Figure 8 shows the results of this evaluation. The horizontal axis indicates the interval frame reception intervals, while the vertical axis indicates the CCDF calculated from the distribution of recorded values for all intervals. By referring to records of actual frame transmissions through an

LTE network, the clogged network emulation programs reproduced delay spike characteristics. As a result of measuring the interval assigned to the packet transmitted from this program, it is confirmed that it is accurate in milliseconds, and the error rarely occurs within 10 ms. These independent records were used in the network emulators of both PC-2 and -3.



**Fig. 8** Complementary cumulative distribution function of video frame arrival intervals in our simulation.

In the figure, single channel characteristics were measured without the fast switching function. This shows an LTE network characteristic, where intervals larger than 200 ms have a probability of less than  $10^{-3}$ . Next, using the fast switching function described in the previous section, we evaluated the CCDF of the arriving intervals.

These characteristics depend on the threshold value ( $Th$ ). When the  $Th$  is less than 20 ms, the probability of interval occurrences larger than 150 ms is reduced to less than  $10^{-4}$ . Thus, we can observe that our proposed fast switching function can significantly decrease the probability of large video frame arrival intervals.



## 4.2 Implementation and Experiment with cellular networks

Next, we implemented the fast switching function on a prototype UV equipped Android tablets and a Windows PC serving as the control station. The prototype UV was controlled via the LTE, as shown in Fig. 9. To evaluate the effectiveness of our proposed method, we carried out a video transmission experiment using the prototype UV and the independent cellular networks of two major mobile network carriers in Japan.

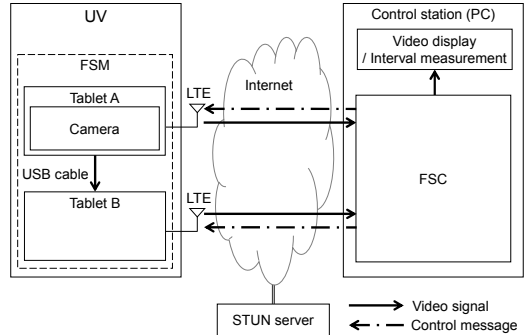


**Fig. 9** Prototype UV controlled via LTE networks.

Figure 10 shows the experimental configuration of this evaluation. Two Android tablets, a PC, and a Session Traversal of UDP through the NATs (STUN) server were connected via cellular networks and the Internet, as shown in the figure. The STUN server was required to allow the tablets to send User Datagram Protocol (UDP) datagrams to a PC outside the cellular network and the Internet<sup>13</sup>. We set the UV on the campus of the Defense Academy. The Control Station and the STUN server are connected to the Internet via LTE. The control station was set up in Yokosuka City outside the Defense Academy. The STUN server was

set in Kawasaki City.

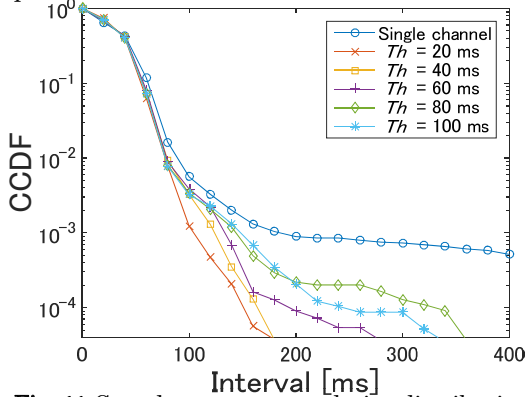
In this experiment, 30 fps MJPEG video signals produced by the camera in Tablet A were transferred to Tablet B via USB cable, and both transmitted the same signal to the control station simultaneously. Figure 11 shows the CCDF characteristics observed during this experiment. To facilitate a comparison of characteristics of the configuration of Fig. 10 with the characteristics shown in Fig. 8, we used the same pre-recorded video signals that were used in the simulation discussed in the previous section. This graph shows that in the actual cellular networks used, there was no significant reduction in the effectiveness between our actual experiment and the simulation.



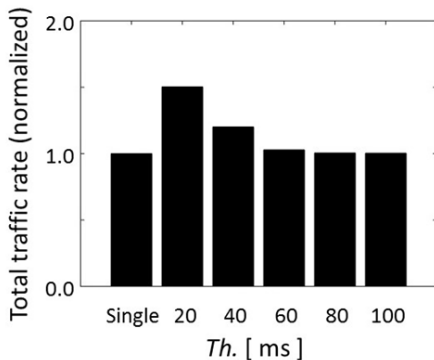
**Fig. 10** System diagram for the proposed implementation. Each tablet uses LTE provided by different carriers.

When the FSC is in a Dual Channel mode, the total packet traffic rate increases to twice the amount of the original video signals. In our experiment, given that we adopted 30 fps MJPEG video signals, the average interval of arriving frames was approximately 33.3 ms. Thus, when the  $T_h$  is set to be smaller than 33.3 ms, the FSC almost always switches to the Dual Channel mode before receiving a frame.

Then, when the FSC receives a packet from one of the channels, it switches to single channel mode. This behavior increases the packet traffic.



**Fig. 11** Complementary cumulative distribution function of video frame arrival intervals in LTE implementation.



**Fig. 12** Normalized data amounts when changing Th.

Next, we investigated the total traffic rate over the same network configuration shown in Fig. 10. Figure 12 shows the overall video traffic characteristics depending on the threshold set. Here, the vertical axis indicates the total traffic rate normalized by the original video signal rate. When we set the threshold to 20 ms, the total traffic rate increased by nearly 50%. However, when we use a 60 ms threshold, the increase was less than 3%, and its

impact on the network appeared to be insignificant.

## 5. Conclusion

In this research, we aim to realize an affordable UV operation system that uses cost-effective mobile cellular networks. We focused on a video signal sent from a camera on a UV to a control station. We began by evaluating the transmission quality of commercially available cellular networks and concluded that the "delay spike" must be overcome to realize reliable and real-time video signal channels.

To overcome this problem, we proposed the simultaneous use of multiple cellular networks and the installation of a fast switching function that adaptively selects the best available channel from multiple networks. Then, we evaluated our proposed method.

Our simulation results show that the fast switching function can significantly reduce the frequency of the delay spike. We also carried out a feasibility study experiment using Android tablets and two actual independent cellular networks. The results of these experiments confirmed that the proposed method is feasible and effective for realizing reliable real-time video signal transmission capability over actual mobile cellular networks. We can expect that the method is also effective for random packet losses on control channels.

The video transmission quality requirements will depend on the UV applications. Accordingly, faster and more precise UV movements will require higher quality videos. Therefore, quantitative evaluations of the transmission quality

requirements for specific applications will be another important part of our future study.

### Acknowledgment

This research was supported by the National Defense Academy's special research subject funds.

### References

- 1) K. Iwata, "Research of Cargo UAV for civil transportation", *J. Unmanned System Technol.*, 1[3] (2013), pp. 89-93.
- 2) G. Tuna, B. Nefzi and G. Conte, "Unmanned aerial vehicle-aided communications system for disaster recovery", *J. Network and Computer Applications*. 41 (2014), pp. 27 -36.
- 3) H. Xiang and L. Tian, "Development of a low-cost agricultural remote sensing system based on an autonomous unmanned aerial vehicle (UAV)", *Biosystems Engineering*, 108[2] (2011), pp. 174-190.
- 4) M. Bhaskaranand, and J. D. Gibson, "Low-complexity video encoding for UAV reconnaissance and surveillance", *Proc. of IEEE Military Communications Conference*, IEEE Communications Society (2011), pp. 1633 -1638.
- 5) P. E. Ross, "When Will We Have Unmanned Commercial Airliners?", *IEEE Spectrum*, <http://spectrum.ieee.org/aerospace/aviation/when-will-we-have-unmanned-commercial-airliners> (2011).
- 6) T. Suriyon, S. Higa, L. Wanamal, S. Nagamine, K. Fukuyama, T. Miyagi, T. Anezaki and K. Nakagawa, "Development of Monitoring System Using Multiple Flying Robots Controlled via 3G Network", *IEEE Transactions on Industry Applications*, 135[2] (2015), pp. 132 -137.
- 7) I. M. Delgado-Luque, F. Blázquez-Casado, M. Garcia Fuertes, G. Gomez, M. C. Aguayo-Torres, J. T. Entrambasaguas, J. Baños, "Evaluation of latency-aware scheduling techniques for M2M traffic over LTE", *Proc. of the 20th European Signal Processing Conference* (2012), pp.989-993.
- 8) 3GPP specifications, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures," TS 36.213 (2009), V8.8.0.
- 9) A. Gurtov and R. Ludwig, "Making TCP robust against delay spikes," University of Helsinki, Department of Computer Science, *Series of Publications C*, No C-2001-53, (2001), pp.1-17.
- 10) R. Yu, J. Yuan, Y. Xiao and G. Du, "Quality-based Jitter Buffer Algorithm Using Adaptive Variable-Size Window," *Proc. of 2012 2nd International Conference on Computer Science and Network Technology*, IEEE, (2012), pp. 151-157.
- 11) H. G. Kim and J. H. Lee, "Enhancing VoIP Speech Quality Using Combined Playout Control and Signal Reconstruction," *IEEE Transactions on Consumer Electronics*, 58[2] (2012), pp. 562-569.
- 12) N. Chiba, M. Ogura, R. Nakamura and H. Hadama, "Dual transmission protocol for video signal transfer for real-time remote vehicle control", *Proc. of the 20th Asia-Pacific Conference on Communicatinos* (2014), pp. 315-320.

- 13) J. Rosenberg, R. Mahy, P. Matthews, D. Wing, "Session Traversal Utilities for NAT (STUN)", *RFC 5389* (2008), <https://tools.ietf.org/html/rfc5389>.