Future Generation Computer Systems 27 (2011) 986–990

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Multi-point 4K/2K layered video streaming for remote collaboration

Daisuke Shirai [a,*], Masahiko Kitamura [a], Tatsuya Fujii [a], Atsushi Takahara [a], Kunitake Kaneko [b], Naohisa Ohta [c]

[a] NTT Network Innovation Laboratories, NTT Corp., 1-1 Hikari-no-oka, Yokosuka, Kanagawa, 239-0847, Japan

[b] Research Institute for Digital Media and Content (DMC), Keio University, West Annex, 2-17-22 Mita, Minato-ku, Tokyo, 108-0073, Japan

[c] Graduate School of Media Design, Keio University, 4-1-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa, 223-8526, Japan

ARTICLE INFO

Article history:
Received 29 March 2010
Received in revised form
31 August 2010
Accepted 29 November 2010
Available online 9 December 2010

Keywords:
Applications
Multi/mixed media
Systems development
Network communication

ABSTRACT

The progress of technologies in image industries and the rapid spread of broadband networks allow us to put high quality video communications to use for remote collaboration. We have developed a 4K/2K layered video streaming system with JPEG 2000 codec. It offers three main advances; support of various video formats to realize a media exchange platform, 4K/2K layered streaming with multicast to utilize 4K images even if the environment has no 4K equipment, and robust transmission using a forward error correction (FEC) scheme that can deal with bandwidths beyond 800 Mbit/s. Using the system, we conducted the world's first multi-point "shared collaborative space" experiment with a large audience in CineGrid International Workshop 2008. It proved the feasibility of the new environment for 4K remote collaboration between multiple points including those that did not have 4K equipment. This paper details the system and implemented technologies.

1. Introduction

The rapid progress in image technologies and the widespread adoption of high bandwidth networks permit us to realize exciting new high quality image communications applications. Those applications allow us to collaborate with remote partners by destroying the barrier of distance. When only small bandwidth streaming services such as 64 kbit/s were available, the best video communication could do was to make it known who was speaking. After that, it became possible to use standard definition (SD, e.g. 720×480 pixels) quality for video communication, which allowed people's expressions and emotions to be conveyed by video. Recent broadband infrastructures allow high definition (HD, e.g. 1920×1080 pixels) video to be transmitted over networks. HD enables an exchange of the detailed appearance of the remote partner and the atmosphere of the other room.

For very high quality motion picture production and media-intensive entertainment, education and research applications, the target of interest is shifting to the so-called "4K" imaging with more than 8 megapixels per frame. We have integrated 4K imaging technologies into video communication to realize an environment that enables the realization of the "shared collaborative space" on the public network, as one application of computer-supported cooperative work (CSCW). 4K images have 4096 or 3840 horizontal pixels and 2160 vertical pixels. 4K offers roughly quadruple the number of pixels of the HD television format, and 24 times that of an SD TV signal. 4K is a particularly significant new image format because it will be widely used for digital cinema theatrical distribution under new specifications proposed by Digital Cinema Initiatives, LLC, a consortium of the major Hollywood studios, and has been standardized by SMPTE (Society of Motion Picture and Television Engineers) [1,2].

One of the key technologies demanded for 4K video communication is real-time transmission via public networks. We have developed the world's first 4K real-time video streaming system with a JPEG 2000 codec [3]. The bit rates of uncompressed 4K video data range from 6.3 Gbit/s (4:2:2/10 bit color, 24 fps) to 9.5 Gbit/s (4:4:4/12 bit color, 30 fps). The codec compresses them in real-time to permit 4K data to be transmitted via common gigabit bandwidth networks. In 2005, we conducted the world's first real-time, international transmission of 4K digital cinema and 4K video at iGrid 2005 [4]. The demonstrations proved that the streaming system makes it feasible to implement networked professional audio/video applications, even at 4K quality, on IP networks over distances of up to 15,000 km.

Our next goal is establishing a 4K-based shared collaborative space between multiple points, even if one site has no 4K shooting/projection devices. For this, we newly developed a video streaming system that has a 4K/2K layered streaming function, supports various standardized video formats, and has a robust forward error correction (FEC) function with low density generator matrix (LDGM) codes.

In this paper, we detail the 4K/2K video streaming system, the FEC function, and the world's first experiment conducted in

* Corresponding author. Tel.: +81 46 859 5143.
E-mail address: shirai.daisuke@lab.ntt.co.jp (D. Shirai).

0167-739X/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.future.2010.11.021

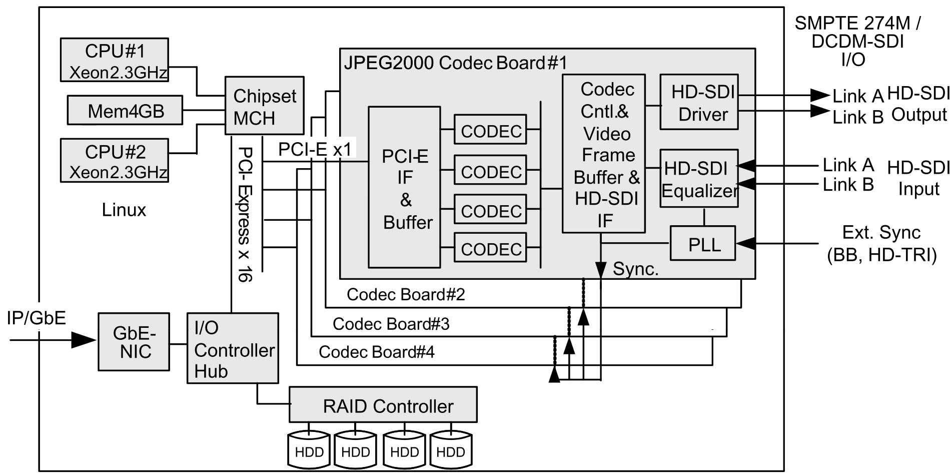

Fig. 1. Block diagram of the video streaming system.

CineGrid International Workshop 2008 [5]. The results of which showed the feasibility of the multi-point shared collaborative space for remote collaboration between 4K-equipped environments and 2K-equipped environments.

2. 4K/2K layered video streaming system with JPEG 2000

The 4K/2K layered video streaming system uses the JPEG 2000 coding algorithm. The codec compresses 4K video to 200–800 Mbits/s without degradation of subjective image quality, which allows us to transmit such high resolution video via common gigabit IP networks; the bandwidth of 4K uncompressed video is more than 6 Gbits/s. The system uses UDP/IP for video streaming over long distance networks and to multiple destinations, i.e. multicast.

The 4K/2K layered video streaming system employs parallel JPEG 2000 processing elements to encode and decode 4K/2K video. We selected Analog Devices Inc. ADV202 as the JPEG 2000 processor. The system consists of 2 circuit blocks, a PC/LINUX part with GbE interface, and newly developed JPEG 2000 codec boards, each of has four JPEG 2000 processers. A total 4 boards are installed on the PCI-Express bus. The system can process 4K images of 36-bit RGB color (4:4:4) at 30 fps or 30-bit YCbCr color (4:2:2) at 60 fps by running those boards in parallel.

Fig. 1 shows the block diagram of the system. The 4K video signal is input/output via multiple HD-SDI. Each codec board has two HD-SDI inputs and two outputs that are compatible with SMPTE 292 (Single Link). These inputs and outputs also can work as an SMPTE 372M (Dual Link) compatible I/O by using all of them simultaneously. The board also has a reference signal input for NTSC black burst and HD tri-level sync for synchronization with external devices.

The input image is separated into its color components of RGB or YCbCr. Each component is sent to a different JPEG 2000 processer to be encoded or decoded in parallel. The system switches the pattern of data flow for the divided 4K image depending on the signal format of the image. In the case of 4K RGB or YCbCr 4:4:4 color at 24 or 30 fps, the image is input via four dual link HD-SDI ports and each of the four codec boards processes a quadrant of the 4K image. Each of the four JPEG 2000 processers on the board processes one component of R, G or B (or Y, Cb or Cr). In the case of 4K YCbCr 4:2:2 color at 24 or 30 fps, the image is input via four single link HD-SDI ports and the system processes the input as described above. It can also be processed by just two of the four codec boards by inputting two quadrants of the 4K image to one codec board and using one of JPEG 2000 processers to process Y or CbCr components. In this case, the system is capable of bidirectional 4K (4:2:2 24 or 30 fps) streaming. In addition, it can support 60 fps of 4K 4:2:2 video in a similar fashion, except that each board is dedicated to processing one quadrant of the 4K image.

Table 1
Specifications of the video streaming system.

Coding method	JPEG 2000
Video I/O interfaces	HD-SDI (SMPTE 292)× 2ch × 4 blocks
Video format	SMPTE 274M, 372M 1920 × 1080 : 24p, 24PsF, 30p, 30PsF, 60i, 60p 3840 × 2160 : 24p, 24PsF, 30p, 30PsF, 60i, 60p DCDM-SDI (SMPTE 428) 2048 × 1080 : 24p, 24PsF 4096 × 2160 : 24p, 24PsF The 1/1 and the 1/1.001 divisor clocks where applicable
Color space/bit depth	RGB 4:4:4/12 bit YCbCr 4:4:4/12 bit YCbCr 4:2:2/10 bit
Audio format	Embedded Audio (SMPTE 299M) Uncompressed PCM 44.1/48 kHz, 16/24 bit, 32ch
External reference clock	NTSC Black Burst, HD Tri-level sync, or HD-SDI
Codec LSI	ADV-202
Wavelet filter	9/7I, 5/3R
OS	Linux kernel 2.6.x
Transmission protocol	TCP/IP, UDP/IP (Multicast supported)

Table 1 shows supported video formats and other specifications. This video streaming system supports such a wide variety of video formats that it realizes a media exchange platform that supports all kinds of contents including movies, computer generated graphics, scientific visualization, a video shot by streaming camera, and so on.

We employed the scalability of JPEG 2000 to establish the scalable streaming function (4K/2K multi-resolution); it allows 4K video sources to be utilized even in 2K-only environments. The process flow of JPEG 2000 coding is as follows. The input image is decomposed into subbands by using discrete wavelet transform (DWT). Each subband is divided into code-blocks, the coding unit of the EBCOT algorithm. The EBCOT algorithm realizes various levels of scalability. There are four basic scalability dimensions in a JPEG2000 code stream: resolution (R), quality (L), spatial location (P), and component (C). Different scalability levels are achieved by ordering packets within the code stream.

Fig. 2 shows an overview of 4K/2K layered multicast. Layered multicast is one of the approaches to achieve optimized multicast in the existing heterogeneous multicast environment [6]. We expanded this approach to the world's first real-time 4K layered multicasting based on the scalability of JPEG 2000.