

Proposal of Next Generation HDTV and Adaptive Video Distribution Model with Tile Structure of JPEG2000

Kenji TANAKA^{*/**}, Yasunari SUZUKI^{*}, Yoshihiro ICHIOKA^{**}, Terumasa AOKI^{***}

and Hiroshi YASUDA^{***}

^{*} Communications Research Laboratory
4-2-1, Nukuikitamachi, Koganei, Tokyo, 184-8795 Japan

^{**} Stanford University

^{***} University of Tokyo

4-6-1, Komaba, Meguro-ku, Tokyo, 153-8904 Japan

E-mail: ken@crl.go.jp

Abstract In Japan, various broadcast form such as terrestrial broadcasting, satellite broadcastings and CATV, are now being converted from analog to digital. In fact, all domestic broadcastings will have become digitized by the summer of 2011, and it is expected that this will promote TV's multi-channelization as well as increase the numbers of interactive programs and High-Definition TV (HDTV) programs. On the other hand, High-speed network services are already becoming available to general households, and, when the "Fiber to the Home (FTTH)" is realized at the end of fiscal year 2005, over 7.7 million households (1/6 of the entire national households) are believed to sign up for the service and enjoy approximately 100Mbps high-speed transmission. Thus, just as the United States and European countries may do, we intend to achieve the full-digitization of broadcasting, which requires the quality of HDTV, while generalizing high-speed networks for individual viewers.

Thus, we authors are aware of the further spread of the high-definition system and have developed the following two (next generation) formats that can provide even higher visual precision: one is WHD format (Wide/Double HDTV:3840*1028 pixels) that has twice as much pixels as existing HDTVs in the horizontal direction. The other is QHD format (the Quadruple HDTV:3840*2048 pixels), which has twice as much pixels in both vertical and horizontal directions and accomplishes the precision equivalent to or beyond that of 35mm films.

This paper describes the overview of the concept of WHD/QHD camera and projector that we have been researching and developing. We also would like to propose an image distribution model as one of the applications of WHD/QHD images for homes. Using the tiling function of JPEG2000, this model offers the contents which are effective, high precision and adapted for a viewers' needs.

Key words HDTV, WHD, QHD, JPEG2000

1. Introduction

As the schedule of Figure 1 shows, Japan has been conducting the digitization of various types of broadcasting facilities including satellite, terrestrial broadcastings and CATV. BS performed a trial digital broadcasting at the 27th Olympic games held in Sydney in September, 2000 and commenced the actual services (three months later) in December. On the other hand, CS broadcasting, which was launched in 1996, has learned from BS's example and will be digitized in 2002. Terrestrial broadcasting will also be digitized limitedly in and around Tokyo, Nagoya and Osaka by the end of 2003 and then go nationwide by the end of 2006, terminating its analog system as of July 24, 2011. It is expected that this will promote TV's multi-channelization as well as increase the number of programs that are designed for interactive programs. Besides, high-definition transmission system will be applied not only for TV production, such as shooting or editing, but also for broadcasting to average households, replacing the conventional NTSC format. Under these conditions, HDTV is expected to explosively spread among the general public in a short while.

On the other hand, with many carriers joining and struggling in the market, the costs for the high-speed network services of several Mbps on xDSL method have become drastically lower for individual use. Besides, the nationwide fiber-laying project called "Fiber to the Home (FTTH)" is being proceeded by the Ministry of Public Management, Home Affairs, Posts and Telecommunications; when this is completed at the end of fiscal year 2005, over 7.7 million households (1/6 of 47million national households) are believed to sign up for the service and enjoy approximately 100Mbps high-speed transmission.

Updating broadcasting system is proceeded not only in Japan, but is a worldwide trend. In fact, it is assumed the United States and several European countries will digitize their broadcasting facilities while generalizing HDTVs and high-speed networks.

Thus, we are aware of the further spread of the high-definition system and have developed the following two (next generation) formats that can provide even higher visual precision: one is WHD format (Wide/Double HDTV:3840*1028 pixels) that has twice as much pixels as existing HDTVs in the horizontal direction. The other is QHD format (the Quadruple HDTV:3840*2048 pixels), which has twice as much pixels in both vertical and horizontal directions and accomplishes the precision equivalent to or beyond that of 35mm films.

This paper describes the concept of WHD/QHD cameras and projectors that we have developed. Also mentioned is an image delivery model for the Internet, which, employing the tiling function of JPEG2000, effectively delivers high-precision contents that meets viewers' individual needs. This model should fulfill the requirements of next-generation contents making, which is to low-costly create quantity of attractive contents that entertain viewers with high-quality image and interactive communication system.

2. Next generation high definition transmission system

Figure 1 compares WHD/QHD formats with HDTV and SXGA, which are the standard formats for TV broadcasting and PC-monitors respectively. WHD format consists of twice more pixels than HDTV and two-and-a-half times more pixels than SXGA, while QHD consists of four times more pixels than HDTV and six times more pixels than SXGA.

Table 1: Shift schedule to digital broadcasting (in Japan).

Communication Satellite	
Broadcasting Satellite	
Terrestrial Broadcasting	
CATV	

- - - ● Analog broadcasting
● —> Digital broadcasting

2.1 WHD image transmission system[1]-[3]

WHD image transmission system is composed of WHD camera, CODEC and WHD projector.

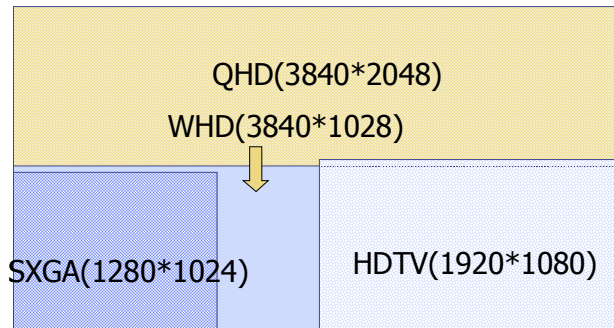


Fig. 1: Comparison of format.

2.1.1 WHD Camera System

Figure 2 shows the composition of the WHD camera system, which consists of two HDTV cameras (Real HD Camera-1, Real HD Camera-2), angled by 36.5 degrees and two mirrors. By connecting images from two conventional HD cameras in sideways, the system compose a virtual WHD camera with the view of 73 degrees. Using commercial cameras, this system was not too expensive to set up; however, it was hard to adjust each factor, such as focus, aperture, zoom, white-balance, location and direction between two independent cameras. Besides, as the cameras were not placed right in front of the object to shoot from oblique directions, when they shoot the soccer stadium, the screen images looked shrunk and distorted along both right/left side-edges (see Figure3 (a)). However, once the distortion was adjusted just

as shown in Figure3 (b), a perfect image was obtained.

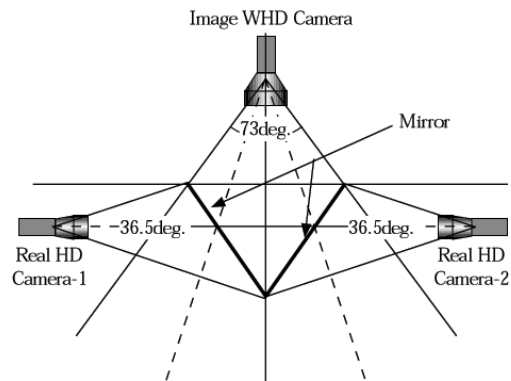


Fig. 2: Configuration of WHD camera.



(a) Before correction.



(b) After correction.

Fig. 3: Correction of distortion on WHD images.

2.1.2 WHD Projector

WHD projector consists of the followings: three display devices which are 1.3 inch of 3.95 million pixels (3840*1028) with the structure of the Direct-Image-Light-Amplifier(D-ILA), an optical block, which converts pixel aspects, enlarges and projects images and a front project screen of which size is 243.5 inch (6m*1.5m) and 165 inch (4 m*1m). Table 2 describes the specifications of WHD projector.

Table 2: Specification of WHD Projector.

Image Device	3*D-ILA (1.30 inch)
Resolution	3,840*1028 pixels
Lamp	750 W Xenon
Brightness	400 ANSI Lumens
Contrast Ratio	500:1
Input Signal	2*YPbPr

2.1.3 WHD image transmission experiments

Using two sets of MPEG2 CODEC for HDTV, we performed experiments to transmit WHD camera

images to a remotely located WHD projector through the Japan Gigabit Network (JGN)[4]. Figure 4 (a) shows the remote transmission experiment of soccer games, while Fig. 4 (b) shows the telemedicine experiment of brain surgery.



(a) Soccer.



(b) Neurosurgeon.

Fig. 4: WHD images transmission experiments.

2.2 QHD image transmission system

QHD image transmission system consists of a QHD camera, CODEC and a QHD projector.

2.2.1 QHD Camera

As pre-mentioned, WHD camera system is the compound of two individual cameras, and, therefore, it requires certain image correction to erase distortion or a seam line as well as careful adjustment in locating cameras. It was, in fact, not handy enough, and this has led to the development of QHD camera. This camera can manage moving images of 8.5 million-pixels ($3,840 * 2,160$), which are four times as high-resolution as those by HDTV with three image pickup chips corresponding to each RGB color.

In developing QHD camera, we decided to mount a small, light 35mm camera lens that could be suitable for any environment or purposes, such as; in/outdoor setting, to be used as an operation microscope or astronomical telescope. CMOS sensor, a light weight imaging device, was employed for its speedy and low power performance. Although it is generally said that its S/N ratio is not sufficiently qualified, we developed a special signal-amplification system to cover the defect.



Fig. 5: Outlook of QHD Camera.

Eventually, as shown in Figure 5, we succeeded in developing a QHD camera, which is about the same size as most studio cameras. Because this camera employs four channel HD-SDI (Serial Digital Interface) for its image output, it can be applied for VTR, editing or transmission devices for conventional HDTV.

2.2.2 QHD projector

Specifications of QHD projector can be seen in Table 3. This projector boasts twice more pixels than the formerly mentioned WHD projector does as well as high-luminance light, newly-employed square pixels and the design enhancement of optics and signal processing system. These improvements have enabled the projector to acquire essential qualities as a display device, including brightness, color reproducibility and high precision display. Appearance of the QHD projector is in Figure 6.

Table 3 Specification of QHD Projector.

Image Device	3*D-ILA (1.7 Inch)
Resolution	3,840*2,048 pixels
Lamp	1,600 W Xenon
Brightness	5,200 ANSI Lumens
Contrast Ratio	750:1
Input Signal	3840*2048/60p LVDS 4*HD-SDI/60i/30p/24p 8*HD-SDI Dual-link/60p 4*DVI



Fig. 6: Outlook of QHD Projector.

2.2.3 QHD image transmission

As shown in Table 4, WHD/QHD images can be enormous amounts of data streams of as large as 3Gps or 6Gps respectively without any compression done. Using the coding method suggested in this paper (JPEG 2000), images look uncompressed when they are actually compressed down to 1/10; this means the method helps images to maintain their original precision after compression. The compressed WHD/QHD images sizes, which are respectively 300Mbps and 600Mbps, are small enough to be transmitted over high-speed networks, such as GbE or OC-12. Currently under consideration is the real-time transmission by CODEC.

Table 4: Amount of data stream.

	Non - compression	1/10 compression
QHD	6 Gbps	600 Mbps
WHD	3 Gbps	300 Mbps
HD	1.5 Gbps	150 Mbps
SXGA	1 Gbps	100 Mbps
NSTC	250 Mbps	25 Mbps

3. Image Distribution Model

3.1 Outline

It is not plausible that high-speed networks of several 100Mbps will prevail among households within a couple of years. Besides, despite the current increase of high-vision programs and spreading HDTV monitors, it will take a while for WHD/QHD monitor to be commonly used. Thus, suppose the upper limit of transmission speed was 100Mbps (which, as mentioned in the first chapter, is what 1/6 of the entire Japanese households would be on after several years), the maximum throughput would be kept as small as

dozens of Mbps. Table 4 indicates that, with 1/10 compression by JPEG 2000, images with a display area around NTSC can be lightened for effective transmission while maintaining their original precision.

3.2 Viewer's needs

Several viewers watching the same TV program may actually be watching different parts of the screen. Figure 7, the same WHD image of a soccer stadium shown in Figure 3 (b), includes most of the essential objects in a soccer telecast; the whole view of the ground, benches for both teams and spectators' seats. Provided three viewers are watching this program at the same time, each of them may have different interests as indicated by the three squares (=NTSC size) in Figure 7; one of them wants to watch conventional TV program of ball chasing and another, a soccer apprentice, wants to study defense strategy while the other is crazy for a famous Paraguay GK called Chilavert. In other words, by providing selected images based on each viewer's interest, broadcasting can truly become satisfactory. This also helps reduce excess information on networks as well as unnecessary burden of decoding process at viewers' terminal such as STB(Set Top Box) or PC.

3.3 JPEG2000

JPEG2000 is an encoding technology for which a call for development contributions was launched in July 1997. It has been standardized by the SC-29, which belongs to the JTC-1 for both ISO and IEC; eventually, part of the method (Part 1:JPEG2000 Image Coding System) became approved standard in January 2001. Unlike conventional JPEG method (Baseline) that uses DCT (Discrete Cosine Transform) and Huffman coding, JPEG2000 standard employs Wavelet transform and block-based entropy/arithmetic coding methods to realize lossless vision under the compression of 1/10. Even with higher compression added, only subtle visual-loss, was observed, compared to JPEG or other method with DCT conversion. Among all the JPEG2000 options, we especially appreciate the tiling function that allows the division of a screen image during encoding; Figure 8 shows the code stream when a image is divided into n tiles.

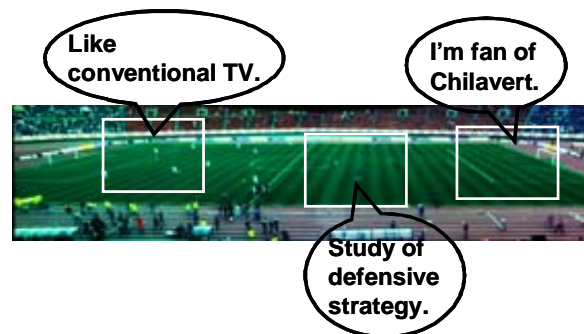


Fig. 7: Viewer's demands (on WHD image).

Header
Tile #1
Tile #2
⋮
Tile #n
EOC

Fig. 8: Code Stream with tile function.

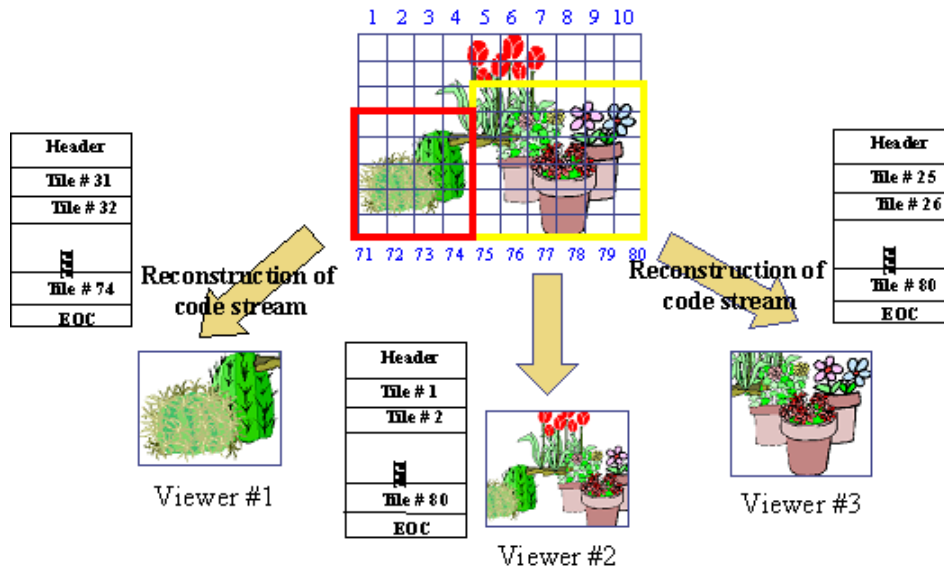


Fig. 9: Concept of adaptive video distribution model.

3.4 On-demand image distribution model

Figure 9 shows the overview of the delivery model that this paper suggests. The image is divided by 10x8 into 80 tiles. The viewer #2, who wishes to see the whole picture, receives the code stream that consists of all the tiles; whereas, for the viewer #1 and #3 who want to watch the left and right bottoms respectively, only the limited amount of tiles are needed to fulfill their requests (for the viewer#1 – 20 tiles between tile #31and #74 and for the viewer #3 – 36tiles between tile #25 and #80); the chosen tiles compose their own code streams and send specific images to those viewers. Hence, the whole picture of the stadium is no more needed when only 25 % or 40 % of it can satisfy the viewer #1 and #3.

3.5 Evaluation of Proposed Mode

3.5.1 Overview

In order to efficiently compose new code streams for individual viewers, it is effective to split a screen into small tiles while encoding. However, the number of tiles inevitably increases when each piece is small, and, it is feared that this may lead to the increase of each tile's header information, which deteriorates the

quality of images. Hence, the subjective evaluation of the relativity between tile sizes and image quality are examined as follows using a static image of QHD format.

3.5.2 Evaluation Image

The QHD image shown in Figure 10 has been compressed by JPEG2000 (JasPer Ver. 1.5) and transformed into image data. Three types of compression ratio (1/5, 1/10 and 1/20) and four tile sizes (3840*2048, 480*256, 128*128 and 64*64) were tested, and it was proved that the smaller the tile size is (=more pieces of tiles), the lower PSNR is, as shown in Figure 11; when compressed to 1/10 with the tile size of 3840*2048 (the number of split tiles = 1), PSNR was 34.7dB. Whereas, with tile size of 64*64 (the number of split tiles = 1920), it was 29.8dB, reduced by approx. 5dB.



Fig. 10: QHD Image.

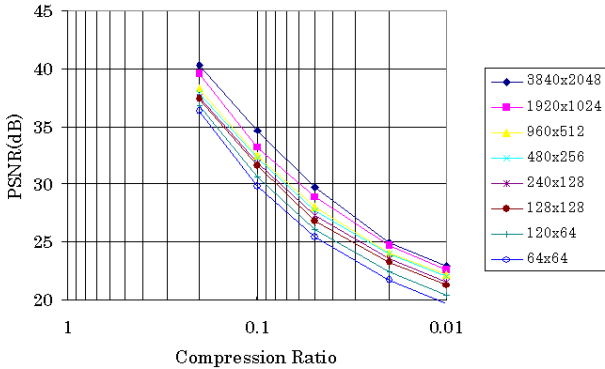


Fig. 11 Compression ratio vs PSNR (DL=5).

Figure 12 shows the relativity between the decomposition level (DL value= the number of wavelet conversion adapted) and PSNR, when the compression ratio is 1/10. This chart indicates that, with smaller tile size and DL value over “2”, PSNR becomes degraded. Therefore, two DL values of “2” and “3” were applied in the evaluation image.

We also generated image data of HDTV and NTSC qualities by processing QHD images through a Lanczos filter.

Setting HD quality images as the standard, we performed an evaluation experiment using the 12 images listed up in Table 5. Those images were randomly presented to subjects, who knew nothing of the

experiment details, including the evaluation of image qualities. An 120 inch front-projection screen (screen gain: 1.2) was used as a display, and subjects were seated or stood 2m to 4m from the screen and answered questions while watching displayed images.

3.5.3 Results

Twenty-three people related to the R&D of display devices participated in the evaluation experiment. The subjects made judgment on the displayed images using the scale of 1 to 5; 1. “very bad”, 2. “poor”, 3. “average”, 4. “very good” and 5. “excellent”. The results (MOS: Mean Opinion Score, S.D.: Standard Deviation) can be seen in Table 6.

The results indicate that, with the compression ratio of 1/10 (image# 3 to 6), even tile-size reduction (= the increase of the number of tiles), which could cause 5dB decrease of PSNR, does not deteriorate the quality of images.

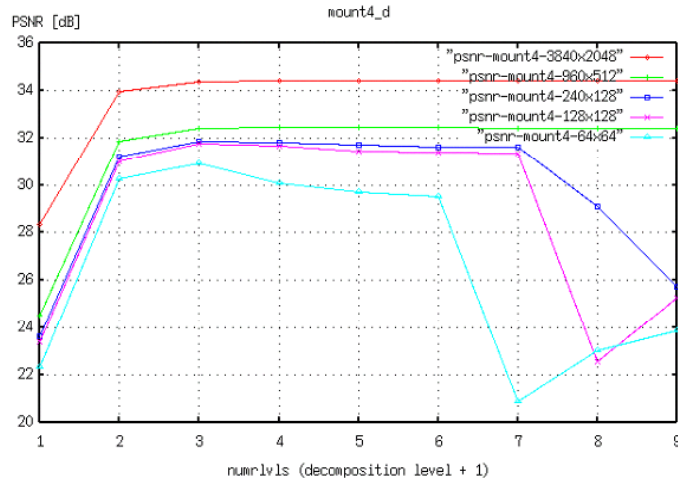


Fig. 12 Decomposition Level vs PSNR.

Table 5: Estimate Images.

No.	Compression Ratio	Tile size	DL
1	Non		
2	1/5	64*64	2
3	1/10	3840*2048	2
4	1/10	480*256	2
5	1/10	128*128	2
6	1/10	64*64	2
7	1/20	3840*2048	2

8	1/20	480*256	2
9	1/20	128*128	2
10	1/20	64*64	2
11	1/20	64*64	3
12	NTSC		

4. Conclusion

This paper proposed WHD/QHD formats as the formula for the next generation HDTV. The concept of the WHD/QHD cameras and projectors we have developed were also mentioned.

Furthermore, we introduced an Internet image delivery model, which is capable to effectively deliver each viewer what he/she really wants to watch. In order to make sure the visual quality of this model, we performed an evaluation experiment on its tiling function and observed no visible deterioration at the compression ratio of 1/10, even when the tile size was reduced down to 64*64.

In future researches, we intend to perform more evaluation experiments on other QHD images to obtain results with objectivity. The delivery model is to be implemented and evaluated before long.

5. Acknowledgement

We, authors, appreciate Dr. Nakagawa of the next generation Internet group and Mr. Arakawa of the image group (both CRL) for the great advice they have occasionally given us. We are also grateful to Dr. Fukuhara of Sony corporation for his precious comments on this project and for the cooperation of Tokyo Electron Ltd. and Victor Company of Japan Ltd. in the system development.

Table 6: Results.

Image No.	MOS	S.D.
1	3.78	0.52
2	3.52	0.73
3	3.35	0.78
4	3.39	0.66
5	3.52	0.65
6	3.48	0.67
7	3.44	0.73
8	3.30	0.77
9	3.13	0.76
10	3.30	0.77

11	2.96	0.93
12	1.22	0.52

6. References

- [1] Tanaka, K., Suzuki, K., Oyama, K., and Arakawa, Y., "WHD (Wide/Double HD) Transmission System", Human Interface Society, June, 2000.
- [2] Tanaka, K., Suzuki, K., Sato, M., and Arakawa, Y., "Super High Definition Images (WHD: Wide/Double HD) Transmission System", ITE Technical Report Vol.24, No.72, PP.37~42, November, 2000.
- [3] Tanaka, K., Suzuki, K., Sato, M., and Arakawa, Y., "Super High Definition Image (WHD: Wide/Double HD) Transmission System", IEICE Trans. D-II, Vol.J84-D-II, No.6, pp.1094-1101, June, 2001.
- [4] T. Komine, F. Kubota, S. Nakagawa and A. Amemiya, "Development of the Japan Gigabit Network, "IEEE Conference 2000 on High Performance Switching and Routing, pp.543-548, June 2000.