

Research on Human Factors in Ultra-high-definition Television to Determine its Specifications

Masayuki Sugawara, Kenichiro Masaoka, Masaki Emoto, Yasutaka Matsuo, and Yuji Nojiri

NHK Science and Technical Research Laboratories
1-10-11, Kinuta, Setagaya-ku, Tokyo, Japan 157-8510
e-mail: sugawara.m-fq@nhk.or.jp

Abstract

High-definition Television (HDTV) has now completely entered the practical stage in many countries throughout the world. We can envisage the future of broadcasting by looking back on the long history on HDTV from when it was first proposed to when it was deployed.

Ultra High Definition Television (UHDTV) is now being studied in Japan as the most promising candidate for next-generation television beyond HDTV. UHDTV consists of extremely high-resolution imagery and multi-channel 3D sound to give viewers a stronger sensation of presence. Various aspects should be taken into account when determining UHDTV specifications. Of these, we believe that human factors, such as how we feel when viewing a video, are one of the most important aspects to be considered so that the system achieves the intended psychological effects. Several research projects have been carried out at NHK's laboratory in accordance with this idea.

They include the following items:

- Dependence of sensation of presence on visual angle, both subjective and objective,
- Required angular resolution based on resolution discrimination threshold and visual realism,
- Negative effects of wide screen video, e.g., motion sickness, and
- Dynamic visual acuity when viewing a wide-angle video.

This paper describes the studies conducted at our laboratory on the human factors of UHDTV.

1. Introduction

The development of HDTV was intended to provide a new media format suited to the delivery of richer information and the creation of a richer visual culture.[1] Now that HDTV has become widely diffused throughout the world, this objective is being achieved. In addition, HDTV has become an internationally standard format. As seen with other media, television evolves in stages, towards the ultimate form, on the basis of available technology and the demands of consumers. Currently, the development of HDTV is being driven by digital compression and flat panel display (FPD) technology, on the technical front, and by the desire for higher quality, more compelling audiovisual experiences on the demand side. History shows that one sign of the maturity of a particular media format is the arising of a desire for new media formats. In the 1990s, several R&D initiatives were launched to explore next-generation video systems of higher definition than HDTV.[2,3] These research efforts, which have led to the development of applications such as digital cinema,[4] expanded large screen digital imagery(LSDI),[5] and ultra-high-definition TV (UHDTV), continue to the present.

At NHK, research into new video systems beyond HDTV began in the mid 1990s, with the commencement of work on a UHDTV camera. One envisioned application for this research was the creation of HDTV programs and advanced HDTV services. Later, we clearly set the goal of formulating a new media format to surpass HDTV specifications, and by 2002 we developed our first UHDTV prototype system.[6] This system consisted of a camera, disk recorder, and projection display. Our UHDTV prototype system was exhibited for half a year at the World Expo 2005 in Aichi, Japan,[7] where it was seen by some 1.56 million people. At the same time the system was being adopted at a museum.[8] Since

2006, the prototype has also been demonstrated in the U.S. (NAB 2006 & 2007) and in Europe (IBC 2006). In this time, work has continued to develop various elemental technologies indispensable to a UHDTV television system, such as optical transmission for uncompressed video [9] and MPEG2 compression devices,[10] and we conducted experiments on indoor transmission in the 21 GHz band,[11] and on real-time transmission of live TV programs using the IP transport protocol.[12]

Along with development of hardware-based elemental technologies and program trials, the secret of success in developing a new media format is giving careful consideration to the question of what's needed for a broadcast system capable of delivering a compelling "highly realistic," audiovisual experience.

When the specifications of HDTV were being examined too, the question was similarly studied from the perspective of ideal requirements.[13,14] However, it was inevitable then, as now, that this question had to be examined on the basis of the technology available at the time, or expected to be available in the foreseeable future. The fact that the demands and expectations of people with respect to image quality tends to change after viewing HDTV for the first times also needs to be noted. That is, one viewing experience conditions subsequent viewing experiences. In the initial stages of HDTV development, CCDs and FPDs, which are now widely diffused, and high-efficiency encoding and digital transmission technologies did not exist. These technologies were useful in the ultimate success of HDTV. At the same time, they also provide us with the potential to develop, in the foreseeable future, image capture/reproduction systems of higher performance and quality than HDTV.

In view of all this, NHK SRTL is examining image formats for UHDTV, focusing on a human-scientific perspective. In this report, we relate some of the results of our research work.

It goes without saying that sound is a vital component of television systems. Since UHDTV is a system designed to deliver enhanced "presence," it is particularly desirable that it incorporates an audio format that is consistent with its high definition visual impact. We are currently proposing a 22.2 multi-channel sound system for UHDTV.[15] A key feature of this system is that the speakers are spread vertically at three height levels, to deliver an expressiveness that covers the entire vertical span of large-screen displays. People who heard this audio system at the demonstrations mentioned above were greatly impressed by this expressiveness. For more details on this 22.2 multichannel sound system, you can refer to another relevant report from this conference.[16]

2. Guidelines for examining basic parameters

Ultimately, imaging systems are for viewing by people, so it is only natural that human visual system (HVS) is given due consideration. However, the extent to which HVS is important depends on the stage in the process of the technology's evolution. At the dawn of the television age for example, image quality was far inferior to the normal vision of natural objects, so it was not necessary to make use of HVS to justify improvements to TV specifications. When the number of scanning lines was increased from 40 to 120, for example, the priority was on technical feasibility and economics. Later, when the specifications of 525 (NTSC) and 625 (PAL) scanning lines were standardized, the most important consideration was bandwidth. It was probably only with HDTV that HVS-related considerations became important in the design of basic parameters. We still have many documents that detail the results of our HVS-related testing. Many of these can yield valuable suggestions even today, even if others might have become outdated due to technological and industrial advances.

UHDTV aims at the ultimate in audiovisual experience. To deliver on this promise it will need to have very advanced specifications. Obviously, however, it is also necessary that the new format is accepted in the marketplace, so there must also be a good balance between technical feasibility and the economic viability of applications. In other words, it is necessary to determine the requirements of UHDTV, based on HVS considerations, in a very efficient way. From the point of view of HVS, the specifications of the media format need to take into account content, viewing environment, and human perception characteristics. Of these, content and viewing environment depend heavily on the particular application of the for-

mat. For television the viewing environment will most commonly be the home. UHDTV will serve as an advanced entertainment space and high-quality information terminal in the home. At the same time, as in the case of HDTV—most likely more than HDTV due its superior performance—it is also expected to find widespread application outside the home. As a result, when we examine the human factors relating to UHDTV, it is advisable to consider not only home applications, but also commercial and industrial applications.

Below are itemized summaries of our findings to now on image formats for UHDTV.

3. Number of pixels

The pixel count necessary for the system is calculated by multiplying the assumed display visual angle with the angular resolution (number of pixels per arc-degree of visual angle). A major challenge in the design of many video systems has been how to improve the sensation of presence without degrading image quality. A very effective way to enhance presence is to widen the field of view and increase the display size. On the other hand, image quality is largely determined by the angular resolution. The UHDTV specifications need to satisfy values that are necessary for both of these items.

3.1 Visual angle

HDTV is optimized for a display visual angle of approximately 30 arc-degrees. At this viewing distance one pixel corresponds to one arc-minute of visual angle. One of the results of psycho-physical experiments conducted when the HDTV specifications were developed, in order to determine the relationship between presence and display visual angle, was that presence starts to improve as the visual angle increases from 20 arc-degrees, reaching a plateau at 80 to 120 arc-degrees.[17] As a result of this, it was determined that a visual angle of 30 arc-degrees was necessary to provide the level of presence required for HDTV. However, this result also demonstrates that visual angles higher than 30 arc-degrees generate even stronger sensation of presence. In the field of cinema, the best viewing positions in a theater, known as “prime seats,” correspond to visual angles of 35 to 55 arc-degrees.[18] In addition, research done in recent years has shown that visual angles higher than 30 arc-degrees give rise to high-order psychological effects similar to presence.[19] In view of the above, we performed objective and subjective evaluations of the “presence” of a video system for display visual angles ranging between 30 and 100 arc-degrees, utilizing currently available technology.

3.1.1 Subjective evaluation

A subjective evaluation of the sensation of presence was conducted.[20] It consisted of two experiments in which the visual angle of the displayed images was manipulated as a between- or within- subjects factor. These experiments were intended to investigate a contrast effect, i.e. any bias caused by comparing different visual angles in the within-subjects evaluation.

Two hundred adults (98 female, 102 male) participated in both experiments. The subjects of the between-subjects experiment were divided into five groups for evaluating each visual angle. The UHDTV prototype with a 450-inch screen was used, and the still scenery pictures shown in Figure 1 were used in both experiments. These were photographed with a 4-inch x 5-inch large-format camera and then digitized with a drum scanner. The subjects viewed the evaluation image for 10 seconds, followed by 10-second gray homogeneous image. Then they evaluated the sensation of presence using 10-cm visual analog scales ranging from a low extreme of "I didn't feel a sensation of presence at all" to high of "I felt an extremely strong sensation of presence".

Figure 2 shows the mean scores of the sensation of presence for picture "statue" used in the within-subjects-factor experiment. The results show that the scores of the sensation of presence increase monotonously as the visual angle becomes wider. They show significant differences between visual angles except the case of 87.3 and 100 arc-degrees for the image with a camera picture angle of 60 arc-degrees.

Figure 3 shows the mean scores of the sensation of presence for the picture "statue" used in the between-subjects-factor experiment. The results differ from the ones of the within-subjects-factor experiment as they level off at a visual angle of around 80 arc-degrees. They show significant differences between the visual angle of 33.2 arc-degrees and larger visual angles. There were no significant differences between the visual angles from 61.6 to 100 arc-degrees.









Picture angle [deg]	'bay'	'path'	'warehouse'	'statue'
60				
100				

Figure 1. Test pictures used in the sensation of presence evaluation tests (compressed for the manuscript).

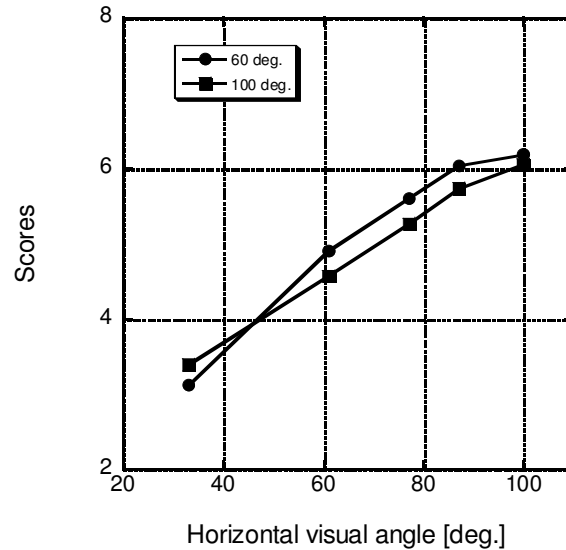


Figure 2. Mean score of sensation of presence for various visual angles they are manipulated with-in-subject factor.

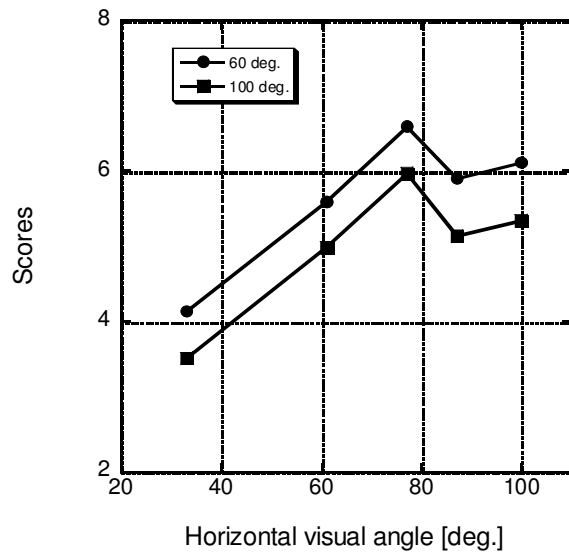


Figure 3. Mean score of sensation of presence for various visual angles they are manipulated between-subject factor.

3.1.2 Physiological index - body sway -

We measured the viewer's body sway while they were viewing still images presented by the UHDTV.[21] It was assumed that the smaller the difference between the real world and the scene presented by video systems becomes, the smaller the difference in the response of human equilibrium would be between viewing the real world and viewing the scene presented by the system.

Twenty adults participated in the experiment. A UHDTV prototype with 320-inch screen displayed three scenery images as the test materials. Viewers stood for 120 seconds on a force platform with the inner sides of their feet in contact, which is called the Romberg foot position. The total body sway distances were calculated by summing the locus of body sway from 30 seconds to 90 seconds during one measurement session.

Figure 4 shows the mean of the differences from the 20 subjects plotted against visual angle. It can be immediately determined from the figure that the wider the visual angle, the shorter the total body sway. A repeated-measurement ANOVA with two variables (picture and visual angle) showed that Mauchly's assumption of sphericity was assumed in the factor 'picture' ($P = 0.708$) but not in the factor 'visual angle' ($P = 0.009$) or their interaction ($P = 0.012$), and the main factor of picture ($P = 0.615$) and their interaction ($P = 0.939$, Greenhouse-Geisser epsilon = 0.484) were not significant, but the main factor of field of view ($P = 0.046$, Greenhouse-Geisser epsilon = 0.595) was significant. This significant decrease in total body sway as a function of visual angle suggests that viewing wide-field images stabilizes human equilibrium. To test whether this stabilizing effect saturates or not with increasing visual angle, Helmert contrasts were performed. The results showed a significant difference between 33.2 arc-degrees and more than 61.6 arc-degrees ($P = 0.014$) and showed a tendency, but not a significant one, between 61.6 arc-degrees and more than 76.9 arc-degrees ($P = 0.071$). There was no difference above 76.9 arc-degrees ($P = 0.476, 0.705, 0.773$). This suggested that the human equilibrium system stabilizes as the still image visual angle increases and their effect on equilibrium tends to saturate at over a visual angle of 76.9 arc-degrees.

From the results above, we concluded that the sensation of presence of two-dimensional large-display images increases as the visual angle increases, reaching saturation at a visual angle of approximately 80 arc-degrees.

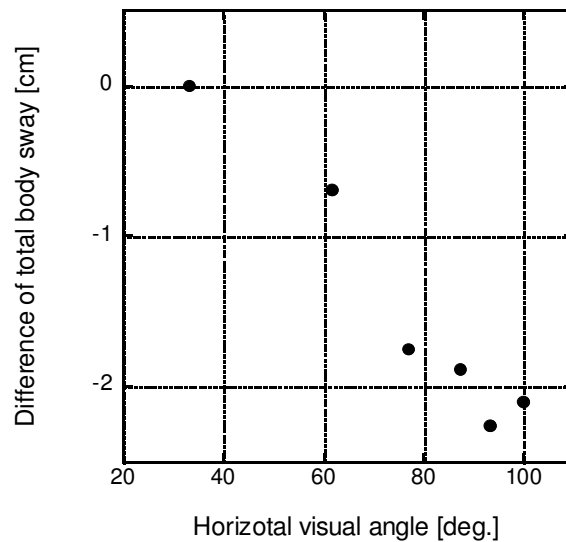


Figure 4. Differential length of total body sway between for a visual angle of 33.2 arc-degrees and the others.

3.1.3 Visually induced motion sickness

Wide-field-of-view video may induce motion sickness in viewers as a result of a strong visual stimulus from a large screen. It is important to take this effect into account when designing such a system.[23]

We evaluated the effects of visual angle on the visually induced motion sickness by using both subjective and physiological indices.[24] The Simulator Sickness Questionnaire (SSQ) and the Single stimulus continuous quality evaluation (SSCQE) were used as subjective indices. SSQ is a method developed by Kennedy to evaluate simulators.[25] It uses 16 symptoms categorized into one or two of three items, Oculomotor, Disorientation, and Nausea. The SSQ scores are obtained by summing these items after weighting. SSQ tests were conducted before and after viewing the test stimuli. SSCQE is a method for subjective evaluation of picture quality that was standardized by International telecommunication union radio sector (ITU-R). It is used to evaluate time-varying picture quality.[26]

We measured heart-rate variability. A change in the balance of sympathetic and parasympathetic nerves is considered to be a sign of motion sickness. It is known that low-frequency component of heart rate variation is affected by both the sympathetic and parasympathetic nerves but high-frequency component is affected by only the parasympathetic nerve. Therefore, LF/HF can be used as an index of the balance state of the autonomic nervous system.

Fifteen adults (14 female, 1 male) participated in the experiments. The UHDTV prototype had a 450-inch screen and it displayed motion images shot from a car running on freeway. Motion vectors of the original images were applied to produce stable and unstable test stimuli from the original one. After 15-minutes of quiet time, the subjects watched 5-minutes of stable images with visual angle of 100 arc-degrees then they watched 10-minutes of unstable images with one of four visual angles.

The results are shown in Figures 5 and 6. The SSQ scores increase as the visual angle becomes wider, and they level off slightly around 83 arc-degrees. SSCQE is worse for the unstable images, and there is no significant difference between visual angles in this case. LF/HF also increases for unstable images, and again, there is no significant difference between visual angles.

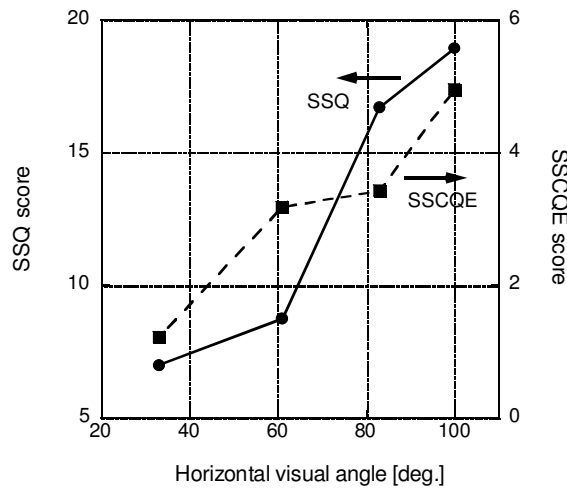


Figure 5. Mean score of SSQ and SSCQE for unstable pictures at various visual angles.

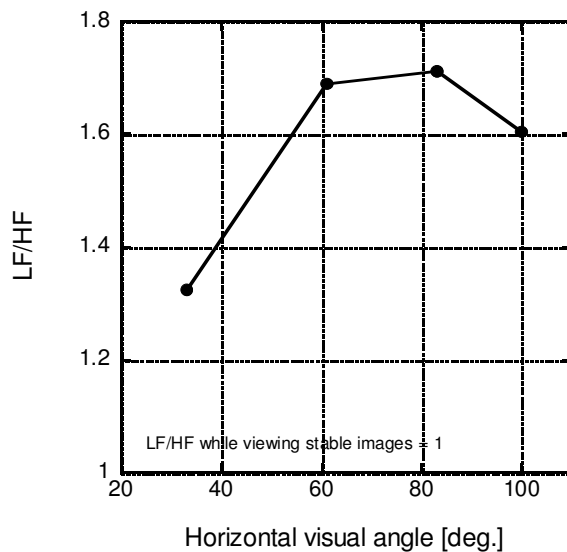


Figure 6. Increase in LF/HF of heart rate variation when viewing unstable pictures at various visual angles.

3.2 Angular resolution

We conducted subjective evaluation experiments to determine the necessary angular resolution, making use of two indicators—the difference in image resolution, and “realness.”

3.2.1 Tests of visual acuity using natural images

Using images of nature as test content, we conducted experiments to determine the discrimination threshold for image resolution, using a FPD. The take of the experiment was to distinguish the still images shown in Figure 7, of varying resolution between 19.5 and 78 cycles per arc-degree (cpd), from an image having a resolution of 156 cpd. The group of test subjects was made up of 11 persons, consisting of both specialists and non-specialists, with an average visual acuity of 2.0. No significant difference was

seen between the results of the specialists and non-specialists. The results are shown in Figure 8. Although the discrimination threshold was fairly low for the image “Brides,” which features few high-frequency components, the limit values for the other two images were 60 to 70 cpd. This corresponds closely to the average visual acuity (minimum separable acuity) of the participants.

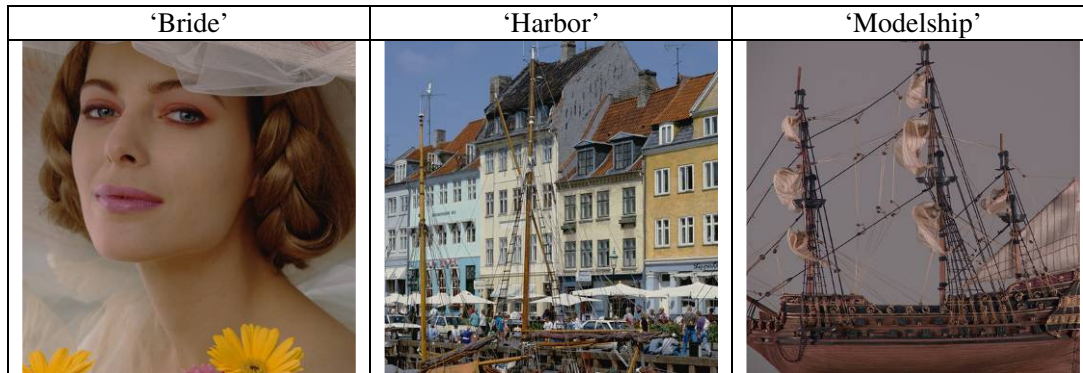


Figure 7. Test pictures used in the resolution discrimination test (compressed for the manuscript). ‘Bride’ and ‘Harbor’ are from Standard High Precision Picture data, the institute of image electronics engineers of Japan.

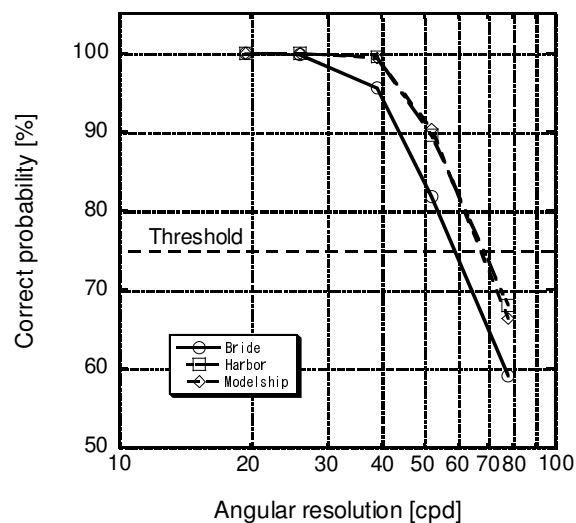


Figure 8. Correct discrimination probability at various angular resolutions.

3.2.2 Comparing realness between real objects and images at various resolutions

We performed subjective assessments to examine the realness of images at six different resolutions, ranging from 19.5 to 156 cpd.[27] A paired-comparison procedure was used to quantify the realness of six images versus each other or versus the real object. Three objects were used (Figure 9). Both real objects and images were viewed through a synopter, which presented the same image to both eyes and removed horizontal disparity. Sixty-five observers were asked to choose the viewed image which was closer to the real object and appeared to be there naturally for each pair of stimuli selected from the group of six images and the real object. It was undisclosed to the observers that real objects were included in the stimuli. The paired comparison data were analyzed using the Bradley-Terry model. The results indicated that realness of an image increased as the image resolution increased up to about 40-50

cpd (Figure 10), which corresponded to the observers' minimum separable acuity, and reached a plateau above this threshold.

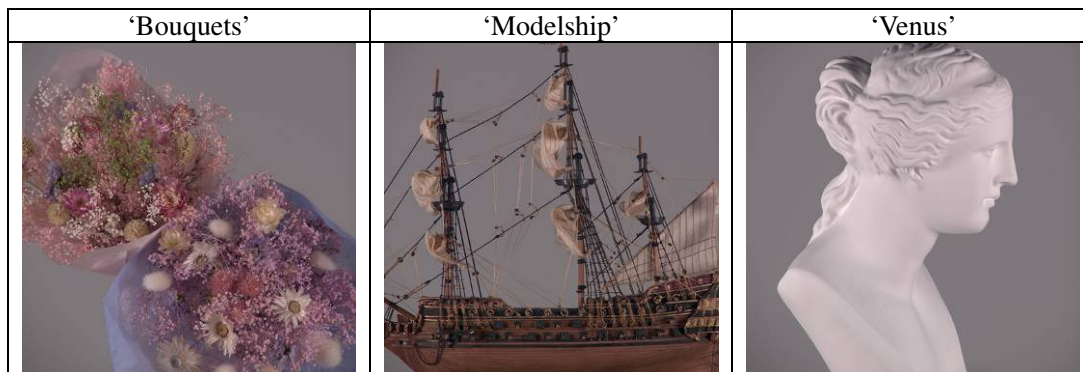


Figure 9. Test objects used in the realness discrimination test.

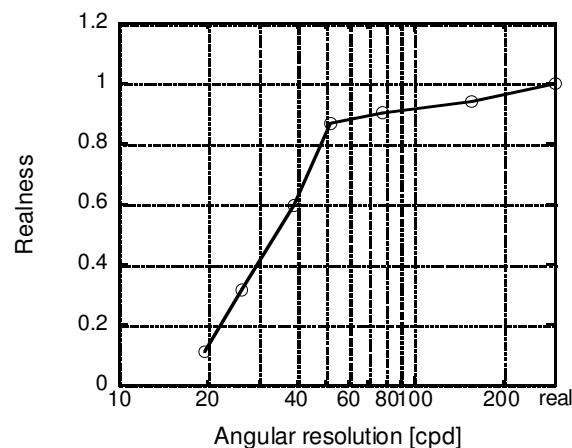


Figure 10. Score of realness for 'Modelship' at various angular resolutions and real object.

4. Parameters other than number of pixels

The primary difference between UHDTV and previous TV formats is the pixel count. However, other parameters are also important. For one, this is because these parameters are influenced by the change in pixel count and/or visual angle. Secondly, even if they are not directly affected by the change in pixel count, these parameters need to be set to produce image quality and performance suitable for the increased pixel count of UHDTV. The frame frequency needs to be reviewed in light of both of these reasons. And while the increase in pixel count does not affect the color space and transfer function directly, it is advisable to give sufficient consideration to the view point of pursuing higher image quality and the fact that display device technology is currently changing rapidly.

Several research efforts have already started at NHK laboratory based on above-mentioned aspects. Of those a recent result of a study on the relationship between dynamic visual acuity and visual angle is described. This investigates how the required frame frequency is affected by widening the visual angle of television.

As displays become wider in terms of visual angle, dynamic visual acuity also may tend to increase. The dynamic visual acuity itself does not change essentially, but rather the region over which the viewed

subjects move on the display expands. This results in a greater chance of capturing the subject, thereby effectively increasing visual acuity. We verified this hypothesis by measuring the dynamic visual acuity for varying display visual angles.[28] Twenty-two adults participated in the experiment. A Landolt ring that corresponds to 0.75cpd was presented on the screen as it moves from left to right with variable speed. The fastest speed at which each observer can read it was measured. The result shows that the threshold speed at which a Landolt ring could be visually perceived increased by a factor of 1.8 as the viewing angle is increased from 10 to 110 arc-degrees. This finding confirms the validity of our hypothesis.

5. Conclusion and Future work

In this report we have described the research efforts at NHK to determine image formats for UHDTV, in particular work done from a human-scientific perspective. Already we have completed our research relating pixel count—the most basic parameter of the new proposed format. In view of our research findings and compatibility with existing formats, we believe that the resolution of UHDTV should be a whole number multiple of 1920 x 1080, up to a maximum resolution of 7680 x 4320 pixels.

Looking ahead, our next major focus of research will be frame frequency. Furthermore, to strive toward a more ideal form of television, further research on color space and transfer function is also vital.

HDTV is becoming more and more widely accepted throughout the world, as the center of a new home entertainment space, and as a home information terminal for our increasingly information-oriented society. It has also undoubtedly played a significant role in the accelerating diffusion of large flat screen displays. UHDTV builds on the foundation of HDTV in pursuit of an even more ideal form for television. Deciding on the image formats is an important and worthwhile undertaking that needs to be implemented at the very initial stages in the process of the developing a large scale audio visual system. At this corresponding stage in the development of HDTV, the technology community played a valuable role. The authors similarly welcome the participation of many other engineers to help us in this process.

6. References

- [1] T. Fujio, "Future broadcasting and high-definition television," NHK Technical Monograph, pp. 5-13, 1982.
- [2] S. Ono, N. Ohta, and T. Aoyama, "All-digital super high definition images," presented at Image Communication, 1992.
- [3] K. Mitani, H. Shimamoto, and Y. Fujita, "2K x 2K Image Acquisition System Using Four CMD Imagers," ITE Journal, vol. 50, pp. 295-301, 1996.
- [4] C. Swartz, "Understanding DIGITAL CINEMA", Focal Press, 2005.
- [5] Report ITU-R BT.2053-1, "Large screen digital imagery", 2006.
- [6] M. Sugawara, et al., "Ultrahigh-Definition Video System with 4000 Scanning Lines", SMPTE Motion Imaging Vol. 112, No. 10 & 11, pp. 339-346, 2003.
- [7] M. Kanazawa, K. Mitani, K. Hamasaki, M. Sugawara, and F. Okano, "Super Hi-Vision Theater at Expo'05," IEICE The First International Workshop on Image Media Quality and its Applications, pp. 45-49, 2005.
- [8] K. Masaoka, M. Kawakita, M. Sugawara, M. Kanazawa, K. Ohzeki, and Y. Nojiri, "Image Quality Management for the Super Hi-Vision System at Kyushu National Museum," IEICE TRANS. FUNDAMENTALS, vol. E89-A, pp. 2938-2944, 2006.
- [9] M. Maeda, Y. Shishikui, F. Suginoshi, Y. Takiguchi, T. Nakatogawa, M. Kanazawa, K. Mitani, K. Hamasaki, M. Iwaki, and Y. Nojiri, "Steps Toward the Practical Use of Super Hi-Vision," NAB2006 BEC Proceedings, pp. 450-455, 2006.
- [10] Y. Nishida, A. Ichigaya, M. Kurozumi, and E. Nakasu, "MPEG-2 Codec for Super Hi-Vision," IEICE General Conference, pp. D-11-26, 2006.
- [11] H. Sujikai, Y. Suzuki, K. Yamagata, S. Tanaka, Y. Nishida, and M. Sugawara, "Indoor Trans-

- mission Experiment of Super Hi-Vision in the 21GHz Band," ITE Technical Report, vol. 30, pp. 13-16, 2006.
- [12] E. Nakasu, Y. Nishida, M. Maeda, M. Kanazawa, S. Yano, M. Sugawara, K. Mitani, K. Hamasaki, and Y. Nojiri, "Technical Development Toward Implementation of Ultra High-Definition TV System," SMPTE Motion Imaging Journal, vol. 116, pp. 279-286, 2007.
 - [13] D. G. Fink, "The Future of High-Definition Television: First Portion of a Report of the SMPTE Study Group on High-Definition Television," SMPTE Journal, vol. 89, pp. 89-94, 153-161, 1980.
 - [14] I. Yuyama and T. Mitsuhashi, "Fundamental requirements for high definition television systems," NHK Technical Monograph, 1982.
 - [15] K. Hamasaki, K. Hiyama, and R. Okumura, "The 22.2 Multichannel Sound System and Its Application," AES 118th Convention, pp. Paper 6406, 2005.
 - [16] K. Hamasaki, et al., "22.2 Multichannel Sound System for Ultra High-definition TV " will be presented at SMPTE Technical Conference & Exhibition, 2007.
 - [17] T. Hatada, H. Sakata, and H. Kusaka, "Psychophysical Analysis of the "Sensation of Reality" Induced by a Visual Wide-Field Display," SMPTE Journal, vol. 89, pp. 560-569, 1980.
 - [18] I. Allen, "Screen Size: The Impact on Picture and Sound," SMPTE Journal, pp. 284-289, 1999.
 - [19] N. Narita, M. Kanazawa, and F. Okano, "Assessment to Determine Format for a UDW Television," SMPTE Motion Imaging, pp. 594-599, 2002.
 - [20] K. Masaoka, et al., "Contrast Effect in Evaluating the Sensation of Presence for Wide Field-of-View", IEICE Technical Report, IE2006-15, pp. 19-24, 2006.
 - [21] M. Emoto, et al., "Viewing angle effects from wide field video projection images on the human equilibrium", Displays Vol. 26, No. 1, pp. 9-14, 2005.
 - [22] M. Emoto, K. Masaoka, T. Shimizu, and M. Sugawara, "Viewing angle dependency of the effects of wide-field images on human equilibrium," ITE Technical Report, vol. 28, pp. 17-20, 2004.
 - [23] ISO IWA 3:2005 "Image safety -- Reducing the incidence of undesirable biomedical effects caused by visual image sequences", 2005.
 - [24] M. Emoto, et al., "Viewing angle effects from wide-field video system on visual-induced motion sickness", ITE Technical Report Vol. 30, No. 22, pp. 25-28 (2006).
 - [25] R. S. Kennedy, et al., "Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness", The International Journal of Aviation Psychology Vol. 3, No. 3, pp. 203-220, 1993.
 - [26] ITU-R Recommendation BT.500-11 "Methodology for the subjective assessment of the quality of television pictures".
 - [27] K. Masaoka, M. Emoto, M. Sugawara, and Y. Nojiri, "Comparing realness between real objects and images at various resolutions," Proc. of SPIE-IS&T Electronic Imaging, vol. 6492, pp. 1-9, 2007.
 - [28] Y. Matsuo, M. Emoto, and M. Sugawara, "Frame Frequency for Wide Field of View Video System - Dependency of Field of View on Dynamic Visual Acuity -", IEICE Forum on Information Technology, 2007.