SUPER Hi-VISION for the Next Generation Television
Determination of System Parameters

Masayuki Sugawara (member)†, Masaki Emoto (member)†, Kenichiro Masaoka (member)†, Yukihiro Nishida (member)†, Yoshiaki Shishikui (member)†

Abstract SUPER Hi-VISION is being developed to create the ultimate two-dimensional television. A set of system parameters is the basis in the research and development of televisions. The system parameters of SUPER Hi-VISION have been determined after the various investigations from the viewpoint of human factors in particular. It features the increased pixel count to widen the field of view, high frame rate for better motion portrayal, and wide-gamut colorimetry for better color reproduction. These system parameters have been standardized as Recommendation ITU-R BT.2020 in 2012.

Key words: SUPER Hi-VISION, HDTV, system parameters, image format, ITU-R.

1. Introduction

It is not too early to conduct research and development (R&D) on the next generation television that will succeed high-definition television (HDTV) at this moment when HDTV is now entering its practical stage all over the world. In fact, it took nearly 40 years from the initiation of HDTV R&D to the practical stage of HDTV that was reached by satellite broadcasting in 2000 in Japan. From this point of view, the R&D on SUPER Hi-VISION as the next generation television is ongoing. SUPER Hi-VISION features the ultra-high definition image with a pixel count 16 times that of HDTV and three-dimensional sound with 22.2 multi-channel.

A new television system will be embraced by the viewers because it not only is a generation change but it also provides them a completely new audio-visual experience different to that of the conventional systems. The technologies for the system should be able to afford to offer such an experience. Bandwidth is one of the major issues. Also, the receivers and the contents have to be provided to the users at reasonable cost. The R&D on SUPER Hi-VISION has to specify these requirements. In other words, the questions about what the new audio-visual experience the viewers want is and, with what technologies such experience will be realized must be addressed. The audio-visual experience provided through television systems depends on the viewing environment and conditions and the system parameters for video and audio. This paper describes the determination of the values of SUPER Hi-VISION system parameters.

Section 2 overviews the research relating to television system parameters in television history. The use cases of SUPER Hi-VISION are discussed in Section 3. Section 4 describes the determination of the basic parameters, i.e., pixel count, frame rate, and system colorimetry. The standardization of these parameters in ITU-R in described in Section 5. Section 6 is the conclusion.

2. History of R&D on Television System Parameters

Technologies progress in generations, and the duration of a generation for television technology is longer than in many other technologies. Major changes have happened only twice since the advent of B&W SDTV. These were the advent of color TV and the spread of HDTV. Generational changes have taken longer for television technology than other technologies such as IT because of the special characteristics of the television system or service. That is, to reduce the over system cost, a TV broadcast system consists of a few rather expensive transmission parts and many inexpensive reception parts. It also requires strict standards to support exchange of programs over as wide an area as possible.

The extent to which the television system covers the ability of the human visual system (HVS) is an index of progress for a television system; television may be regarded as expanding the HVS spatially and temporally. From this viewpoint, the development of color
television was an effort to give the television one of the important abilities of the HVS, and the development of HDTV was an effort to expand the field of view (FOV) that the television can provide. That is, even though “high-definition” literally means an expansion of the definition of the image, the true aim of the developers of HDTV was to expand the percentage of the viewing field covered by the screen while increasing the pixel count so as not to degrade the picture quality.

The color television was invented a half century ago, and HDTV is becoming commonplace. Three-dimensional television (3DTV) and ultra-high definition television (UHDTV) are the candidates for the next generation of television. 3DTV tries to add depth information to television while UHDTV tries to expand the FOV beyond that of HDTV. Compared with SDTV, HDTV substantially expands the FOV. However, the horizontal FOV of HDTV is still only 30 arc-degrees $^3$. This means that there is quite a bit of a room for improvement so that viewers may receive an enhanced visual experience, which we call the sensation of reality or an immersive feeling. This is the main concept of UHDTV.

The R&D on an image system with a pixel count beyond that of HDTV started in the early 1990’s $^4$. NHK’s R&D on UHDTV started with the goal of developing a system with a 4000-scanning-line (pixels in the vertical direction) image and 22.2-channel three-dimensional sound in the 2000’s $^1$. The system was called “SUPER Hi-VISION”. The determination of base-band format for the video and audio was the important issue because this information would become the basis of the subsequent system development. Determining the image parameters consisted of two parts, selection of parameter values $^5$ and their standardization $^6$.

3. Use-case of SUPER Hi-VISION

Compared with current HDTV, the application of SUPER Hi-VISION should bring considerably better benefits to its viewers.

$^*$ Standard viewing distance is sometimes used as the reference distance at which the field of view of one pixel is one arc-minute. This distance is sometimes understood as the pixel structure becomes discernable at this distance because one arc-minute corresponds to the slit of Landolt ring for 1.0 or 20/20 visual acuity. This distance for each television system is represented as the ratio of the absolute distance and screen height. The standard viewing distance for a 1920 x 1080 system is 3H. This is sometimes called design viewing distance $^7$ or optimal viewing distance $^3$. It corresponds to the angular resolution of 30 cycles per degree.

These benefits may include:
- A stronger sensation of reality or presence,
- Higher transparency to the real world, and
- More information.

It may be presented in:
- Living rooms,
- Personal spaces in mobile and non-mobile environments, and
- Collective viewing locations such as theatres.

Each screen should be individually adjusted to the particular form of use.

Rec. ITU-R BT.1845 provides guidelines on the relationship between screen size and viewing distance given that the standard viewing distance is one at which one pixel corresponds to the visual angle of one arc-minute. The standard viewing distance is, for example, around 100 cm when viewing an SUPER Hi-VISION image on a 100-inch screen. This means that the pixel sampling structure cannot be detected when the screen is viewed under the conditions on the right and below the line in Fig. 1.

A viewer typically watches television while remaining stationary at around two to three meters from the display. Such a viewing style for SUPER Hi-VISION corresponds to a large screen application (Fig. 1) that may include theatrical environments as well as home environments. This should provide a stronger sensation of reality and a stronger feeling of immersion to viewers by offering them a far wider FOV than current systems can offer.

The second category in Fig. 1 is a medium-size screen application with screen sizes from around 50 to 150
inches. The respective standard viewing distances are 50 to 150 cm. The typical viewing distance for television is longer than this distance. The viewing styles for this application may not only include such conventional ones but also a new style in which viewers change their viewing distance in accordance with their preferences or the content creator’s intentions. For example, fine art can be appreciated by standing as close to the screen as the painters who painted the originals. A SUPER Hi-VISION image of fine art can be appreciated in the same way.

The third category is a small-screen application. Similar applications for text data or still images have recently emerged and are known as electronic paper. The standard viewing distance for a 20-inch screen is 20 cm, at which the HVS reaches the limits of accommodation. The size of a 20-inch screen is close to A3, and a 7680 x 4320 pixel screen would have approximately 350 pixels per inch.

4. Basic System Parameters for SUPER Hi-VISION

The primary determination of parameter values depended on what kind of visual experience the system could provide or to what extent the system performance matched the HVS. Rough values were determined through this process. Then exact values were determined from the engineering point of view. This section describes the pixel count, which is regarded a primary parameter for SUPER Hi-VISION, and the frame rate, which has been added to those of HDTV, and the system colorimetry, which has changed from that of HDTV.

4.1 Pixel Count

The primary aim of SUPER Hi-VISION is to expand the FOV. The pixel count is the major consideration when determining the system parameters because it fulfills the primary aim while maintaining the picture quality. Subjective and objective experiments were conducted to determine the required pixel count for SUPER Hi-VISION. They focused on the relationship between the FOV and the sensation of reality (sense of being there) 5(10). The results show that sensation of reality goes up as the FOV increases to around 100 degrees 5(10). Another use-case can be assumed where the increase in pixel count is used to increase not only the FOV but also to increase the angular resolution. For example, the increase in pixel count results in the increase in angular resolution when the screen size and the absolute viewing distance are the same as those for the system with a smaller number of pixels. In conventional television design, the standard viewing distance where the angular resolution is 30 cycles per degree (cpd) gives the condition where the picture quality is rated “five” on a five-grade quality scale 71. The conditions exceeding this were not considered. Two experiments were conducted to investigate the impact of the increase in angular resolution on the perception of the image. One was on the visual acuity when natural images are used 5(10). The other was on the relationship between the angular resolution and visual fidelity (realness) 11. The results show that discrimination ability and increase in realness were observed even when the angular resolution exceeds 30 cpd.

Fig. 2 plots the horizontal FOV and angular resolution that three systems with different pixel counts provide. The optimal viewing distance for the 1920 x 1080 system provides the angular resolution of 30 cpd and FOV of 30 degrees. The 7680 x 4320 system can increase the FOV up to 100 degrees while keeping the angular resolution. This leads to the stronger sensation of reality as mentioned above. When the increase in pixel count is used for increasing the angular resolution, it leads to improved realness. It was concluded that the required pixel count for SUPER Hi-VISION was 8000 horizontally according to the series of experiments. The exact number of 7680 x 4320, which is just four times larger both in horizontal and vertical direction than that of HDTV, was determined by the harmonization with existing systems.
4.2 Frame Frequency

Flicker and motion blur were investigated to determine the frame frequency of SUPER Hi-VISION. These two characteristics for a particular frequency conflict in terms of the duty ratio, i.e., the ratio of the lighting period of the display to the whole frame period. Flicker becomes more visible and motion blur becomes smaller as the duty ratio decreases, while flicker becomes less visible and motion blur becomes larger as the duty ratio increases.

Although many studies on flicker had been reported, we conducted a new experiment in which we used conditions such as FOV, screen luminance, and duty ratios appropriate for the current display technologies and the SUPER Hi-VISION application.

The result shows that the wider FOV angle requires a higher frame rate to suppress flicker perception. The frame frequency of 60 Hz (the frequency of HDTV) almost satisfies the critical fusion frequency (CFF) limit at a FOV of 30 degrees, i.e., at the standard viewing distance of HDTV. However, SUPER Hi-VISION requires a frame frequency higher than 80 Hz to satisfy the CFF limit at a FOV of 100 degrees.

Motion blur is one of the major artifacts caused by temporal sampling and finite temporal aperture. Conventional television systems have been designed on the basis of the acquisition with 100% temporal aperture and impulse-type display. This means that the motion blur is attributed to the acquisition end. However, the technology shift from CRT to non-CRT has brought motion blur to the display end. A subjective test was conducted. Test sequences consisting of still and moving parts were used to evaluate the impairment in the moving part from the still part. The scroll speed varied from 8 to 32 degrees per second. The results of the experiment show that a simulated motion blur of 6~11 pixels/frame is an acceptable limit. This corresponds to 1/320th of a second for an object moving at 32 degrees/second.

This can be realized by reducing the duty ratio or increasing the frame frequency. We decided that the frame frequency should be increased up to the point where no flicker is perceptible at any duty ratio. Accordingly, motion blur can be reduced by shortening the duty ratio at the camera and/or display end. Harmonization with the existing HDTV system is also desirable. Consequently, we decided to use 120 Hz for the frame frequency of SUPER Hi-VISION.

4.3 System Colorimetry

The system colorimetry of conventional television was limited by the type of display device. In short, the chromaticity coordinates of the RGB primaries were determined on the basis of the properties of cathode ray tube (CRT) phosphors. However, CRT technology is not likely to be used for SUPER Hi-VISION displays. Moreover, it is difficult to imagine a particular display technology that will be predominant in the future. Under such circumstances, we decided on three requirements for the system colorimetry of SUPER Hi-VISION.

- Wide-gamut television colorimetry should handle all the colors covered by existing television systems and other related non-broadcasting systems.
- The color coding efficiency in wide-gamut television colorimetry should be comparable to that in the currently used broadcasting systems.
- Every color used in wide-gamut content should be displayable on a reference monitor so that broadcasters can monitor and control the image quality.

Monochromatic colors at wavelengths of 467 nm, 532 nm, and 635 nm were selected for the blue, green, red primaries, respectively.

5. Standardization in ITU-R

5.1 ITU-R Recommendations prior to UHDTV

The International Telecommunication Union Radiocommunication sector (ITU-R) established the first version of the HDTV Recommendation in 1990. It initiated a study on a system beyond HDTV, terming it extremely high resolution imagery (EHRI), three years later. The study led to Recommendation ITU-R BT.1201. Rec. BT.1201 does not specify all the parameters, but recommends using multiple integer values of the HDTV pixel count for EHRI. Another study on image system formats “beyond HD” started in the early 2000’s and was called large screen digital imagery (LSDI). LSDI aimed at an application that would present sports or cultural events in theatrical environments, and it assumed the use of the same system parameters as HDTV. An expanded version of LSDI was studied, and Recommendation BT.1769 was subsequently established in 2006. It recommends two sets of system parameters: 7680 x 4320 and 3840 x 2160 (pixel count). The other parameter values specified by it are almost the same as those of HDTV.
5.2 Standardization of UHDTV

Standardization of SUPER Hi-VISION as a television application started in ITU-R in 2008, after a Japanese proposal was put forward. The study covered not only the pixel count; rather, all parameters were now subject to review. The results of this study were compiled in Recommendation BT.2020 19, which was established in August 2012. The major parameter values specified in it are listed in Table 1.

Regarding the pixel count, first, some members were not fully convinced of the feasibility of the 7680 × 4320 format from the viewpoint of its necessity for home use. However, after the Japanese contribution was submitted, it was finally agreed that there should be two pixel counts. The Japanese contribution explained that an 85-inch LCD with 7680 × 4320 pixels had been developed 20 and that a subjective evaluation using the LCD had shown a significant improvement in user visual experience 21.

Second, regarding the frame frequency, 120 Hz was added in accordance with a Japanese proposal explaining its necessity and technical feasibility.

Third, regarding the system colorimetry, the one mentioned in section 4.3 was proposed by Japan. Another administration proposed first a way to use imaginary colors as primaries, then to use real colors with different wavelengths. It was eventually decided that the red primary would be slightly changed from 635 nm to 630 nm as a compromise between the two proposals.

5.3 Luminance and Color-difference Equations

The discussion on the luminance and color-difference equations took the longest time. It was common sense among television engineers that the luminance equations for SDTV and HDTV did not carry exact luminance information; this is called the “non-constant luminance” issue. Although many engineers thought that constant luminance should be introduced, a study by ITU-R revealed previously unrecognized merits of the non-constant method. As a result, BT.2020 specifies both conventional and constant luminance methods.

6. Conclusion

This paper described the determination of SUPER Hi-VISION system parameters. SUPER Hi-VISION is considered as the next-generation television after HDTV. A new television system is required to provide worthwhile audio-visual experience to the users. The FOV may match the HVS’s characteristics the least among various attributes of the visual experience provided by HDTV. Increasing the FOV will lead to the new visual experience with a much stronger sensation of reality. This is the primary concept of SUPER Hi-VISION. FOV as well as the angular resolution relates to the pixel count. Increase in angular resolution can improve the visual fidelity while the increase in FOV strengthens the sensation of reality. It was also desirable to review other parameters such as frame frequency and colorimetry for better motion portrayal and color.

Table 1 Major system parameter values for UHDTV and HDTV specified in ITU-R Recommendations.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>UHDTV (BT.2020)</th>
<th>HDTV (BT.709 PART2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pixel count</td>
<td>7680 x 4320, 3840 x 2160</td>
<td>1920 x 1080</td>
</tr>
<tr>
<td>Frame frequency [Hz]</td>
<td>120, 60, 60/1.001, 50, 30, 30/1.001, 25, 24, 24/1.001</td>
<td>60, 60/1.001, 50, 30, 30/1.001, 25, 24, 24/1.001</td>
</tr>
<tr>
<td>Scanning</td>
<td>Progressive</td>
<td>Progressive/Interlace</td>
</tr>
<tr>
<td>Colorimetry</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(CIE, 1931)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>0.708</td>
<td>0.640</td>
</tr>
<tr>
<td>G</td>
<td>0.197</td>
<td>0.300</td>
</tr>
<tr>
<td>B</td>
<td>0.131</td>
<td>0.150</td>
</tr>
<tr>
<td>W</td>
<td>0.3127</td>
<td>0.3290</td>
</tr>
<tr>
<td>Non-linear transfer function</td>
<td>$E' = \begin{cases} 4.5E, &amp; 0 \leq E &lt; \beta \ \alpha E^{0.15} - (\alpha - 1), &amp; \beta \leq E \leq 1 \end{cases}$</td>
<td>same as on the left</td>
</tr>
<tr>
<td>Color equations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Y''<em>{B} = \begin{cases} 0.2627R + 0.6780G + 0.0593B \leq 0 \ 6.9702 \leq B' - Y''</em>{C} \leq 0 \end{cases}$</td>
<td>$Y'' = 0.2627R' + 0.6780G' + 0.0593B' \leq 0$</td>
<td></td>
</tr>
<tr>
<td>$C''<em>{R,C} = \begin{cases} B' - Y''</em>{C}, &amp; 0 \leq B' - Y''<em>{C} \leq 0.7908 \ 1.5816 - B' - Y''</em>{C} \leq 0 \end{cases}$</td>
<td>$C'' = \begin{cases} B' - Y', &amp; 0 \leq B' - Y' \leq 0.7908 \ 1.5816 - B' - Y' \leq 0 \end{cases}$</td>
<td></td>
</tr>
<tr>
<td>$C''<em>{G,B} = \begin{cases} R' - Y''</em>{C}, &amp; 0 \leq R' - Y''<em>{C} \leq 0.7908 \ 1.7184 - R' - Y''</em>{C} \leq 0 \end{cases}$</td>
<td>$C''_{R,C} = \begin{cases} R' - Y', &amp; 0 \leq R' - Y' \leq 0.7908 \ 1.7184 - R' - Y' \leq 0 \end{cases}$</td>
<td></td>
</tr>
<tr>
<td>Chroma sub-sample</td>
<td>4:4:4, 4:2:2, 4:2:0</td>
<td>4:2:2</td>
</tr>
<tr>
<td>Bit depth</td>
<td>10, 8</td>
<td>10, 8</td>
</tr>
</tbody>
</table>
reproduction, respectively. The research on the system parameters of SUPER Hi-VISION has been conducted against these backgrounds. Various sorts of system parameters were reviewed. The major results are shown in this paper.

The determination of system parameters values is the first step in R&D of television systems. The determined set of system parameters and its international standardization will accelerate the development of technologies and equipment for SUPER Hi-VISION. Meanwhile, more than a few people have said that they felt a “depth” feeling. This is very interesting and might be a good example for new study items. It is expected that the development of SUPER HI-VISION will stimulate the research field of the psychological effects of high-picture-quality images.

References
3) Recommendation ITU-R BT.1845-1 “Guidelines on metrics to be used when tailoring television” (2010)
7) Recommendation ITU-R BT.1127 “Relative quality requirements of television broadcast systems” (1994)

Masayuki Sugawara received his B.S. and M.S. degrees in electric communication engineering, Ph.D. degree in electronic engineering from Tohoku University, Sendai, Japan. He joined NHK in 1983. Since 1987, he has been researching solid-state image sensor and HDTV camera at NHK STRL. He was an associate professor at the University of Electro-Communications in Tokyo from 2000 until 2004. He has been attending ITU-R SG6 meeting since 2004 and been active in the working parties that deal with programme production. At present, he is an executive research engineer at NHK STRL and is engaged in the research of ultra high definition TV system. He is a member of IEEE, SID, SMPTE, IEICE and Institute of Image Information and Television Engineers of Japan (ITE).

Masaki Emoto received his B.S. and M.S. degrees in electronic engineering and his PhD degree in human and environmental studies from Kyoto University in 1986, 1988, and 2003, respectively. He has been working on future television systems and human science, especially human visual system, at NHK Science and Technical Research Laboratories. His research interests include future television systems incorporating desirable effects such as a heightened sensation of presence, and eliminating undesirable side effects to the viewer such as photosensitive response, visually induced motion sickness, and visual fatigue.

Kenichiro Masaoka received his B.S. in electronics engineering and M.S. in energy engineering from the Tokyo Institute of Technology, Japan. He joined NHK (Japan Broadcasting Corporation) in 1996. He is a Principal Research Engineer for the Advanced Television Systems Research Division, NHK Science and Technology Research Laboratories (STRL), Tokyo. He received his Ph.D. in Engineering from the Tokyo Institute of Technology in 2009. He worked with Professors Mark Fairchild and Roy Berns for a six-month residency as a Visiting Scientist at the Munsell Color Science Laboratory at the Rochester Institute of Technology in 2012. His research interests include color science, human vision, and digital imaging systems. He is a member of IEEE and the Institute of Image Information and Television Engineers of Japan (ITET).

Yukihiro Nishida is Senior Research Engineer, Advanced Television Systems, at NHK Science and Technical Research Laboratories in Tokyo. He has been involved in many areas of broadcasting technologies including HDTV, digital broadcasting, video coding, multiplexing, quality evaluation and UHDTV. He is also active in the standardization of broadcast technologies in ITU-R, ARIB and the EBU. Mr. Nishida is Vice-Chairman of ITU-R Study Group 6 on broadcasting services, and is also Chairman of Working Party 6B covering source coding, interfaces, multimedia and interactivity. He was a visiting research engineer at RAI Research Center in Turin, Italy, in 1992. He graduated from Keio University, Japan, with an M.Sc. in Electrical Engineering. He is a member of ITE of Japan.
Yoshiaki Shishikui received the B.S., M.S. and Ph.D. degrees in electrical engineering from the University of Tokyo, Tokyo, Japan, in 1981, 1983, and 1997, respectively. He joined NHK (Japan Broadcasting Corporation), Tokyo, Japan, in 1983. Since 1986 he has been with NHK Science and Technology Research Laboratories, where he has been engaged in research on digital signal processing, picture coding, HDTV broadcasting systems, IPTV systems and advanced data broadcasting systems. Currently, he is the head of Advanced Television Systems Research Division of NHK STRL and leads UHDTV and 3DTV research activities. From 2001 to 2003, he was with NHK Engineering Services Inc., where he helped to develop video archives and video-on-demand systems. He was a part-time instructor at several universities in Japan. He has been actively involved in standardization activities at SMPTE and ISO-IEC (MPEG).