

International Telecommunication Union

ITU-R
Radiocommunication Sector of ITU

Report ITU-R BT.2042-5
(11/2011)

**Technologies in the area of extremely
high resolution imagery**

BT Series
Broadcasting service
(television)



International
Telecommunication
Union

Foreword

The role of the Radiocommunication Sector is to ensure the rational, equitable, efficient and economical use of the radio-frequency spectrum by all radiocommunication services, including satellite services, and carry out studies without limit of frequency range on the basis of which Recommendations are adopted.

The regulatory and policy functions of the Radiocommunication Sector are performed by World and Regional Radiocommunication Conferences and Radiocommunication Assemblies supported by Study Groups.

Policy on Intellectual Property Right (IPR)

ITU-R policy on IPR is described in the Common Patent Policy for ITU-T/ITU-R/ISO/IEC referenced in Annex 1 of Resolution ITU-R 1. Forms to be used for the submission of patent statements and licensing declarations by patent holders are available from <http://www.itu.int/ITU-R/go/patents/en> where the Guidelines for Implementation of the Common Patent Policy for ITU-T/ITU-R/ISO/IEC and the ITU-R patent information database can also be found.

Series of ITU-R Reports

(Also available online at <http://www.itu.int/publ/R-REP/en>)

Series	Title
BO	Satellite delivery
BR	Recording for production, archival and play-out; film for television
BS	Broadcasting service (sound)
BT	Broadcasting service (television)
F	Fixed service
M	Mobile, radiodetermination, amateur and related satellite services
P	Radiowave propagation
RA	Radio astronomy
RS	Remote sensing systems
S	Fixed-satellite service
SA	Space applications and meteorology
SF	Frequency sharing and coordination between fixed-satellite and fixed service systems
SM	Spectrum management

Note: This ITU-R Report was approved in English by the Study Group under the procedure detailed in Resolution ITU-R 1.

Electronic Publication
Geneva, 2012

© ITU 2012

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without written permission of ITU.

REPORT ITU-R BT.2042-5

Technologies in the area of extremely high resolution imagery

(Question ITU-R 40-2/6)

(2004-2006-2007-2009-10/2010-10/2011)

TABLE OF CONTENTS

	<i>Page</i>
1 Introduction	3
1.1 EHRI systems under development in Japan	4
1.2 1 920 × 1 080/60 Hz progressive technologies in Japan in the year 2003.....	5
1.2.1 1 920 × 1 080/60P current technology status	5
1.2.2 The technologies and products within the foreseeable range	5
1.2.3 Summary	6
1.2.4 Bibliography.....	6
2 Overview of current EHRI technologies	6
2.1 Still and picture-by-picture image processing (current practice in programme making).....	6
2.2 Computer graphics (CG).....	7
3 Technologies and devices for EHRI realization	7
3.1 Display devices	7
3.1.1 CRT displays.....	9
3.1.2 Projection-type displays	9
3.1.3 Display panels	10
3.1.4 Consumer displays	10
3.1.5 7 680 × 4 320 display	10
3.2 Acquisition technology	11
3.2.1 Electronic picture camera.....	11
3.2.2 Telecine	12
3.2.3 Electronic still camera.....	13
3.2.4 7 680 × 4 320 image sensor	13
3.3 Transmission technology	14

	<i>Page</i>
3.3.1 Optical transmission.....	14
3.3.2 Satellite broadcasting	14
3.3.3 CATV	17
3.4 Storage technology	18
3.4.1 Tape streamers	18
3.4.2 Disks.....	19
3.5 Coding and image processing technology	20
3.5.1 General	20
3.5.2 MPEG-4 studio profile.....	21
3.5.3 H.264/MPEG-4 AVC high profile	23
3.5.4 H.264/MPEG-4 AVC high 4:4:4 intra profile	28
4 Parameters	33

1 Introduction

Throughout this Report a hierarchy of spatial resolutions, which is recommended in Recommendation ITU-R BT.1201 and also given in Table 1, is adopted to classify spatial resolution of pictures in extremely high resolution imagery (EHRI). The limitation of available technologies in this area used to force us to stay mainly in still (non-real-time picture) image applications for higher resolutions. Recently real-time systems for higher resolution systems are reported though those are still in the experimental stage. Basically real-time applications in this area can be defined in terms of frame repetition rates independent of the spatial resolution hierarchy.

The attention of the reader is drawn to Table 21 where some questions have been raised that need further study.

TABLE 1

A hierarchy of spatial resolution in EHRI

	EHRI-0	EHRI-1	EHRI-2	EHRI-3
Spatial resolution (number of samples)	1 920 × 1 080	3 840 × 2 160	5 760 × 3 240	7 680 × 4 320

The hierarchy is based on the well accepted 16:9 picture aspect ratio.

EHRI-1 to 3 are simple integer multiples of EHRI-0 pixel counts, namely 1 920 × 1 080, in horizontal and vertical directions, i.e. the multiplier is the suffix value plus 1.

The EHRI hierarchy in Table 1 is in spatial domain and is independent of the temporal axis. In the real-time case, images are classified by specifying the frame rate in the temporal axis.

1.1 EHRI systems under development in Japan

Recent findings on EHRI technology development have proved that real-time systems in the area of EHRI-1, 2 hierarchy defined in Table 1 are possible. They are still under development and dissemination of devices for EHRI and products to support practical applications is considered to be still several years away. However, an advent of a killer application of EHRI will surely accelerate the development of essential devices and thus system components.

TABLE 2
EHRI hierarchy and major system parameters EHRI systems
under development in Japan (September 2002)

Affiliation	System parameters						Developed hardware
	EHRI hierarchy	Aspect ratio	Horizontal resolution (pixels)	Vertical resolution (pixels)	Frame rate (Hz)	Scanning	
CRL and JVC	EHRI-1	16:9	3 840	2 160	30/60	Progressive	Camera and display
NTT	EHRI-1	16:9	3 840	2 160	24/48/(96)	Progressive	Display
NHK	EHRI-1	16:9	3 840	2 160	60	Progressive	Camera (developed one year before in 2001)
NHK	EHRI-3	16:9	7 680	4 320	60	Progressive	Camera and display

CRL: Communications Research Laboratory

JVC: Victor Company of Japan

NTT: Nippon Telegraph and Telephone Corporation

NHK: Japan Broadcasting Corporation

NOTE 1 – The experimental systems are reported to ITU-R as a contribution in September 2002.

CRL and JVC have jointly developed a camera and display system with 2 000 scanning lines called Quadruple HDTV. The camera system employs three CMOS sensors of $3\,888 \times 2\,192$ pixels and outputs the video signals in four channels of high-definition television (HDTV) signals. The projector employs three LCD panels of $3\,840 \times 2\,048$ pixels. The light output of the projector is 5 200 lm and the contrast ratio is more than 750:1. The resolution of this system corresponds to 2×2 times of $1\,920 \times 1\,080$ pixels.

NTT has also developed a digital cinema system that can store, transmit, and display images of 2 000 scanning lines, with 10-bit each for *R*, *G*, and *B* components. The projector of the system is the same as that of CRL-JVC. Image sources of the system are 35 mm motion films of 24 Hz and the system operates at a frame rate of 24 Hz or 48 Hz. The projector displays the images with a refresh rate of 96 Hz in order to avoid the flicker disturbance. The resolution of this system also corresponds to 2×2 times of $1\,920 \times 1\,080$ pixels.

NHK has developed an EHRI-3 system including a video camera and a projector display succeeding their previous system based on EHRI-1. In order to realize this system, four panels for both CCD and LCD are employed. As the maximum number of panel pixels currently available is $3\,840 \times 2\,048$ for both CCD and LCD, four panels (two panels for greens, one for red and one for

blue) are combined to realize a resolution of $8\text{ k} \times 4\text{ k}$ pixels. The two green panels are arranged by the diagonal-pixel-offset method to achieve the resolution. The resolution of this system corresponds to 4×4 times of $1\,920 \times 1\,080$ pixels.

1.2 1920 × 1080/60 Hz progressive technologies in Japan in the year 2003

1.2.1 1920 × 1080/60P current technology status

– Camera system bases on 2/3 inch CCD technology

A $1920 \times 1080/60\text{P}$ (60 frames/s) camera with three CCD devices for each *RGB* colour, has been developed as an experimental progressive scan HDTV camera in NHK of Japan in 2003. The horizontal and vertical resolutions of this camera are about 1000 TV-lines each, and the vertical MTF (modulation transfer function) response is about 57% on 700 TVL and 30% on 1000 TVL.

– 60P display devices available as products

It had been long believed that it is difficult to realize 1080/60 Hz progressive CRT monitors since the response of horizontal deflection of CRT tube needs certain amount of time to settle itself in a stable condition. A novel technique can overcome this problem without changing the response of the deflection circuit of monitors. With a little bit higher response of the video circuit and the use of higher memory readout speed, the picture part of video signal can be squeezed in time domain and will leave a wider horizontal blanking period in the video signal. With this technique 1920×1080 60 Hz progressive scanning is realized. A professional monitor product is available from one of the broadcast products manufactures in Japan using this scheme. The scanning specification of the CRT monitor covers not only 24P but also 60P.

– 1080/60P interface

To make a 1080/60P system feasible, interface for the system components is considered to be essential. Fortunately there is an SMPTE standard, SMPTE 372M-2002, to use for the links between the equipment. The title of this SMPTE standard is “Dual Link 292M Interface for 1920×1080 Picture Raster”. The SMPTE standard uses two HD-SDI connections to transmit 2.970 Gbit/s data. The specification includes 1920×1080 60P/4:2:2/10-bit interconnections. Here, each link is specified in Recommendation ITU-R BT.1120 and can carry a 10-bit serial data stream defined in Recommendation ITU-R BT.709.

1.2.2 The technologies and products within the foreseeable range

– Projectors available before the end of 2004

The availability of 1080/60P projectors is a product planning issue, and not so much a technological issue. The processing speed is a key technological issue for projecting progressive signals. However, this issue is not difficult, and is rather straightforward. It does not require a novel technique to achieve.

The real issue is to develop a projector that meets the demand and the competitive pricing of the market. One of the broadcast products manufactures in Japan is currently planning to release a full 1920×1080 projector before the end of 2004. This multi-scan projector covers 50P and 60P projection in its specification.

– CCD and CMOS devices for 1080/60P cameras

For acquisition purposes, it is necessary for us to be provided with 60 Hz progressive cameras to have a genuine 60 Hz progressive environment. It is a well-known claim that an optical sensor for the 1080/60P camera will be realized with the refinement of a current CCD device. Around this frame rate the CMOS optical sensor which can provide higher processing speed need not be

required. It is also understood that a camera system with the CMOS device will also be available in parallel with CCD based 60 Hz progressive cameras.

– **Storage devices**

The data rate of $1920 \times 1080/60$ Hz progressive format is two times higher than that of $1920 \times 1080/60$ Hz interlace. In order to record $1920 \times 1080/4:2:2/10$ bit/60 Hz progressive signals on tape it is necessary for a digital VTR to handle approximately 1.24×2 Gbit/s of data for net video only. Compression technology is widely applied to video recording and the picture quality is well accepted. Under the current product line-up of VTRs in several manufactures there are recorders which can record 880 Mbit/s of net video rate. The combination of these technologies makes a recorder for 1080/60 Hz progressive quite feasible. One of the broadcast products manufacturers in Japan has released the specifications of a VTR product which is a portable VTR of the HDCAM series of products. The VTR can record $1920 \times 1080/4:2:2/10$ bit/60 Hz progressive signals with a compression factor of 2.7.

1.2.3 Summary

Japan contributed a progress report to the Radiocommunication Study Group 6 block meetings in the year 2002 on the subject of EHRI. In this Report several EHRI systems are reported to be progressive and have adopted the frame rate of 60 Hz. The systems reported are under experiment but several products which support 60 Hz progressive are already available. As the voices of customers accumulate toward 60 Hz progressive applications, it is a natural tendency that the family of products suitable to those applications should increase. There are clear technology trends to respond to such expectation.

1.2.4 Bibliography

Document 6-9/52, Document 6P/137 – Progress report on extremely high resolution imagery applicable to digital cinema, contribution by Japan.

Contribution Document to AHG on D-cinema in September 2001 – Ultra-high definition video camera, by NHK Science & Technical Research Laboratories.

SMPTE 372M-2002 – Dual Link 292M Interface for 1920×1080 Picture Raster, Society for Motion Picture and Television Engineers.

2 Overview of current EHRI technologies

2.1 Still and picture-by-picture image processing (current practice in programme making)

It is well known that in films of recent release digital film optical effects are often used intensively and the advanced picture processing makes the films very attractive to the majority of audiences. The digital film optical effects, i.e. electronic processing on film, set a new stage for film-making, efficiently replacing the previous film optical processes by the cost-effective and well-established studio post-production techniques. These are compositing with computer-generated graphics, film matting and compositing by blue-screen keyer, retouching of scenes to remove unwanted landscapes and colour and gradation changes for old and decayed films.

There are several such systems available in the market and they are successfully used. The whole system comprises a CCD film scanner, an output film recorder and a signal processing facility based on high-speed workstations. Workstations and relevant software packages are usually used to realize these effects. The equipment can process film quality pictures in the area of EHRI; that is more than 40 times conventional TV signal resolutions.

2.2 Computer graphics (CG)

Various high quality graphic images are generated on computers. The images are generated in non-real-time, and there are no serious problems involved in this technology area. If disk storage capacity to store the images is large enough and a high-speed computer is used, parameters such as spatial resolution, screen aspect ratio, temporal resolution and others, can be set, in principle according to the demands. However, creation of moving images on a real-time basis is difficult to realize with current technology. It depends on the complexity of the image to be produced and the CG technology used. Image generation by a simple CG technology makes some applications, such as virtual reality systems, flight simulators and game machines, possible in real-time.

For current HDTV programme production, approximately 0.25 h is required using an 800 MIPS computer to generate one frame of a human image. If an EHRI-3 level of image is to be produced with the same technology, four hours will be needed to generate a 4×4 times higher resolution image. Availability of huge CPU power in terms of MIPS and an adoption of dedicated graphics engines are always the key for generating high resolution images in CG.

3 Technologies and devices for EHRI realization

3.1 Display devices

The number of HDTV display monitors for high-grade home use in Japan has begun to increase following the successful introduction of digital satellite broadcasting service for HDTV. The price of such monitors is becoming significantly lower compared to the past.

Personal computers are also becoming popular not only in the office but also in each individual home all over the world. The phenomenon has coincided with the wide penetration of the Internet. The GUI for the "Windows" machines requires much higher display capabilities than VGA (640×480), such as XGA (1024×768), and SXGA (1280×1024). Displays for typical workstations hold a resolution of SXGA or UXGA (1600×1200). Toward the year 2005 WUXGA (1920×1200) and QXGA (3200×2400) TFT liquid crystal display (LCD) monitors will be available in the market and will be used in certain applications.

With the advent of multimedia age and especially after the emergence of interactive applications on TV, requirements for a display have changed. Such a display has to have a characteristic of both TV and PC display. Those two are different in the following respects:

- Gamma non-linearity of a cathode ray tube (CRT) display is pre-equalized before broadcasting, while pictures generated by a PC do not have any pre-equalization. Simultaneous display of those two different pictures on the same screen is therefore a compromise.
- Uniformity of picture resolution across the screen is the essential requirement for any PC monitor. On the other hand a TV monitor does not require uniformity but rather requires higher luminosity. Those two characteristics are difficult to maintain on the same display monitors.
- Generally, TV displays have around 10% over-scan. PC displays do not have any. Besides the CRT, there are several other alternative new display technologies available now. Projection-type displays and panel displays have been developed to provide a larger screen size, which is important for sharing common pictures among a large number of audiences. CRT, liquid crystal on silicon (LCOS), and digital micro-mirror device (DMD)

technologies are used for projection-type displays. Plasma display panel (PDP) technology is used widely for flat panel displays. Various sizes of liquid crystal (LC) panel displays are becoming popular. Fifteen inch XGA, 17 inch SXGA, and 24 inch WUXGA LC panel displays are available for computer display. TV display applications of the LC panel are also becoming popular and 28 inch panels are currently available for the applications.

Both CRT and PDP use electro-luminescent effect of G/B/R phosphors. On the other hand, the LC device controls the amount of light which is generated by a light bulb, and the DMD device, by switching mirrors on and off, reflects the light projected on it to a lens block. For colour display, red, blue and green lights are separated from a single light source by “dichroic prism” and are subsequently led to the modulation block of each projector.

3.1.1 CRT displays

For CRT displays with an image size of about 20 inches, a resolution of around 1000 lines is achievable at a shadow mask pitch of about 0.3 mm. In high-level workstations, a pitch of 0.15 mm has already been achieved. The mask pitch depends on many technical factors, like the thickness of the mask and manufacturing conditions. With the present technology level the limit is estimated to be about 0.16 mm in the 40 inch size of CRT. Current spot size of the electron beam is around 1-2 mm. To have higher resolution it is necessary to reduce the size of the spot to around 0.5-1 mm.

It is also necessary to increase the driving speed of CRT deflection circuitry. This is achieved by reducing the width of the deflection yoke wire and by lowering the loss at the core. To reduce deflection errors a digital compensation circuit will be necessary.

Table 3 shows typical high-resolution applications of the CRT display and some of the parameters of available products in the market.

TABLE 3

Some CRT display products available for high-resolution applications

Area of application		Medical	Graphic display	HDTV display monitor
Display	Size	21 inch	32 inch	30 inch
	Aspect ratio		1:1	19:6
Pixel number		2 048 × 2 560	2 048 × 2 048	1 920 × 1 080
Phosphor pitch		None colour continuous (black and white)	0.31 mm	
Contrast ratio		10 bit D/A		
Scanning	Horizontal	186 kHz	126.8 kHz	
	Vertical	72 Hz non-interlace	60 Hz non-interlace	
Physical size	Depth		1 000 mm	
	Weight		100 kg	

3.1.2 Projection-type displays

By way of projecting light there are several technologies and thus product models available on the market. Following are typical examples of various types of the projectors in the area of high resolution applications.

TABLE 4

Some projection-type display products available for high-resolution applications

Model name	MARQUEE 9500LC	JVC DLA-M4000L	Sony VPL-FE100J	Barco ELM R12
Projection type	3 CRT and 3 proj.-lenses	3 D-ILA devices 1 600 W Xenon	3 LC panels 120 W UHP × 4	3 DLP devices 3 000 W Xenon
Light-mod. device	9 inch CRT	0.9 inch (1 397, 760) D-ILA	1.8 inch (1 310 × 720) Poli-silicon device	1.1 inch DMD (1 310 × 720)
Resolution	2 500 × 2 000	1 365 × 1 024	1 280 × 1 024	1 280 × 1 024
Light output (lumen)	1 300	4 000	3 500	12 000
Screen size (inch)		60~600	40~500	~800
Contrast	–	–	–	500:1
Scanning frequency (horizontal/vertical)	15~152 kHz/ 38~180 Hz	15~82 kHz/ 50~78 Hz	15~100 kHz/ 50~120 Hz	
Power consumption	650 W	2 200 W	770 W	3 750 W
Weight (kg)	80	70	34.5	143

3.1.3 Display panels

In the urban area of a heavily populated city, such as Tokyo, a large-size display panel for advertisement or for public announcement is often observed attached to the outside surface of the buildings in a busy square. Although the display is large, it is primarily designed for SDTV or lower quality pictures. The number of light-emitting devices required on the display is the primary limitation.

Plasma display panel technologies have been studied for a long time. After a long survey period, quite recently, 50 inch full colour panels are available and 60 inch panels are announced these days. They have 640-1 024 vertical resolution.

For the LCD of a direct-view type the availability of a larger size liquid crystal panel is a fundamental problem in terms of technology and cost involved. For a high resolution image display, a larger size screen is requested by most of the viewers.

3.1.4 Consumer displays

Technologies that deal with extremely high resolution images are emerging in the field of consumer electronics as well. Several displays that have more pixels than HDTV were exhibited at the 2008 International Consumer Electronics Show (CES), one of the largest tradeshows for consumer technology. Panasonic showcased a 150-inch PDP prototype, which has 4 096 × 2 160 pixels. Sony exhibited an 82-inch LCD prototype, which has 3 840 × 2 160 pixels and operates at a frame rate of 60 Hz with 10-bit precision.

3.1.5 7 680 × 4 320 display

The development of a direct-view type display for the 7 680 × 4 320 system of UHD TV was announced and exhibited in May 2011. It is based on LCD technology. A photograph and major specifications are shown in Fig. 1 and Table 5 respectively.

FIGURE 1

Newly developed display for 7 680 × 4 320 system



TABLE 5

Specifications

Screen size	85 inches (approx. 1.9 × 1.05 m)
Pixel count	7 680 (H) × 4 320 (V)
Frame frequency	60 Hz
Brightness	300 cd/m ²
Gradation	10 bits for each RGB colour

3.2 Acquisition technology

3.2.1 Electronic picture camera

The marginal spatial resolution of a typical lens system is assumed to be about 100 lines/mm. Therefore, the achievable vertical resolution by a 1 inch lens system (CCD scanning area of 14 × 7.8 mm) is $7.8 \times 100 \times 2 = 1\,560$ lines, and it is considered that an optical system that is larger than 1 inch size would be required in a system above a EHRI-1 level (3 840 × 2 160). NHK, the public broadcaster of Japan is studying extremely high resolution camera systems. The objective is to realize a camera system producing in excess of 4 000 scanning lines. Table 7 shows specifications for their current camera system under development.

TABLE 6

Some panel display products available for high-resolution applications

	PDP	PDP
Display size (inch)	42	60
Aspect ratio	16:9	16:9
Pixel pitch (mm)	0.90 × 0.51	0.972 × 0.972
Number of pixels	1 024 × 1 024	1 366 × 768
Quantization levels	256	256
Number of colours (× 10 ⁴)	1 677	1 670
Luminance (cd/m ²)	500	450
Contrast	500:1	More than 500:1
View angle (degrees)	160	Not available
Power (W)	250	Not available
Comments	Available products	Under development

TABLE 7

Intermediate specifications for a future camera system by NHK of Japan

Horizontal pixel (/line)	4 400
Number of vertical lines	2 250
Number of pixels (active)	800 × 10 ⁴
Aspect ratio	16:9
Frames (/s)	60
Scanning system	Progressive
Imaging system	<i>RGB</i> 3CCD
CCD imaging size (mm)	32.2 × 17.2 (2.5 inch equivalent)
Lens system	Fix focal lens (<i>f</i> : 50 mm)

Higher resolution requires a smaller pixel size with the same size of image pick-up device. The low sensitivity which comes from a smaller pixel size is alleviated by enlarging the light-receiving surface, adopting a high sensitivity device, and reducing the device noise level. As for the number and size of pixels, 2 million-pixel (2/3 in optical system) CCDs have been available for HDTV television. The wider surface of the image pick-up devices can cover up to EHRI-1 but some new technologies would be required for further increase of resolution. The reduction of the *S/N* ratio of a camera lowers the compression rate. Thus, lowering the noise level is of prime importance.

3.2.2 Telecine

Three different image pick-up methods are currently used in telecine. These are image pick-up tube camera or area sensor, flying spot scanner, and laser scanner. Most of the problems originating in high resolution imagery with these techniques come in real-time telecine operations. If the systems

are operating in non-real-time, almost all the problems will disappear because scanning operations can be performed more slowly.

3.2.3 Electronic still camera

The image quality of silver salt photography using 35 mm film is almost equivalent to that of the EHRI-1 class. Handling of much higher resolution is possible by enlarging the size of the film used.

A still image CCD of $100 \times 50 \text{ mm}^2$ size with 51 million pixels, which corresponds to higher resolution than EHRI-3, has been realized. It has 10080 horizontal elements and 5040 vertical elements and can function up to 5 frames/s.

In 2001, 3 million pixel electronic still cameras are widely available in consumer electronics shops.

3.2.4 7 680 × 4 320 image sensor

One of the essential technologies to realize UHDTV is an image sensor. Recent improvements in CMOS image sensor technologies have resulted in large-pixel-count and high-speed sensors. A UHDTV image sensor that operates at 120 frames per second has been reported¹. The primary difficulty with such kind of sensors is the short analogue-to-digital conversion time. The sensor realized the conversion by using a two-stage column-parallel cyclic analogue to digital converter. Table 8 lists the design specifications of the sensor. It outputs 7 680 × 4 320 UHDTV images at 120 frames per second with an optical image size of 21.5 × 12 mm (3/2 inch). This shows the feasibility of realizing a UHDTV camera that operates at 120 Hz in the near future.

TABLE 8

Design specifications of 33 M-pixel CMOS image sensor

Process	0.18 μm 1P4M*
Chip size	26.5 (H) × 21.2 (V) mm
Power supplies	1.8 V (Digital), 3.3 V (Analogue)
Number of active pixels	7 680 (H) × 4 320 (V)
Total number of pixels	7 805 (H) × 4 336 (V)
Pixel size	2.8 × 2.8 μm
Frame frequency	120 fps
Optical format	3/2 inch
Shutter	Rolling
Gain	× 1, × 2, × 3.5, × 8
ADC**	Column-parallel two-stage cyclic
ADC resolution	12 bit
Output interface	96 parallel LVDS***
Output data rate	533 Mbps
Package	896 pin BGA****
Power consumption	Approx. 2.5 W

* 1P4M: 1 polycrystalline silicon, 4 metal

** ADC: analogue to digital converter

*** LVDS: low voltage differential signalling

**** BGA: ball grid array

¹ See K. Kitamura et al., "A 33 M pixel, 120 fps CMOS Image Sensor for UDTV Application with Two-stage Column-Parallel Cyclic ADCs," IISW 2011, pp. 343-346, 2011.

3.3 Transmission technology

3.3.1 Optical transmission

In optical transmission using 1.55 μm wavelength a rate of more than 2.5 Gbit/s and a relay distance of more than 100 km per span have been achieved. As the optical transmission system has a very large transmission capacity compared to other transmission schemes, it will form the fundamental transmission infrastructure for digital imaging in future.

Table 9 shows several potentially important fields of concern in developing for optical transmission technology to convey future high bit-rate signals in the EHRI real-time applications. It is obvious that some innovative break-through technologies are needed but dense wavelength division multiplexing (DWDM) technology in optical transmission has already been established. Large capacity optical networks based on DWDM are becoming widely available in various parts of the world.

TABLE 9

Issues on technology development of optical relay transmission

	In the case where 150 Mbit/s is the applied transmission ratio for real-time EHRI-0 and 1⁽¹⁾	In the case where 600 Mbit/s is the applied transmission ratio for real-time EHRI-2 and 3⁽¹⁾
Optical relay transmission technology	Optical transmission technique up to 100 Gbit/s Coherent light wave transmission technology Light modulation technology DWDM (10 waves) Light amplification technology	Optical transmission technique up to Tbit/s bit level Coherent light wave transmission technology Light modulation technology DWDM (100 waves) –

⁽¹⁾ See Table 19 for definitions of the real-time transmission hierarchy.

3.3.2 Satellite broadcasting

WARC-92 relocated the band 21.4-22.0 GHz in Regions 1 and 3 to the broadcasting-satellite service (BSS) to be implemented after 1 April 2007.

As of 1 April 2007 the introduction of HDTV systems in this band is to be regulated in a flexible and equitable manner until such time as a future competent world radiocommunication conference has adopted definitive provisions for this purpose in accordance with Resolution 507 (Rev.WRC-03). WRC-07 also approved Agenda item 1.13 together with Resolution 551 (WRC-07), which:

“considering

h) that *a priori* planning is not necessary and should be avoided as it freezes access according to technological assumptions at the time of planning and then prevents flexible use taking account of real world demand and technical developments;

i) that interim arrangements for the use of the bands are on a first-come-first-served basis;

...

resolves

1 that ITU-R continue technical and regulatory studies on harmonization of spectrum usage, including planning methodologies, coordination procedures or other procedures, and BSS technologies, in preparation for WRC-12, in the 21.4-22 GHz band and the associated feeder-link bands in Regions 1 and 3, taking into account *considering h) and i);*”

3.3.2.1 21 GHz-band indoor transmission experiment

One study was carried out in Japan in May 2007² in the 21 GHz band of an indoor-experiment nature in which 7 680 × 4 320/60P format video and 22.2 multi-channel audio were successfully transmitted. Figure 2 shows the arrangement of the experimental system.

FIGURE 2
Transmission experiment in the 21 GHz-band



Report BT. 2042-01

Major parameters are listed in Table 10. Measured bit error ratio vs. C/N characteristics is shown in Fig. 3. The results showed good video and audio quality.

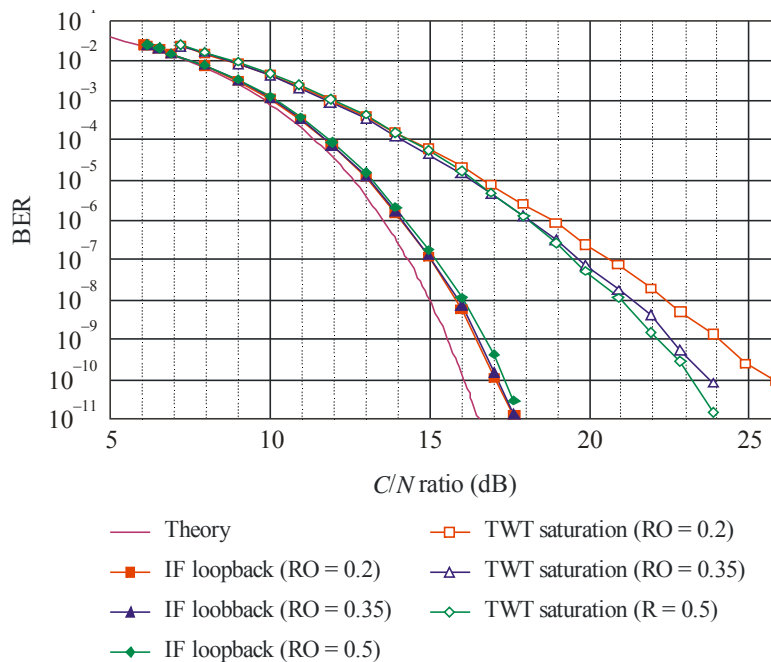
² See SUJIKAL H., SUZUKI Y., TANAKA S. and SHOGEN K. [November, 2007] Super Hi-Vision Transmission Experiment in the 21 GHz band with Prototypes of a Wideband Modulator and a Demodulator. IEICE Tech. Report, SAT2007-54, p. 221-226.

TABLE 10

Parameters of transmission experiment in 21 GHz-band

Source video format	7 680 × 4 320/60/P
Source audio format	22.2 multi-channel
Input signal	MPEG-2 TS at 250 Mbit/s
Modulation	QPSK
Error correction	Reed Solomon
Symbol rate	250 MSymbol/s
Occupied bandwidth	295 MHz
Information rate	500 Mbit/s (250 Mbit/s was used)
Center frequency	21.85 GHz
Transmitting antenna	Horn antenna
Receiving antenna	45 cm ø

FIGURE 3
Bit error rate at the data rate of 500 Mbit/s



Report BT. 2042-02

3.3.2.2 Experiment on advanced satellite broadcasting in the 12 GHz-band

The Association of Radio Industries and Businesses (ARIB) is currently studying a new broadcasting system that can be applied for the services that may start after the end of analog satellite broadcasting in 2011. The study includes transmission coding, video source coding, audio source coding, multiplexing, data broadcasting and the interim report was published in January 2008. It features the following items:

- Data rate is increased to 70 Mbit/s from conventional 52 Mbit/s, while keeping the bandwidth and rain attenuation-tolerance.
- 126 Mbit/s can be achievable using 32-APSK modulation.

- 3 840 × 2 160/60P, 7 680 × 4 320/60P format will be studied in the project.
- H.264 has been adopted as a source coding scheme.

Indoor transmission experiment of 7 680 × 4 320/60P format video and 22.2 multi-channel audio was successfully carried out based on the system (see Fig. 4 and Table 11).

FIGURE 4
Indoor transmission experiment in the 21 GHz-band: Codec (left) and modem and transmission-path simulator (right)



Report BT.2042-03

TABLE 11
Parameters of indoor transmission experiment
in 12 GHz-band

Source video format	7 680 × 4 320/60/P
Source audio format	22.2 multi-channel
Input signal	H.264 TS at 126 Mbit/s
Modulation	32-APSK
Error correction	LDPC, BCH
Symbol rate	32.6 MSymbol/s
Occupied bandwidth	34.5 MHz
Information rate	126 Mbit/s
Center frequency	12.03436 GHz

3.3.3 CATV

Compared with the present analogue transmission over CATV networks, transmission of EHRI signals over CATV will need some of the following new measures:

- use of multiple analogue TV channels;
- realization of a high quality transmission channel;
- much higher speed and broader bands;
- use of digital and optical technology.

Table 12 lists examples of a possible combination of bandwidth and modulation levels for each member of the EHRI transmission hierarchy.

TABLE 12

Bandwidth and modulation levels for EHRI transmission

Real-time EHRI transmission hierarchy⁽¹⁾ (after compression)	Combination between a bandwidth and modulation levels
EHRI-0 (50 Mbit/s)	12 MHz/64-QAM
EHRI-0 and 1 (65-130 Mbit/s)	24-36 MHz/64-QAM 18 MHz/256-QAM
EHRI-2 and 3 (500 Mbit/s)	100 MHz/256-QAM (optical fibre cable required)

⁽¹⁾ See Table 19 for definitions of the real-time transmission hierarchy.

3.4 Storage technology

3.4.1 Tape streamers

The technology trend extrapolated from some of the current tape streamers (8 mm, 1/2 inch) shows that the maximum data storage capacity can be anticipated at around 400 Gbytes and 1 000 Gbytes (see Table 13).

TABLE 13

Maximum data capacity of some tape streamers in the year 2005

Tape streamer	8 mm cassette	1/2 inch cassette
Available data capacity (Gbytes)	400	1 000

Real-time recording of EHRI signals on magnetic tape may not be feasible. Compression is considered to be mandatory to reduce the total amount of data and also the data rate which is otherwise too high to record. Table 14 shows the estimated recording capacity of each data streamer format under consideration.

In Table 14, real-time recording of an EHRI-3 signal clearly indicates that it needs compression whose ratio is higher than 1/30 from the view point of recording capacity. The estimated values are based solely on the total capacity of available media for the streamers. It is also important to consider the data rate for actual recording of the EHRI data streams but this point was left for more detailed discussions.

TABLE 14

Estimated recording capacity of tape streamers by the year 2005

EHRI hierarchy ⁽¹⁾	Real-time EHRI 60 frames/s bit rate (Gbit/s)	Tape streamer Cassette type	Real-time EHRI (h)			Still picture (No. of sheets)
			Compression ratio			Compression ratio
			1/60	1/30	1/4	1/10
EHRI-0 2 million pixels	2.5 4:2:2 10 bit/pixel	8 mm 1/2 inch	21.3	10.7	1.4	7.68×10^5
			53.3	26.8	3.5	1.92×10^6
EHRI-1 8 million pixels	10 4:2:2 10 bit/pixel	8 mm 1/2 inch	5.3	2.67	0.35	1.92×10^5
			13.3	6.68	0.88	4.8×10^5
EHRI-2 19 million pixels	40 4:4:4 12 bit/pixel	8 mm 1/2 inch	1.3	0.67	0.09	4.8×10^4
			3.25	1.68	0.23	1.2×10^5
EHRI-3 33 million pixels	72 4:4:4 12 bit/pixel	8 mm 1/2 inch	0.74	0.37	0.05	2.6×10^4
			1.85	0.93	0.12	6.7×10^4

⁽¹⁾ See Table 19 for definitions of the real-time transmission hierarchy.

3.4.2 Disks

The technology trend extrapolated from the current disk technologies shows that four to nine times increase in recording capacity by the year 2005 can be expected. Table 15 indicates available recording capacity for each size disk currently on the market.

TABLE 15

Recording capacity to be obtained by the year 2005

Storage media	Size (mm)	Current recording capacity (Gbyte)	Future recording capacity (Gbyte)
MD	64	0.14	0.56-1.25
CD-ROM, CD-R	120	0.64	2.56-5.76
DVD-ROM, DVD-R	120	4.7	18.8-42.3

Real-time recording of EHRI signals on disks may not be feasible in terms of recording time and available data rate. Compression is considered to be mandatory to reduce the total amount of data and also the data rate which is otherwise too high to record. Table 16 shows estimated recording capacity for each disk format under consideration.

TABLE 16

Calculated recording capacity of video disks at the year 2005 EHRI hierarchy ⁽¹⁾	Real-time EHRI 60 frames/s bit rate (Gbit/s)	Disk storage media	Real-time EHRI (h)			Still picture (No. of sheets)
			Compression ratio			Compression ratio
			1/60	1/30	1/4	1/10
EHRI-0 2 million pixels	2.5 4:2:2 10 bit/pixel	MD	0.06	0.03	–	2×10^3
		CD	0.3	0.1	0.02	9×10^3
		DVD	2.3	1.1	0.15	8.1×10^4
EHRI-1 8 million pixels	10 4:2:2 10 bit/pixel	MD	0.01	0.01	–	5×10^2
		CD	0.06	0.03	–	2×10^3
		DVD	0.56	0.28	0.04	2×10^4
EHRI-2 18 million pixels	40 4:2:2 12 bit/pixel	MD	–	–	–	1×10^2
		CD	0.02	0.01	–	6×10^2
		DVD	0.14	0.07	–	5.1×10^3
EHRI-3 32 million pixels	72 4:2:2 12 bit/pixel	MD	–	–	–	7×10
		CD	0.01	–	–	3×10^2
		DVD	0.08	0.04	–	2.8×10^3

⁽¹⁾ See Table 19 for destinations of the real-time transmission hierarchy.

Table 16 shows the recording capacity when technology improvement is expected to be nine times the present level. Table 16 clarifies that in motion images, a compression ratio less than 1/30 will make the recording time too short, and 1/60 compression with EHRI-0 will realize a recording time close to that of current analogue LD.

3.5 Coding and image processing technology

3.5.1 General

Ultra-definition television in the real-time EHRI category contains enormous amounts of data. While maintaining high image quality, effective and economical reduction of the bit rate to fit to the available bandwidth of transmission and storage media is quite important.

Table 17 shows the magnitude of compression rate expected at each processing stage in the total bit-rate reduction scheme.

TABLE 17

Picture data compression rate of each element in the total compression scheme

Compression rate in the special frequency domain: discrete cosine transform	5-10
Compression in the temporal domain: motion compensation	2-3
Compression by the statistical characteristics of data: variable length coding	1.3-1.5
Average compression ratio	15-30

3.5.2 MPEG-4 studio profile

MPEG-4 has a wider perspective. It can be applied not only for high compression applications for band-limited transmission based on a new object coding scheme but also for high quality picture compression; i.e. picture compression based on 10/12 bit per pixel coding, 4:4:4 components coding, and higher resolution coding.

Table 18 shows a proposed definition of the levels of the MPEG-4 studio profile.

TABLE 18

Definition of levels of the MPEG-4 studio profiles

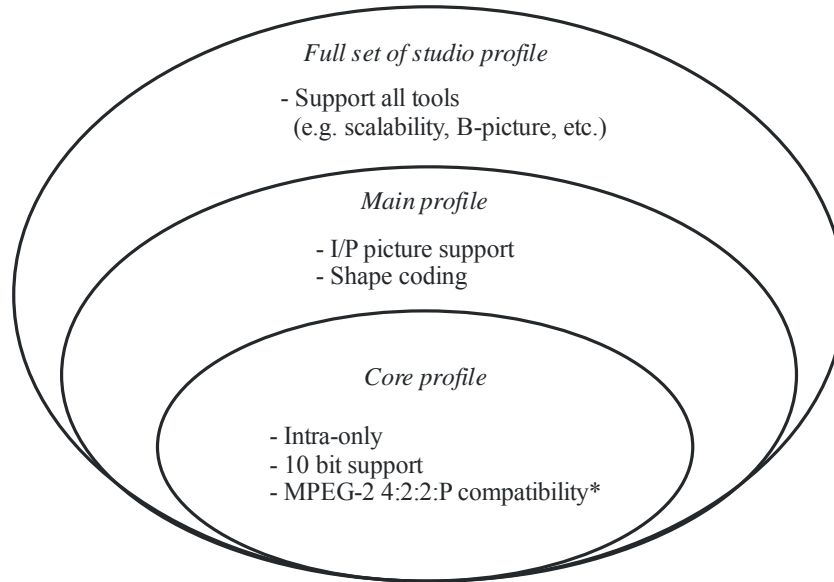
Level		Maximum picture size	Maximum total sample rate ⁽¹⁾	Maximum bit rate	Other aspects
Low (compatible with MPEG-2 high)		H: 1 920 pixels V: 1 088 lines Frame rate: 30 Hz	125 337 600 = $1\,920 \times 1\,088 \times 30 \times 2$	300 Mbit/s	4:2:2 10 bit
Main	422	H: 2 048 pixels V: 2 048 lines Frame rate: 60 Hz	250 675 200 = $1\,920 \times 1\,088 \times 60 \times 2$ > $2\,048 \times 2\,048 \times 30 \times 2$	600 Mbit/s	4:2:2 10 bit
	444		376 012 800 = $1\,920 \times 1\,088 \times 60 \times 3$ > $2\,048 \times 2\,048 \times 30 \times 3$	800 Mbit/s	4:2:2 4:4:4 (YPbPr and RGB) 10 bit
High	422	H: 4 096 pixels V: 4 096 lines Frame rate: 120 Hz	805 306 368 = $4\,096 \times 4\,096 \times 24 \times 2$ > $1\,920 \times 1\,088 \times 120 \times 2$	1.2 Gbit/s	4:2:2 10 bit/12 bit
	444		1 207 959 552 = $4\,096 \times 4\,096 \times 24 \times 3$ > $1\,920 \times 1\,088 \times 120 \times 3$	2.5 Gbit/s	4:2:2 4:4:4 (YPbPr and RGB) 10 bit/12 bit

⁽¹⁾ The performance rating of MPEG-2 decoder is evaluated by the maximum luminance sample rate. In case of MPEG-4, the total sample rate will be a proper measure for chip performance since MPEG-4 studio profile is likely to handle more chrominance samples as 4:4:4.

- *Low level:* This level is basically compatible with the MPEG-2 high level. The difference resides in the 10 bit support in MPEG-4. This level is useful to convert and reuse the assets coded by MPEG-2 [HL@4:2:2P](#) in studio applications.
- *Main level:* This level is intended to cover DTV production and telecine applications. The production system for DTV will require the $1\,920 \times 1\,080 \times 60$ progressive form of programme sources. The telecine machine should support higher resolution picture format such as $2\,048 \times 2\,048 \times 24/25/30$, 10 bit per pixel.
- *High level:* This level is to support super motion systems and high end telecine format. The super motion system will support 120 Hz in future. The high end telecine machine should cover high resolution formats such as $4\,096 \times 4\,096 \times 24$, 10 or 12 bit.

The proposed structure of MPEG-4 studio profiles can be illustrated in Fig. 5.

FIGURE 5
The proposed structure of MPEG-4 studio profiles



* *Note 1* – The compatibility with MPEG-2 4:2:2 profile includes the following two kinds of functionality:

- forward compatibility: MPEG-4 decoder has the MPEG-2 VLD and decoding tools;
- transcoding transparency: minimum quality loss in transcoding process from MPEG-4 to MPEG-2 4:2:2:P. If possible, backward compatibility is preferred.

Report BT. 2042-04

The core profile is a minimum set of the studio profile and includes the simple tools for production requirements. This profile should provide the compatibility with MPEG-2 4:2:2 profile.

Table 19 shows the compression ratio required for transmission for each real-time EHRI image.

Viewers will seldom notice image quality degradation after secondary distribution if the compression ratio is around 15 to 30, as was indicated in Table 17. Additional reduction of the bit rates shall be possible by utilizing human visual sensitivity characteristics or filtering. Therefore, compression up to 1/25 to 1/50 is considered possible to achieve secondary distribution quality. However, as far as the image quality of contribution is concerned, a compression ratio of about 1/6 might be the limit.

In the case of top-level EHRI hierarchy, it is necessary to realize a compression ratio of 300-500 to send the signals through a transmission path. As to this level of compression, some form of technology breakthrough is required. A knowledge-based coding, which is still in the research phase, is one candidate.

TABLE 19

Required compression ratio for transmission

Image hierarchy	MPEG-2 4:2:2 profile	Real-time EHRI-0	Real-time EHRI-1	Real-time EHRI-2	Real-time EHRI-3
Number of effective pixels	720 × 512 (for 525) 608 (for 625)	1 920 × 1 080	3 840 × 2 160	5 760 × 3 240	7 680 × 4 320
Sampling frequency ratio	4:2:2	4:2:2	4:2:2	4:4:4	4:4:4
Gradation (luminance colour difference) (bit)	8	10	10	12	12
Frame rate/s	30	60	60	60	60
Source signal bit rate (Gbit/s)	0.216	2.5	10	40	72
Transmission rate (Mbit/s)	5-50	60-80	100-150	150-600	150-600
Compression ratio	20-40	30-40	70-100	70-270	120-480

3.5.3 H.264/MPEG-4 AVC high profile**3.5.3.1 Level extension conforming to EHRI**

ITU-T H.264/MPEG-4 AVC high profile can achieve better coding performance compared to MPEG-2 and conventional MPEG-4, and has the potential to be utilized in a band-limited transmission environment. Table 20 shows representative levels of the ITU-T H.264/MPEG-4 AVC high profile. Although ITU-T H.264/MPEG-4 AVC high profile is applicable to 4:2:0 8-bit coding, ITU-T H.264/MPEG-4 AVC support other chroma-formats and higher bit-depth by employing other profiles. The ITU-T H.264/MPEG-4 AVC profiles that are relevant to encoding EHRI signals are summarized in Fig. 6.

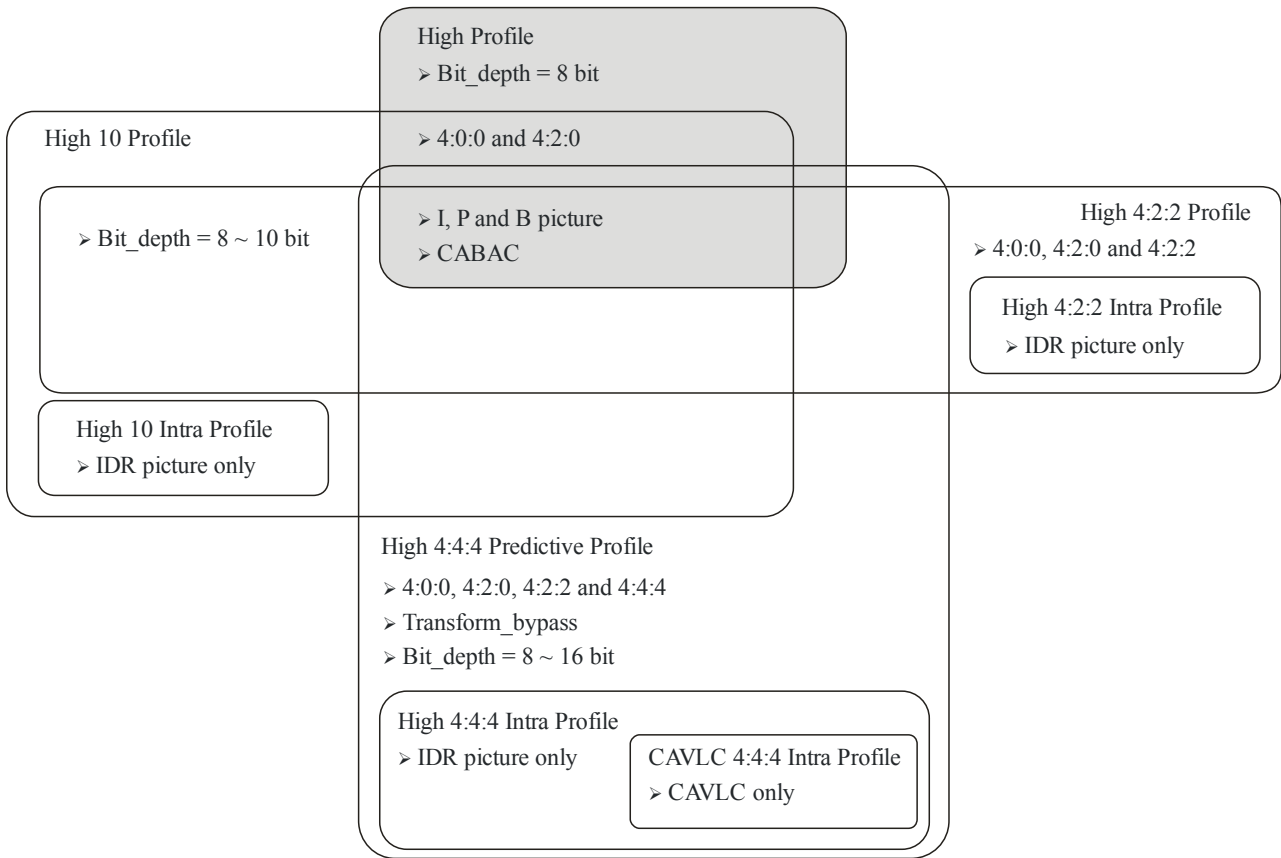
TABLE 20

Representative levels of the ITU-T H.264/MPEG-4 AVC high profile

Level	Max macroblock processing rate MaxMBPS (Macroblocks/s)	Max frame size MaxFS (Macroblocks)	Max video bit rate MaxBR (1 000 bits/s, 1 200 bits/s, cpbBrVclFactor bits/s, or cpbBrNalFactor bits/s)	Max CPB size MaxCPB (1 000 bits, 1 200 bits, cpbBrVclFactor bits, or cpbBrNalFactor bits)
4	245 760	8 192	20 000	25 000
5	589 824	22 080	135 000	135 000
5.1	983 040	36 864	240 000	240 000

FIGURE 6

Structure of profiles relevant to encoding EHRI signals



Report BT.2042-05

In order to evaluate the coding performance of ITU-T H.264/MPEG-4 AVC high profile for EHRI-3 (7 680 × 4 320), a reference encoder was assumed by extending the limitation of level 5.1 to support such large size picture. The extended level limit is defined as shown in Table 21. As for the encoding process, it is assumed that a picture corresponding to a video frame is processed as a single slice, thereby improving the coding performance.

TABLE 21

Extended level limits of ITU-T H.264/MPEG-4 AVC

Max macroblock processing rate MaxMBPS (Macroblocks/s)	Max frame size MaxFS (Macroblocks)	Max video bit rate MaxBR (1 000 bits/s, 1 200 bits/s, cpbBrVclFactor bits/s, or cpbBrNalFactor bits/s)	Max CPB size MaxCPB (1 000 bits, 1 200 bits, cpbBrVclFactor bits, or cpbBrNalFactor bits)
7 776 000	129 600	400 000	400 000

3.5.3.2 Extended coding functions

In order to achieve further coding performance improvement for EHRI-3(7 680 × 4 320) sequences, new coding functions by extending ITU-T H.264/MPEG-4 AVC standard have been studied. In order to evaluate the performance by incorporating promising functions under the framework of ITU-T H.264/MPEG-4 AVC, the following three new functions are additionally incorporated in the encoder. The functions are included in the test model maintained by ITU-T SG 16³, and are selected considering their potential effectiveness for EHRI signals.

1) Extended block size

From the preliminary experiment for 7 680 × 4 320 test sequences, an extended macroblock (MB) size larger than 16 × 16 could achieve significant coding gain under adaptive selection from possible candidates including 16 × 16⁴. In the reference encoder, MB size is adaptively selected from 16 × 16 and 32 × 32, and the MB size decision is conducted picture by picture. Extended DCT size is also employed as a suitable solution for dealing with the residual signal from such large MB size. The available combinations of MB size and DCT block size are summarized in Table 22. In the table, the circled case is the extension to ITU-T H.264/MPEG-4 AVC. Although many approaches have been studied for mechanisms to extend MB size, we focus on a simple mechanism that evaluates the effect while maintaining the block partition style of ITU-T H.264/MPEG-4 AVC.

TABLE 22

Assignment of DCT size under extended MB size scheme

		Luma				Chroma
		Intra			Inter	
		Half	Quarter	Full		
MB size	16x16	8x8	4x4		4x4/8x8	4x4
	32x32	16x16	8x8		8x8/16x16	8x8

2) Adaptive interpolation filter

Applying an in-loop filter to motion compensated prediction reduces prediction error and improves coding efficiency. In ITU-T H.264/MPEG-4 AVC, an interpolation filter based on the Wiener filter algorithm is adopted for fractional-pel motion compensation. The filter coefficients employed in ITU-T H.264/MPEG-4 AVC are constant independent of the image features, and the adaptive interpolation filter (AIF) has been studied. The underlying concept of AIF is to optimize filter coefficients picture by picture⁵.

³ <http://iphome.hhi.de/suehring/tml/>.

⁴ See NAITO S., MATSUMURA A. and KOIKE A. [January, 2006] Efficient coding scheme for super high definition video based on extending ITU-T H.264 high profile. Proceedings of VCIP2006, Vol. 6077.

⁵ See “Enhanced Adaptive Interpolation Filter”, Doc. COM16-C464 (2008), ITU-T.

3) Adaptive loop filter

The in-loop filter for locally decoded image is effective in improving the reconstructed picture quality. As an example employed in ITU-T H.264/MPEG-4 AVC, the de-blocking filter is applied for decoded pixels on the block boundary to suppress the blockiness artifact. In order to decrease the coded noise signal itself, an in-loop filter whose filter coefficients are designed in a similar manner to AIF is promising. From this perspective, the adaptive loop filter (ALF) has been studied. In the typical ALF implementation, in-loop filter coefficients are updated picture by picture⁶.

3.5.3.3 Coding performance evaluation

In order to evaluate the coding performance by the possible extensions addressed in §§ 3.5.3.1 and 3.5.3.2, a subjective evaluation of coded picture quality was conducted. Coding parameters employed in the experiment are shown in Table 23. In advance of the encoding process, $7\ 680 \times 4\ 320/59.94/P$ full resolution 4:4:4 10-bit images were generated from the original test material composed of G1/G2/B/R components with $3\ 840 \times 2\ 160/59.94/P$ resolution, and were then converted into images of $Y/C_B/C_R$ 4:2:0 8-bit for coding. For the conversion process, the luminance and colour-difference matrices compliant with Recommendation ITU-R BT.1361 were applied under the conventional colour gamut system. In the experiment, seven video sequences were tested. A single video sequence is comprised of 480 frames (8 s). Table 24 describes the test sequences and the critical features of each that should be observed when evaluating the coded picture quality.

TABLE 23

Coding parameters for EHRI-3

Input format	$7\ 680 \times 4\ 320/59.94/P$
Video bit-rate	60 Mbit/s, 80 Mbit/s, 100 Mbit/s
Basic coding scheme	ITU-T H.264/MPEG-4 AVC High Profile
Extended functions	Extended block size Adaptive interpolation filter Adaptive loop filter
GOP	IBBPBBP...
IDR picture interval	60 frames
MV search range	$\pm 64 \times \pm 32$, 1/4 pixel precision
Number of reference frame	1
Entropy coding	CABAC
RD optimization	Enabled
De-blocking filter	Enabled

⁶ See “Adaptive Loop Filter for Improving Coding Efficiency”, Doc. COM16-C402 (2008), ITU-T.

TABLE 24

Input video sequences

A	A steam locomotive train passing from right to left at high speed.
B	Underwater shooting of water weeds. Flowers and weeds are swaying in the water. No camera action.
C	The aerial cityscape featured by fine textures. The view is gradually scrolling.
D	A view of a rice field. Many rice ears are trembling slowly. No camera work.
E	Huge and colorful “Nebuta” figures moving in night view.
F	The field is burned off. A big tree appears as the foreground fire gradually disappears. No camera work.
G	The rotating wheel and the body of a steam locomotive train. The camera is gradually zooming out.

Conditions for the subjective evaluation of video quality are shown in Table 25. In the screening of subjects, we rejected subjects who gave higher scores for the coded sequence by 20 points or more compared to the original on at least five occasions.

TABLE 25

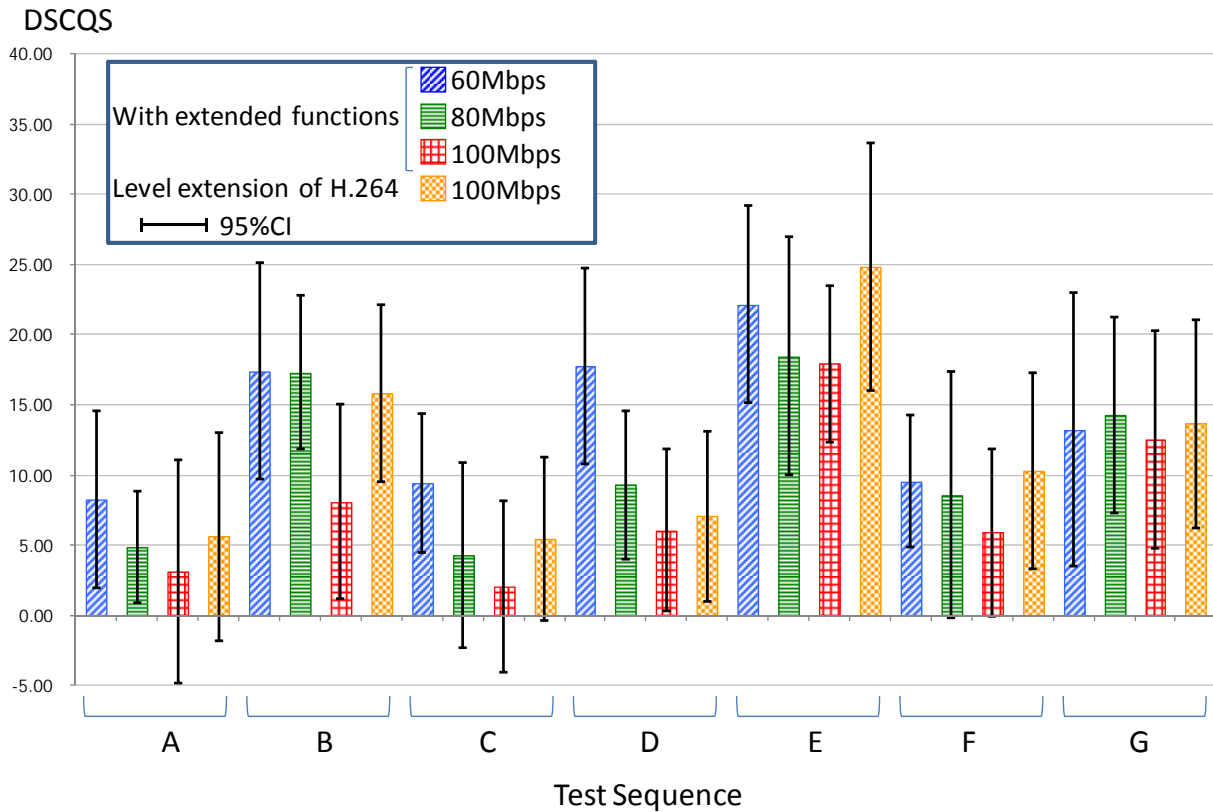
Conditions for subjective evaluation of video quality

Evaluation method	DSCQS
Reference picture	Uncompressed original
Viewing condition	Recommendation ITU-R BT.500-12
Viewing distance	0.75H (H : picture height)
Display	DLA-SH4K projector × 2, 190-inch screen
Subjects	29 non-experts
Date of evaluation	31 August – 1 September, 2010

The results of subjective picture quality evaluation are shown in Fig. 7. DSCQS⁷ results calculated for each of three coding bit-rates are illustrated for each sequence. Figure 7 also shows the 95% confidence interval. The results by the encoder with extended functions confirmed that the subjective score was improved in proportion to the assigned bit-rate. Comparison between results at the bit-rate of 100 Mbit/s revealed that those extended functions contributed to the encoder outperforming the level extension approach.

⁷ DSCQS: Double stimulus continuous quality scale.

FIGURE 7
Results of subjective evaluation of picture quality



3.5.4 H.264/MPEG-4 AVC high 4:4:4 intra profile

3.5.4.1 Level extension conforming to EHRI

ITU-T H.264/MPEG-4 AVC high 4:4:4 intra profile can achieve better coding performance compared to MPEG-2 and conventional MPEG-4, and has the potential to be utilized in a band-limited transmission environment. The level limitation of the ITU-T H.264/MPEG-4 AVC high 4:4:4 intra profile is identical to that of the ITU-T H.264/MPEG-4 AVC high profile as shown in Table 20. In order to evaluate the coding performance of ITU-T H.264/MPEG-4 AVC high 4:4:4 intra profile for EHRI-3 (7 680 × 4 320), a reference encoder was assumed by extending the limitation of level 5.1 to support such large size pictures. The extended level limit is defined as shown in Table 26. As for the encoding process, it is assumed that a picture corresponding to a video frame is processed as a single slice, thereby improving the coding performance.

TABLE 26

Extended level limits of ITU-T H.264/MPEG-4 AVC

Max macroblock processing rate MaxMBPS (Macroblocks/s)	Max frame size MaxFS (Macroblocks)	Max video bit rate MaxBR (1 000 bits/s, 1 200 bits/s, cpbBrVclFactor bits/s, or cpbBrNalFactor bits/s)	Max CPB size MaxCPB (1 000 bits, 1 200 bits, cpbBrVclFactor bits, or cpbBrNalFactor bits)
7 776 000	129 600	4 000 000	4 000 000

3.5.4.2 Extended coding functions

In order to achieve further coding performance improvement for EHRI-3 (7 680 × 4 320) sequences, new coding functions by extending ITU-T H.264/MPEG-4 AVC standard have been studied. In order to evaluate the performance by incorporating promising functions under the framework of ITU-T H.264/MPEG-4 AVC, three new functions, extended block size, adaptive loop filter, and inter channel prediction, were additionally incorporated in the encoder, considering their potential effectiveness for coding 4:4:4 UHD TV signals. The extended block size and the adaptive loop filter are included in the test model maintained by JCT-VC⁸.

1) Extended block size

See §§ 3.5.3.2. As for Table 22, the block size specified for luminance signals is applied to every color component signal in 4:4:4 coding.

2) Adaptive loop filter

See §§ 3.5.3.2.

3) Inter channel prediction for intra coding

Additional coding gain can be achieved by eliminating the correlation between RGB components for the compression coding of 4:4:4 (RGB) video signals⁹. From the result of a preliminary experiment for RGB (4:4:4) 10-bit test materials, as the compression scheme for residual signals from conventional intra prediction, G component is selected as the reference signal, and the B and R components are predicted from the reconstructed signal of G component. Consequently, the prediction process of B and R samples is composed of two steps. At the first step, the B and R samples are predicted by the mode derived from intra prediction of G component. At the second step, residual B and R samples are predicted by reconstructed residual G samples of the same block by the following equation.

$$\text{ResiPred}_{BR}[x, y] = \alpha \cdot \text{ResiRec}_G[x, y] + \beta$$

In the equation, ResiPred_{BR} and ResiRec_G indicate the predicted residual value and the reconstructed residual value, respectively. Parameter α is derived at the encoder and is then quantized, and coded. The quantization step size is defined as 1/8, and the range of parameter α is set as $[-1, +1]$. Parameter β is not coded explicitly because it can be handled as constant value in a transform unit. It means that parameter β is implicitly added to a DC coefficient. As the bit-stream syntax, parameter α is coded as MB side information of B and R component.

3.5.4.3 Coding performance evaluation

In order to evaluate the coding performance by the possible extensions addressed in §§ 3.5.4.1 and 3.5.4.2, a coding experiment was conducted. Coding parameters employed in the experiment are shown in Table 27. In advance of the encoding process, 7 680 × 4 320/59.94/P full resolution 4:4:4 10-bit images were generated from the original test material composed of G1/G2/B/R components with 3 840 × 2 160/59.94/P resolution. In the experiment, seven video sequences as shown in Table 24 were tested. A single video sequence comprises 480 frames (8 s).

⁸ See “Common test conditions and software reference configurations”, Doc. JCTVC-E700 (2010), JCT-VC.

⁹ See “Chroma intra prediction based on residual luma samples”, Doc. JCTVC-F095 (2011), JCT-VC.

TABLE 27

Coding parameters for EHRI-3

Input format	7 680 × 4 320/59.94/P, RGB (4:4:4) 10-bit
Video bit-rate	2 Gbit/s, 4 Gbit/s
Anchor coding scheme	ITU-T H.264/MPEG-4 AVC High 4:4:4 Intra Profile
Extension to the anchor coding scheme	1) Extended block size 2) Adaptive loop filter 3) Inter channel prediction
GOP	All IDR picture
IDR picture interval	1 frame
Entropy coding	CABAC
RD optimization	Enabled
De-blocking filter	Enabled

As the objective picture quality, the PSNR results at the bit-rate of 2 Gbit/s and 4 Gbit/s are shown in Figs 8 and 9, respectively. In the figure, PSNR results are illustrated for every video sequence. PSNR was calculated as the averaged value over all colour components. In order to evaluate the performance for the application of the contribution and the primary distribution, three codecs in tandem connection are assumed. The results by the anchor coding scheme are also shown for comparison. The comparison between the coding schemes revealed that the extended coding functions contributed to outperforming the anchor coding scheme at both the bit-rates. Furthermore, the extended functions suppressed the picture quality degradation against the increase in the number of codecs in tandem connection.

FIGURE 8

PSNR results at the bit-rate of 2 Gbit/s

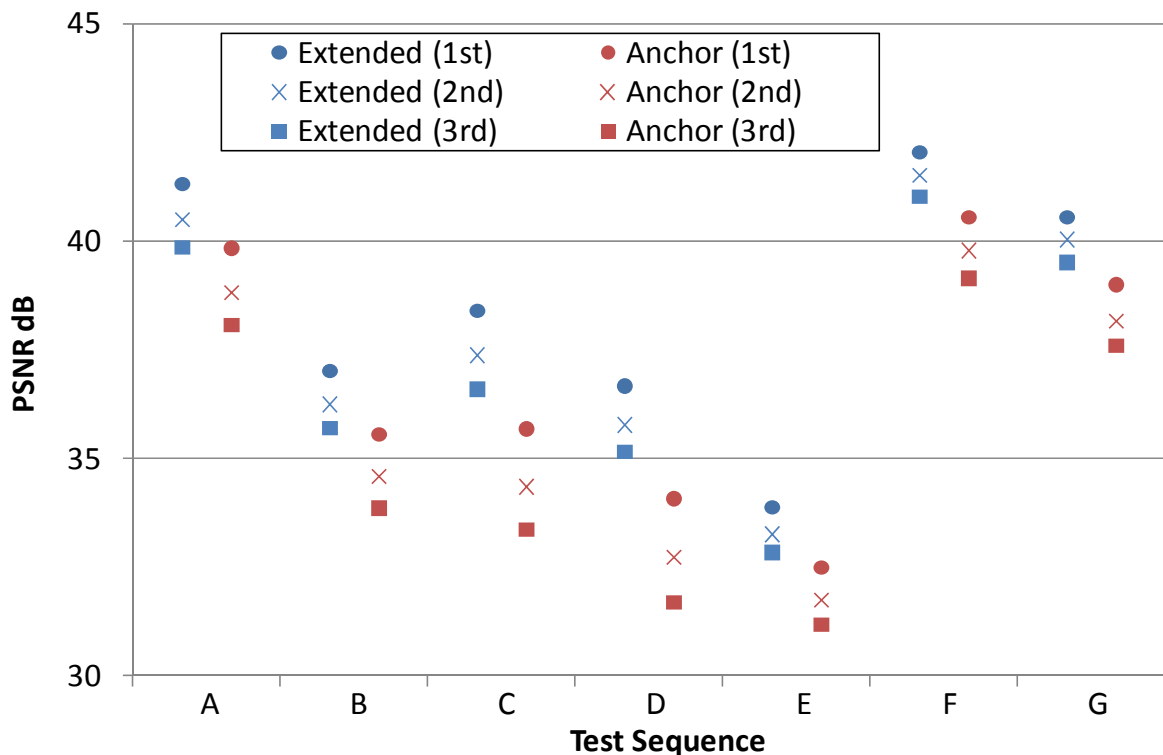
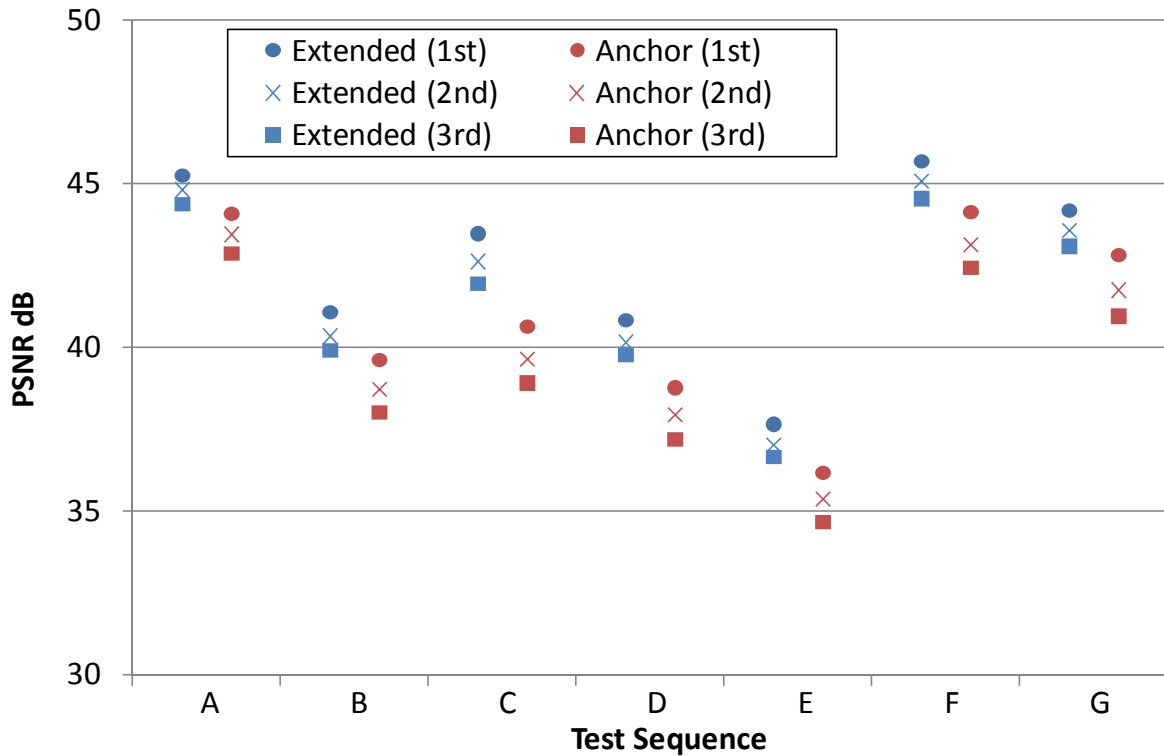


FIGURE 9
PSNR results at the bit-rate of 4 Gbit/s



Conditions for the subjective evaluation of video quality are shown in Table 28. In the screening of subjects, we rejected subjects who gave higher scores for the coded sequence by 20 points or more compared to the original on at least five occasions.

TABLE 28

Conditions for subjective evaluation of video quality

Evaluation method	DSIS (Double Stimulus Impairment Scale)
Reference picture	Uncompressed original
Viewing condition	Rec. ITU-R BT.500-12
Viewing distance	0.75H (H: picture height)
Display	DLA-SH4K Projector x 2, 190-inch screen
Subjects	30 non-experts
Date of evaluation	3-5 August, 2011

The results of subjective picture quality evaluation are shown in Fig. 10. DSIS¹⁰ results calculated for each of three coding bit-rates are illustrated for each sequence. In this experiment, a single transmission link by one codec is assumed. Figure 10 also shows the 95% confidence interval. The results by the anchor coding scheme are also shown for comparison. From the comparison at each bit-rate, the extended encoder outperformed the anchor coding scheme. The significant gain is confirmed especially for the sequence C and E where detailed textures are frequently included.

¹⁰ DSIS: Double stimulus impairment scale.

FIGURE 10
Results of subjective evaluation of picture quality

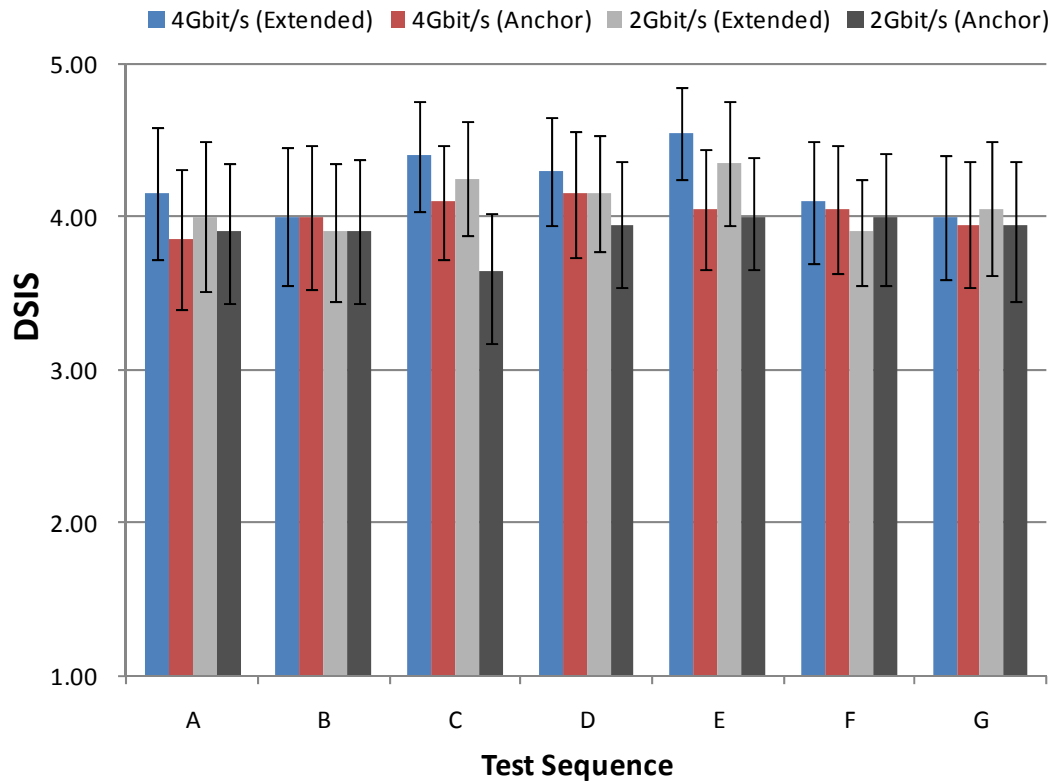


TABLE 29

Correspondences between EHRI and TV and LSDI formats

EHRI hierarchy ⁽¹⁾	Image format	TV Recommendation	LSDI Recommendation
EHRI-0	1 920 × 1 080	ITU-R BT.709	ITU-R BT.1680
EHRI-1	3 840 × 2 160		ITU-R BT.1769
EHRI-2	5 760 × 3 240		
EHRI-3	7 680 × 4 320		ITU-R BT.1769

⁽¹⁾ The classification of EHRI is given in § 1 of this Report.

4 Parameters

TABLE 30
A set of parameters for EHRI

Parameters	Values
Screen aspect ratio	16:9 is the fundamental ratio but other ratio values may also be possible for various applications
Spatial resolution	1 920 × 1 080 and/or its integer multiples are preferable on 16:9 screens. Squareness of the pixel is also important
Temporal resolution	With respect to the scanning system, progressive scanning must be adopted as this system shows characters and figures containing lateral stripes, and it also attains easier image coding or image processing than interlaced scanning systems. It should be noted that higher spatial resolution generally requires higher temporal resolution. A system which adopts approximately 60 frames/s and progressive scanning system is considered to be appropriate
Gradation	8 bits for moving images and 10 bits for still images are essential. It may be necessary to utilize 12-bit gradation to correspond with sophisticated signal manipulations such as image compositions, video editing and secondary uses
Colorimetry	It seems the colorimetry described in Recommendation ITU-R BT.709 is appropriate for a while, but it may be necessary to utilize a new method which can realize a wider range of colour reproduction