HD Video | Remote Monitoring | Transcode

Technology Reports

Transcoder/Camera-control Gateway for High-definition Cameras toward High-quality Remote Monitoring from Mobile Terminals

The demand for remote monitoring by high-definition, high-quality video has been growing as high-definition surveillance cameras come on the market and mobile networks for transmitting that video become capable of higher speeds and larger capacities. We have developed a prototype high-definition surveillance-camera system to meet this demand using cameras capable of shooting high-definition video and a transcoder/camera-control gateway. This system will enable remote monitoring from mobile terminals by high-quality video on the level of HDTV.

Service & Solution Development Department

Takeshi Kato
Tomoyuki Osano
Norihiro Ishikawa

1. Introduction

A growing awareness of the need for crime prevention and disaster prevention has led to the widespread use of remote monitoring using video cameras. For crime prevention, surveillance cameras are installed in shopping districts, financial institutions, and public transportation facilities to monitor for and record intruders and suspicious individuals. For disaster prevention, cameras are installed alongside rivers, near volcanoes, and alongside roads to check for abnormal activity. In this way, remote monitoring has come to be used in a wide range of fields, but users are now expressing a desire to capture situations with even more detail by high-quality, high-definition video.

In response to these market needs, the development of high-definition surveillance cameras has been progressing and diverse products have come to be developed and commercialized. Surveillance cameras using a Video Graphics Array (VGA)[*1] at approximately 300,000 pixels or even Quarter VGA (QVGA) have been standard up to now, but recent years have seen the mounting of mega-pixel cameras on the order of 1-million pixels or more reflecting an increase in the number of surveillance cameras suitable for networks capable of delivering high-definition video.

At the same time, the pixel count of mobile-terminal screens has been increasing and the speed and capacity of the mobile network have been rising making it possible to achieve remote monitoring by mobile terminals over the network with a picture quality on par with high-definition video.

Compared with a traditional centralized monitoring system using monitoring desks, a remote monitoring system using high-definition video displayed on mobile terminals enables remote, high-definition monitoring from just about anywhere, which can reduce the number of personnel needed for

*1 **VGA**: An IBM-developed display-circuit standard whose typical display mode features a screen resolution of 640×480 pix and 16 colors. A screen resolution of 640×480 pix is often referred to as “VGA size” or simply “VGA.”

monitoring. With remote monitoring, monitoring personnel can be notified by e-mail of an emergency situation at some location simply by carrying around a mobile terminal. Then, if need be, they can proceed to the problem site while sharing high-definition video of the situation. This quick sharing of information at the time of an emergency–such as an incident or accident at a construction site or public facility or an overflowing river or erupting volcano–can be very useful.

In view of the trend toward high-definition video and high-speed, large-capacity mobile networks, we here present the results of studying several systems to determine the optimal one for enabling the control of high-definition surveillance cameras and the remote monitoring of captured video from mobile terminals. We also describe a prototype high-definition surveillance-camera system that we constructed for evaluating the system that we eventually proposed. With this prototype system, we aim to enable remote monitoring from mobile terminals by high-quality video on par with HDTV using high-definition surveillance cameras and a real-time video encoder.

2. Study on Achieving a High-definition Remote-monitoring System using Mobile Terminals

2.1 Requirements

In this study, we establish the following requirements for a remote monitoring system that combines high-speed, large-capacity FOMA-network mobile terminals and high-definition surveillance cameras.

- Quality

  Must support the viewing of video up to a display size corresponding to Wide VGA (WVGA) as installed in FOMA terminals
- Supported terminals

  Must enable the use of not only smart phones but FOMA terminals as well
- Camera control

  Must enable the user to control surveillance cameras as in pan[*2], tilt[*3] and zoom functions in addition to viewing video
- Extendibility

  Must be adaptable to larger screen sizes in future mobile terminals and to speed and capacity upgrades in the network

With these requirements established, we performed a study to search out an optimal system.

2.2 System Study

Original systems for delivering video to mobile terminals can be broadly divided into i-appli, videophone, and i-motion systems. We compared these systems in terms of the four requirements introduced above: quality, supported terminals, camera control and extendibility (**Table 1**).

1) Quality

Although the drawing region of an i-appli depends on the terminal, an i-appli is capable of displaying video up to WVGA (854×480 pix). The videophone system can display video up to Quarter Common Intermediate Format (QCIF : 176×144 pix) regardless of the model. The i-motion system can deliver video up to 30 frames per second (fps) by QVGA under MobileMPEG-4 (MobileMP4)[*4] Version 7.

2) Supported Terminals

The i-appli system targets terminals supporting Mega i-appli given the use of large quantities of video data. The videophone system supports all FOMA terminals with the exception of Raku-Raku Phone S. In the case of i-motion, FOMA terminals compatible with the FOMA high-speed system are required to handle unlimited streaming.

3) Camera Control

On evaluating the feasibility of camera control among the three systems, we found that the i-appli system can perform pan and tilt controls by packet communications and that the videophone system can control cameras by Dual-Tone Multi-Frequency (DTMF) signals. The i-motion system, meanwhile, provides for only video viewing–it cannot control the camera simultaneously.

4) Extendibility

Both the i-appli and i-motion systems allow for advances in picture quality according to the mobile-terminal's

*2 **Pan**: Moving the camera's orientation left or right.

*3 **Tilt**: Moving the camera's orientation up or down.

*4 **MobileMP4**: An MP4 video file format used especially for mobile terminals; also used as a 3GPP file format.

Table 1 Comparison of video delivery systems

	i-appli	Videophone	i-motion
Quality	Good High-definition video viewing according to display size	No good Up to QCIF (176×144 pix) regardless of model	Good Video delivery up to QVGA 30 fps; simultaneous delivery of audio also possible
Supported terminals	Good Models supporting Mega i-appli	Good All FOMA terminals excluding Raku-Raku Phone S	No good FOMA terminals compatible with the FOMA high-speed system (unlimited streaming)
Camera control	Good Pan and tilt control by packet communications	Good Control by DTMF (push) signals	No good Video viewing only
Extendibility	Good Picture quality can advance according to mobile-terminal screen size	No good Picture quality cannot advance in a timely manner with upgrades in screen size	Good Picture quality can advance but video can be viewed only on a native player

screen size, but the videophone system cannot upgrade picture quality in a timely manner according to screen size.

This system comparison against the above four requirements revealed that the existing videophone and i-motion systems have limiting aspects. We therefore decided to propose an original i-appli-based video delivery system as a means of controlling high-definition surveillance cameras and viewing high-definition video from mobile terminals.

3. Design of a High-definition Remote Monitoring System using Mobile Terminals

3.1 System Configuration

The configuration of the proposed high-definition surveillance camera system using mobile terminals is shown in **Figure 1**. This system consists of high-definition cameras, a transcoder[*5]/camera-control gateway (GW), and mobile terminals. The system transfers high-definition video from the surveillance cameras to the GW by H.264[*6] and transfers video between the GW and mobile terminals by Motion JPEG. The GW converts H.264 video to Motion JPEG in real time. The control protocol used for making the surveillance cameras pan, tilt and zoom is the P2P Universal Computing Consortium (PUCC) protocol [1].

3.2 Video Transfer System

Considering that the main purpose of this system is remote monitoring, we surmised that frame quality would have priority over detailed object movement and consequently adopted a video transfer system based on Motion JPEG. Furthermore, as an i-appli is limited to a maximum data-transfer size of 150 kB for every HTTP sequence, it became necessary to establish and release a HTTP connection when sending out frames. Thus, to improve frame rate, as many frames as possible are inserted and transferred for every HTTP connection (**Figure 2**). As shown in the figure, the system transmits multiple frames up to a maximum of 150 kB before releasing the connection. This sequence is repeated as long as there are more frames to transmit. Here, the number of frames that can be accommodated by one HTTP sequence varies according to picture quality, which means that a high frame rate can be selected for low picture quality and a low frame rate for high picture quality. In other words, picture quality and frame rate can be varied as needed according to the application and situation in question.

3.3 Camera Control Protocol

This system uses the PUCC protocol for camera control. Drafted by PUCC, this is an application layer protocol that provides interconnections between communicating devices via different lower-level transport networks. To control devices, the PUCC protocol makes use of device metadata written in eXtensible Markup Language

*5 **Transcoder**: Equipment or system having a function for converting video data to another video format.

*6 **H.264**: A coding system for moving-picture data capable of high rates of compression compared to MPEG-2 and MPEG-4. It can support a wide range of applications such as high-definition broadcasting.

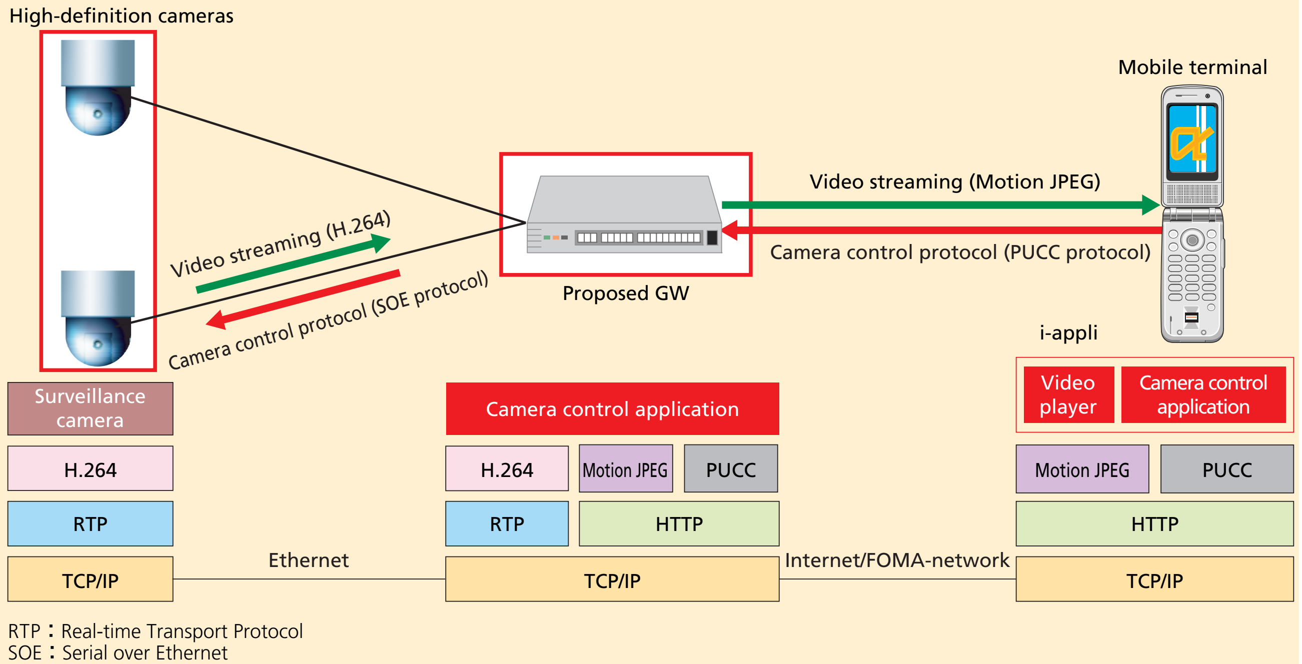

Figure 1 Configuration of high-definition surveillance camera system

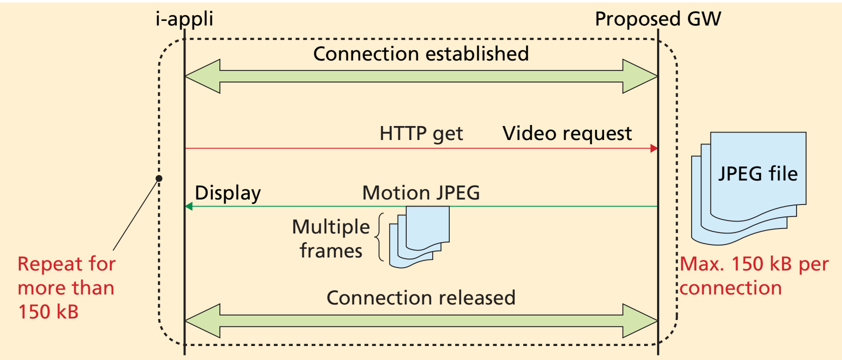

Figure 2 Video transfer system

(XML)[*7] describing device name, type, attributes and provided services. In this study, we designed device metadata for the high-definition surveillance cameras and the control protocol in accordance with PUCC specifications. Using PUCC protocol in this way enables differences in camera functions (e.g., zoom control can be performed but pan/tilt control cannot) to be described by device metadata, which means, in turn, that cameras with a variety of functions can be manipulated by the same application on the mobile-terminal side.

The PUCC protocol sequence is shown in **Figure 3** and summarized below.

- Step (1):

  Perform connection processing and authentication from the i-appli on the mobile terminal side to the proposed GW and get SessionID (Hello Message).
- Step (2):

  Get metadata for camera devices targeted for connection (Discover Message). This action enables a list of cameras to be obtained.
- Step (3):

  Get information on the video delivery systems supported by each camera (send GetSupportedMonitoringMethods command by the Invoke message).
- Step (4):

  Get camera usage right for cam-

*7 **XML**: A markup language for describing information such as documents and data using a hierarchical structure of elements called tags.

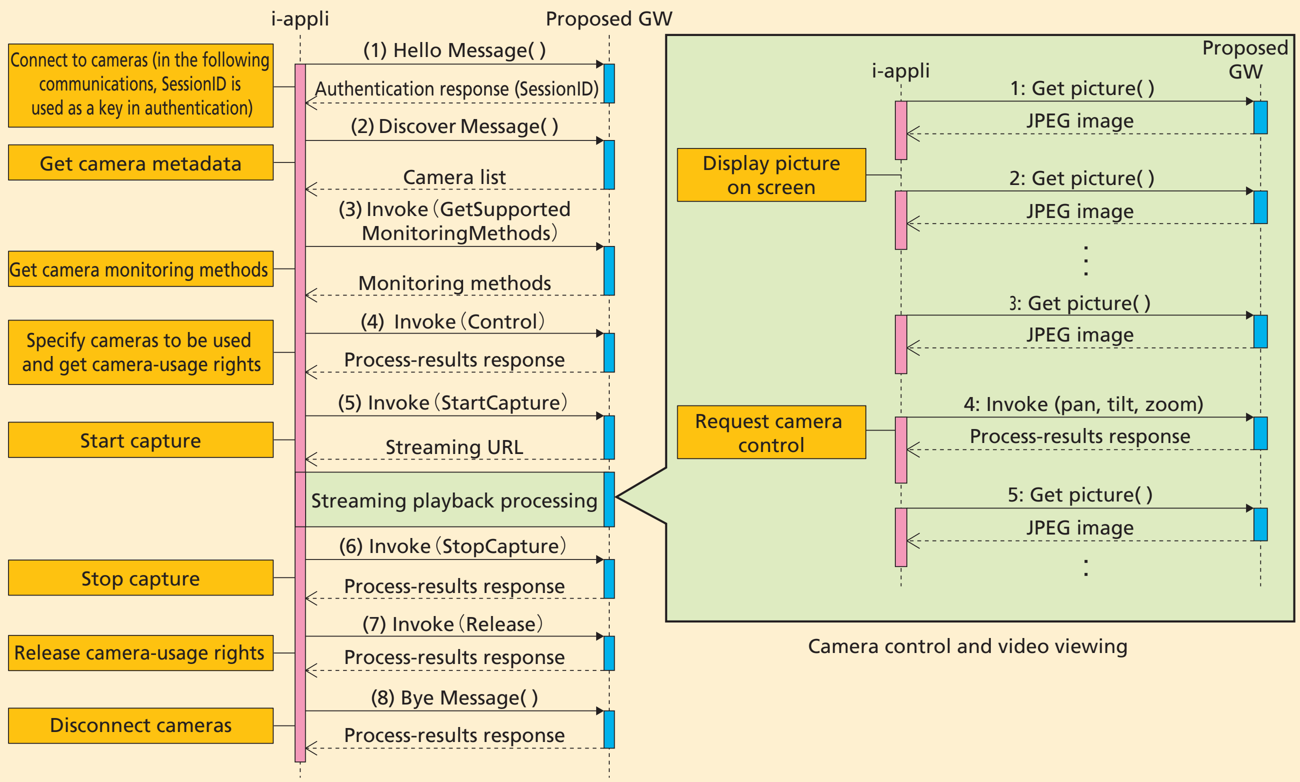

Figure 3 PUCC protocol sequence

era device selected by the user (send Control command by the Invoke message).

- Step (5):

  Request video capture to begin (send StartCapture command by the Invoke message). The GW returns a URL as a response to the Invoke message and the i-appli requests the GW to get Motion JPEG with respect to that URL. This action transfers the camera video and displays it on the mobile terminal. The cameras can be controlled at this time by camera-control requests (i.e., by sending pan, tilt and zoom commands by the Invoke message).

- Step (6):

  Request video-capture stop-processing to terminate viewing of camera video (send StopCapture command by the Invoke message).

- Step (7):

  Release camera usage rights (send Release command by the Invoke message).

- Step (8):

  Perform disconnect processing from the i-appli on the mobile terminal side to the GW (Bye Message).

4. Implementation of Prototype System

We implemented a prototype system based on the system design described in the previous chapter. An overview of this system is shown in **Figure 4**. The cameras used here are surveillance cameras capable of outputting H.264 high-definition video with an output picture size of 1,280×720 pix. The application on the mobile-terminal side was implemented using a Mega i-appli. The proposed GW incorporates an application for converting 1,280×720 pix H.264 high-definition video to WVGA-equivalent 832×468 pix Motion JPEG in real time.

The screen-transition sequence is shown in **Figure 5**. After starting up the i-appli, connecting to the GW, and

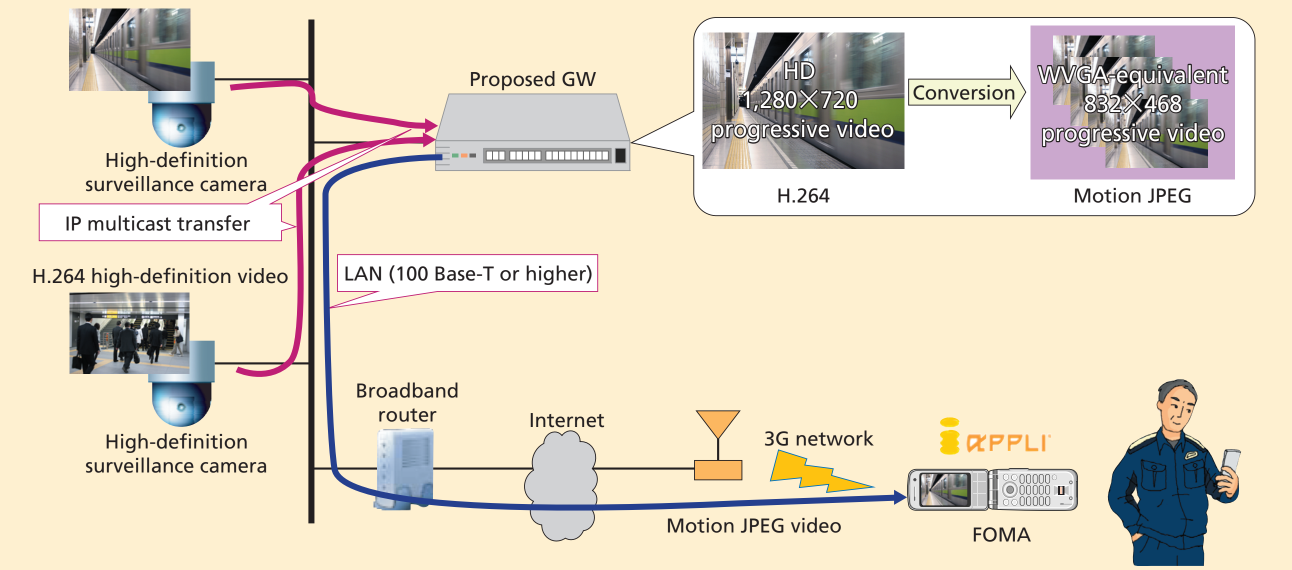

Figure 4 Overview of prototype system

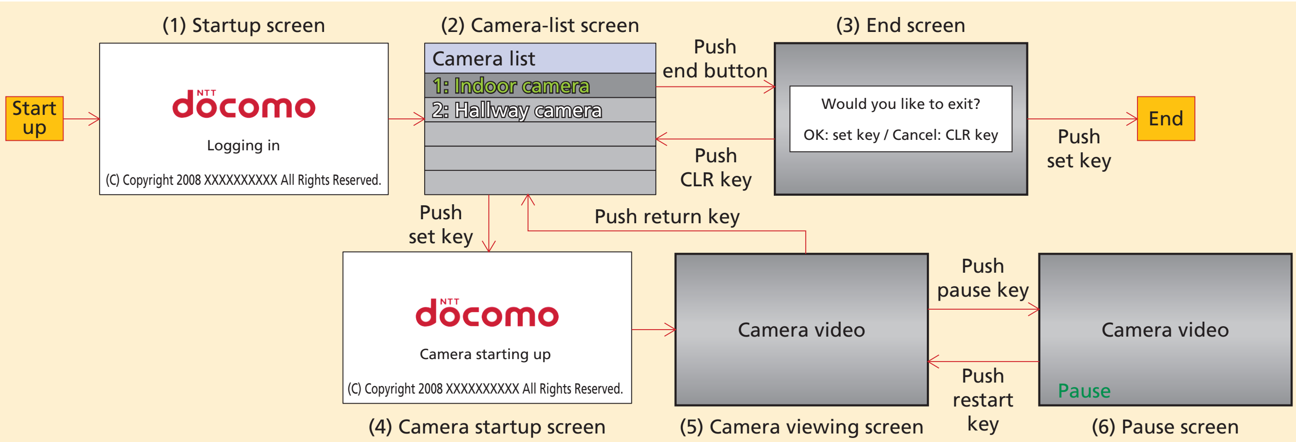

Figure 5 Screen transitions

getting camera metadata (Fig. 5 (1)), a list of available cameras is displayed (2). The user can now select a camera thereby initiating video-capture processing in the camera startup screen (4) and enabling the viewing of surveillance camera video in real time (5). For example, the user can view Motion JPEG video matching the screen size of the mobile terminal (**Figure 6**).

Implementation of this prototype system demonstrated the feasibility of viewing high-definition, high-quality surveillance camera video from mobile terminals.

Figure 6 Video screenshot

5. Conclusion

This article presented the design of

a high-definition surveillance camera system for achieving high-definition, high-quality remote monitoring by mobile terminals and described the implementation of a prototype system. Looking to the future, we plan to study systems that will enable the viewing of even higher quality video as mobile networks increase in speed and mobile-terminal screens reach higher levels of resolution. We also plan to study ways of improving the video frame rate by User Datagram Protocol (UDP) communications using the i-appli Star profile.

Reference

[1] N. Ishikawa et. al: "Home Network Control from Mobile Terminals for Creating New Use Scenarios of Life-Style Mobile," NTT DoCoMo Technical Journal, Vol. 10, No. 1, pp. 26-36, Jun. 2008.