博 士 学 位 論 文

氏 名（本 籍）	森　千 夏 （埼 玉 県）
学 位 の 種 類	博　士 （情 報 学）
学 位 記 番 号	博 甲 第 149号
学位授与年月日	平 成 30年 3月 31日
学位授与の要件	学 位 規 則 第 4条 第 1項
学 位 論 文 題 目	再構成超解像における高解像度能力の 理論的限界と主観映像評価による検証の研究 -Study of Limitation of Super Resolution Image Reconstruction and Image Quality Assessment of Displays with Super Resolution-

論 文 審 査 委 員	主 査	教 授　合 志 清 一
	副 査	教 授　於 保 英 作
	〃	教 授　馬 場 則 男
	〃	教 授　浜 本 隆 之
	〃	
	〃	
	〃	

工 学 院 大 学 大 学 院

再構成超解像における高解像度能力の理論的限界と主観映像評価による検証の研究

-Study of Limitation of Super Resolution Image Reconstruction and Image Quality Assessment of Displays with Super Resolution-

工学院大学大学院

工学研究科　情報学専攻

森 千夏

2018 年 3 月

Study of Limitation of Super Resolution Image Reconstruction and Image Quality Assessment of Displays with Super Resolution

Chinatsu Mori

Department of Informatics

Graduate School of Engineering

Kogakuin University

March, 2018

Acknowledgments

I would like to express my deepest gratitude to the many people who so generously contributed to the work presented in this dissertation.

I would like to express my special appreciation and thanks to my supervisor, Prof. Seiichi Gohshi, who always gave me suggestions and encouraged me to keep moving forward. His guidance was great help for me to obtain current achievements and finish this dissertation. I had a lot of invaluable experiences during my Ph.D. course, and I am grateful to him not only for his tremendous academic support, but also for giving me so many wonderful opportunities.

Besides my supervisor, I would like to thank the rest of my dissertation committee: Prof. Eisaku Oho, Prof. Norio Baba, and Prof. Takayuki Hamamoto, for their insightful comments and encouragement.Thanks for their help and patience during their extremely busy schedule.

My sincere thanks also go to Mr. Yuji Watanabe, Mr. Hiroyuki Tabata, Mr. Keisuke Ohashi, Mr. Shinichiro Nakamura, from Keisoku Giken Co., Ltd., and Mr. Yoshikatsu Tatsumi, from Fujitsu Connected Technologies Limited, for providing experimental equipment and for technical support. Also I would like to thank to Prof. Kenkichi Tanioka from Tokyo Denki University, for huge contribution of his works to this dissertation.

I would like to thank all the members of Gohshi lab, for giving me support and assistance to my research and daily life, and for all the fun we have had in the last six years. In particular, I am grateful to Mr. Masaki Sugie, and Mr. Hirohisa Takeshita, for enormous contribution with this dissertation.

Last but not the least, I would like to thank my beloved family for their constant support. They gave me the courage to challenge and finish the Ph.D. course and I dedicate this dissertation to them.

Contents

1 Introduction **5**

1.1 Introduction 5

1.2 Super Resolution Image Reconstruction 6

1.2.1 Reconstruction Process for Super Resolution Image Reconstruction . . 6

1.2.2 Essential Problems of Super Resolution Image Reconstruction 9

1.3 Assessment of Super Resolution 10

1.4 Super Resolution Image Reconstruction and Super Resolution with Nonlinear Signal Processing 15

1.5 Super Resolution with Nonlinear Signal Processing for Resolution Enhancement of 4K Videos 16

1.6 Super Resolution for Smartphones 16

2 Super Resolution Image Reconstruction and Imaging Device **17**

2.1 Introduction 17

2.2 SRR and Aliasing 19

2.2.1 Nyquist Sampling Theorem and Subsample 19

2.2.2 Still Image and Video 23

2.3 Essential Issue of SRR 24

2.4 Imaging Device 26

2.5 Block Shaped Aliasing 29

2.6 Conclusions 30

3 Subjective Assessment Method for Multiple Displays with Super Resolution **31**

3.1 Introduction 31

3.2 Subjective Assessment 32

3.3 Proposed Method 33
3.3.1 Scheffe's Paired Comparison 33
3.3.2 Observers and Test Sequences 33
3.3.3 Experimental Environments 34

4 Subjective Assessment of 4K TV Set with Super Resolution Image Reconstruction 35
4.1 Introduction 35
4.2 Subjective Assessment Experiments 36
4.2.1 Experiments 36
4.2.2 Assessment Method 37
4.2.3 Experimental Set Up and Environments 37
4.3 Results and Discussions 37
4.3.1 Experiment 1 38
4.3.2 Experiment 2 43
4.3.3 Discussions 45
4.4 Conclusions 45

5 Subjective Assessment of 4K videos with super resolution with nonlinear signal processing 46
5.1 Introduction 46
5.2 Experiment 47
5.2.1 Subjective Assessment Method 47
5.2.2 Apparatus 47
5.2.3 Test Sequences 48
5.2.4 Observers 48
5.2.5 Experimental Conditions 48
5.3 Results and Discussion 49
5.4 Conclusion 52

6 Subjective Assessment of Smartphone Displays with Super Resolution using Best-worst Method 53
6.1 Introduction 53
6.2 Super Resolution with Nonlinear Signal Processing 55

6.3 Subjective Assessment Method 56
6.3.1 Best-worst Method 57
6.4 Experiment 57
6.4.1 Apparatus 57
6.4.2 Test Sequences 57
6.4.3 Observers 59
6.4.4 Experimental Conditions 59
6.5 Results 59
6.6 Discussions 63
6.7 Conclusion 63

7 Conclusions **64**

Chapter 1

Introduction

1.1 Introduction

The resolution improvement technologies for images/videos are desired for many imaging and viewing applications. Quality of a video is important when we enjoy watching various content. Resolution has been frequently represented as quantitative quality. Resolution of TV displays has been improved from analog (640 × 480) to high definition (1,920 × 1,080). In 2011, 4K (3,840 × 2,160) TV sets with the four times resolution of HDTV have been released. Recently, 8K (7,680 × 4,320) TVs and broadcasting systems for practical use are under development for the Tokyo Olympic Games in 2020.

A TV set is a conventional video device for watching video content. Recently, new devices for watching video, such as smartphones have been introduced. Resolution of smartphone displays has been increasing similarly to TV sets; the resolution has been improved from half-size video graphics (HVGA) (480 × 340) to high-definition (HD) (720 × 1,280), and to wide quad HD (WQHD (2,560 × 1,440). Recently, a smartphone with a 4K display is also introduced.

In order to improve resolution, all of the systems need to be improved as well. Figure 1.1 shows imaging and viewing system for image/video. The system shown in Figure 1.1 progresses from left to right. The scene with the object on the left is captured with an imaging device within still/ video camera. The captured data is encoded with any format to reduce the cost of storage or transmission. The coded data is then transmitted through broadcasting or network, and received by the display devices. The received data is decoded and they are processed with size conversion or quality correction. After these processes, the data is finally displayed on the monitor.

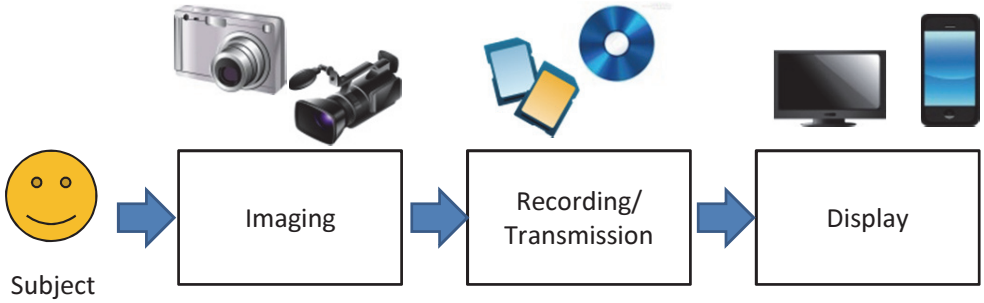


Figure 1.1: Progress of imaging and viewing system

Although 4K displays are currently available, we cannot enjoy watching videos in 4K resolution. 4K video content is insufficient and the provided content is almost entirely recorded with HDTV. Playing HDTV content on a 4K display does not improve resolution theoretically, as the content is merely enlarged to fit to the display, resulting in an interpolated image. Therefore, alternative ways of improving content resolution are required.

1.2 Super Resolution Image Reconstruction

Super resolution (SR) is one of the ways of improving the resolution of recorded images and videos. SR is different from the previous image enhancement technology called enhancer (or unsharp mask) [1]. SR can create high-frequency elements an enhancer cannot create. There are numerous SR methods previously reported and some different types of SR method are practically used; super resolution image reconstruction (SRR) [2][3][4][5][6], super resolution with nonlinear signal processing (NLSP) [7][8][9], and learning-based super resolution[10][11][12] [13][14]. SRR is the most popular SR method with the most extensive body o research. The following is a description of process which may potentially work with SRR.

1.2.1 Reconstruction Process for Super Resolution Image Reconstruction

SRR creates a high-resolution image (HRI) from multiple low-resolution images (LRIs). Figure 1.2 shows the process of SRR. In Figure 1.2, the signal is processed from the top to the bottom. The HRI at the top of the figure is prepared for SRR processing. The low-pass-filter (LPF) limits the bandwidth of the HRI. The HRI processed with LPF is subsampled, and multiple LRIs are created. LRIs are generated by subsampling pixels at different phases. The example of Figure 1.2 is the 2:1 subsample. The pixels of four LRIs can compose all the pixels of the HRI,

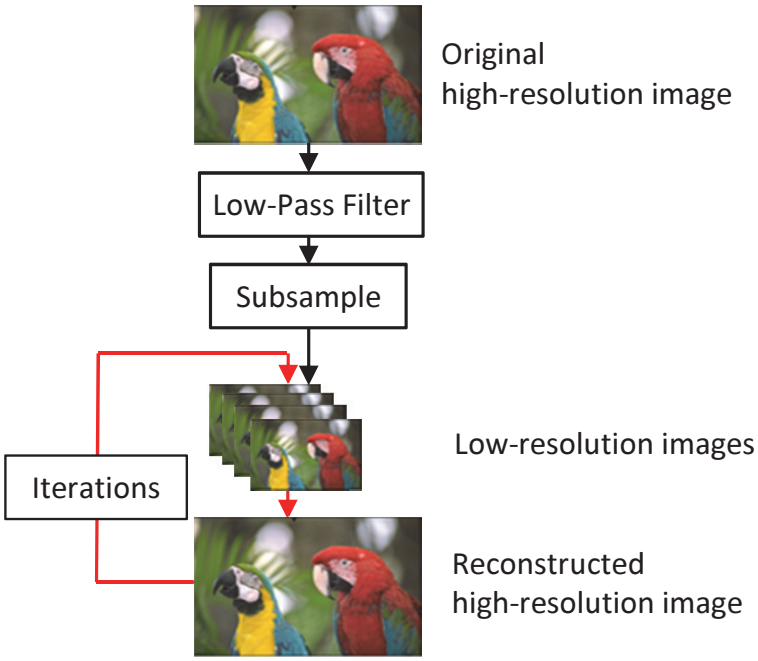


Figure 1.2: Super resolution image reconstruction

but SRR requires many LRIs with duplicated pixels. The reconstruction process is solving an inverse problem and this process is iterated until the value of the cost function is minimized.

The process for generating LRIs with SRR is shown in Figure 1.3 [2]. The top of Figure 1.3 is a real world scene as a continuous HRI. The HRI is warped because of the motion. The HRI is also blurred by the camera lens with continuous point spread function (PSF), and then discretized by the imaging device such as CCD. The additive Gaussian noise is added with the digitalized image, and the blurred and noisy LRIs are generated. The process in Figure 1.3 can be expressed by a product matrix as shown in Equation 1.1.

$$Y_k = D_k H_k F_k X + V_k \quad (k = 1, 2, ..., N) \tag{1.1}$$

The matrix Y_k represents the k^{th} LRI, and matrix X represents HRI. The matrix F represents the motion effect operator of the warp, the matrix H_k represents the PSF of the camera lenses, and the matrix D represents the down sampling operator for the discretization. The matrix V is the additive Gaussian noise.The HRI $\underline{\hat{X}}$ is reconstructed by the error minimization between available LRIs (Y_k) and the LRIs generated from estimated HRI $\underline{X}$:

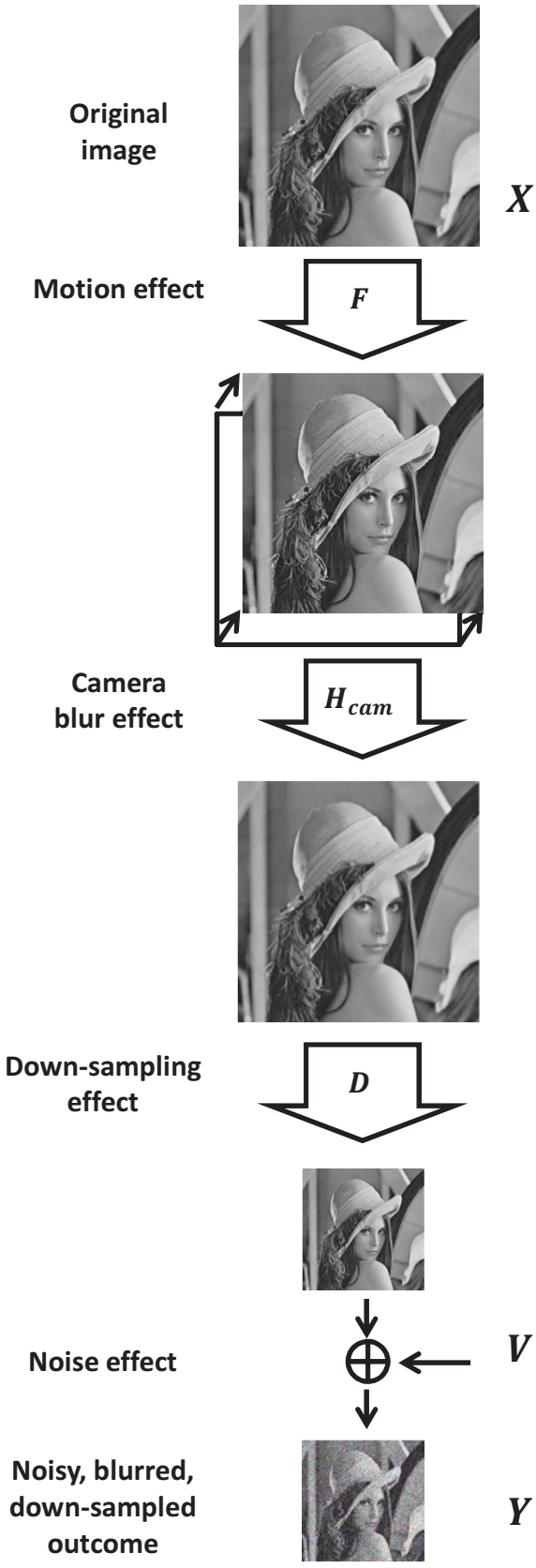


Figure 1.3: Degradation process for SRR

$$\underline{\hat{X}} = \arg\min_{\underline{X}}[\sum_{k=1}^{N} \|D_k H_k F_k \underline{X} - \underline{Y_k}\|_1] \tag{1.2}$$

where $D_k H_k F_k \underline{X}$ is the LRIs generated from the temporal high-resolution image $\underline{X}$. The argmin function gives the $\underline{X}$ at which the error is minimized.

Some forms of regularization must be introduced in the cost function to solve the problem and to minimize the effect of the noise. There are various forms of regularization, such as bilateral total variation (BTV) [2], maximum a posteriori (MAP) [15], projection onto convex sets (POCS) [16], maximum likelihood (ML) [17] iterative backward projection (IBP) [18]. In this study, SRR using BTV method, which is a most common method [2], is focused. BTV is a regularization constraint to remove the noise and to preserve the edges. The regularization term is added to the cost function shown in Equation 1.2.

$$\underline{\hat{X}} = \arg\min_{\underline{X}}[\sum_{k=1}^{N} \|D_k H_k F_k \underline{X} - \underline{Y_k}\|_1 + \lambda \sum_{l=-P}^{P} \sum_{m=-P}^{P} \alpha^{|m|+|l|} \|\underline{X} - S_x^l S_y^m \underline{X}\|_1] \tag{1.3}$$

The second term of Equation 1.3 is the regularization term using BTV. P represents the kernel radius of the selection window, and S_x^l and S_y^m are pixels shifted l and m pixels in the x and y directions. α $(0 < \alpha < 1)$ is the scale weighted coefficient. λ represents the weighted coefficient for regularization.

1.2.2 Essential Problems of Super Resolution Image Reconstruction

SRR algorithm has been proposed by Farsiu and Robinson [2]. They demonstrated the simulation results for still images and videos and claimed to have come up with a working SRR. However, there is no guarantee for the SRR to be able to stand its ground as for as the general use is concerned [19][20]. SRR algorithm requires the existence of the original HRI. In the practical applications, an HRI cannot be prepared and only the LRIs are given: It is unknown whether the HRI as a correct solution for SRR exists. Moreover, the LRIs involving aliasing are created from an original HRI with the insufficient bandwidth limitation and sub sampling. The ability of SRR for the general content cannot be proven since the general content does not have aliasing.

Many SRR with different cost functions have been reported [21][22][23][24][25][26][27][28]

and their performances with peak signal to noise ratio (PSNR) have been demonstrated, however, the improvement of their performance is very slight.

Some papers mentioned the limitation of SRR [21][22][29]. The limit of magnificaton for SRR was analyzed [21][22], and it was 1.6-5.6 depending on conditions [22]. Moreover, the theoretical relationship between the limit of the magnification and the point spread function (PSF) was reported [29]. In these references [21][22][29], a still HRI as the original image was used to create LRIs. They also assume that the shape of the aliasing in LRIs was a sinc function and that the aliasing has the infinite frequency characteristics[29]. However, there is no evidence for these assumptions. Although there is a deep relationship between the aliasing and the performance of the limitations of SRR, their relationship has not been theoretically analyzed. The theoretical limitation of SRR is discussed in Chapter 2.

1.3 Assessment of Super Resolution

SR technology has advanced during this century, and it has been applied to TV or smartphone displays. Image quality is an important feature for display products such as TV and smartphones. SR is used as a marketing term meaning the high performance of the display, and it has become familiar to consumers; however, the performance of the SR function to improve the resolution quality of content is unclear. There is no knowing if there is a difference in the qualities of the displays with different SR functions.

Image quality assessments are generally divided into two types: objective and subjective. An objective assessment is a mathematical evaluation with signal analysis, whereas a subjective assessment is a statistical evaluation with human perceptions.

There is a strong demand for objective assessments, as they are convenient and timesaving owing to the possibility of automatic evaluation using computerized programs. Peak signal-to-noise ratio (PSNR) is most commonly used to measure the quality [30]. PSNR means the ratio between the maximum possible power of a signal and the power of corrupting noise. PSNR is defined via the means squared error (MSE). MSE and PSNR are defined as Equations 1.4 and 1.5.

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(P(i,j) - Q(i,j))^2 \tag{1.4}$$

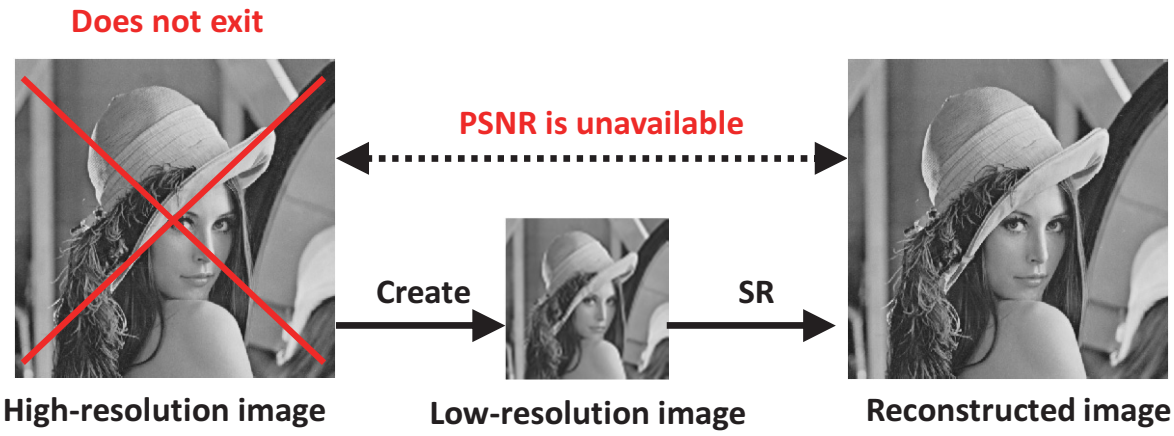


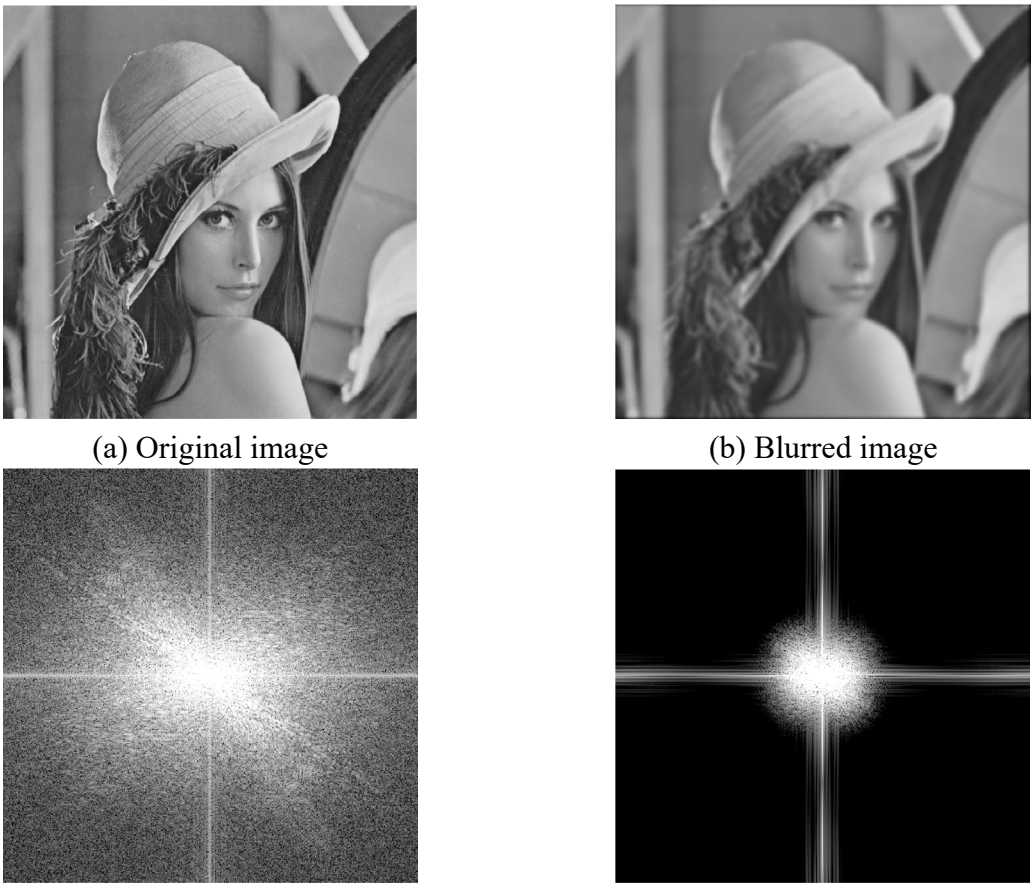
Figure 1.4: PSNR assessment of super resolution

$$PSNR = 20log_{10}\frac{MAX_p}{\sqrt{MSE}} \tag{1.5}$$

where P and Q are two images before and after signal processing. i and j are the coordinate of the image on the horizontal/vertical axis. $P(i, j)$ and $Q(i, j)$ mean the pixel values at the coordinate (i, j). m and n are horizontal and vertical pixel size of P and Q, and MSE are calculated between the corresponding pixels of P and Q. MAX_P in Equation 1.5 is the maximum possible value of the pixels of P, and PSNR can be calculated as the division of the MAX_P by the square of the MSE. The higher value of the PSNR indicates the better image quality. PSNR is the evaluation value for the corruption from the original image, and thus the original image is required for evaluation with PSNR.

Figure 1.4 shows the process for assessing SR capability using PSNR. The image at the left of Figure 1.4 is an original HRI. The low-resolution image (at the center of Figure 1.4) is created from the HRI by applying a low-pass-filter and down sampling. The LRI is also processed with SR and HRI at the right of Figure 1.4 is reconstructed. SR capability can be measured by PSNR between the original HRI and the reconstructed HRI. However, original HRI generally do not exist in the practical situations and thus the PSNR assessment is not applicable.

For evaluating SR performance, it is possible to theoretically analyze resolution improvement by comparing before and after SR processed signals in the frequency domain. Figure 1.5 shows examples of spectral analysis with the 2-dimensional discrete Fourier transform (2D-DFT). Figure 1.5 (a) is an original image, and Figure 1.5 (b) is the blurred image of Figure 1.5 (a). Figure 1.5 (c) shows the 2D-DFT result for Figure 1.5 (a), and Figure 1.5 (d) shows the 2D-DFT result for Figure 1.5 (b). In the spectra, each point represents a particular frequency



(a) Original image (b) Blurred image

(c) 2D-DFT result of (a) (d) 2D-DFT result of (b)

Figure 1.5: Spectrum analysis with 2- dimensional Fourier transform (2D-DFT)

contained in the image. The center point of the spectra means the direct current component. The further point from the center point is the higher frequency. By comparing with Figure 1.5 (c) and Figure 1.5 (d), it can be observed that the resolution of Figure 1.5 (a) and Figure 1.5 (b) can be quantitatively assessed.

Objective assessments have a problem that the results do not reproduce subjective personal opinions completely. PSNR has been criticized for not correlating with the perceived quality measurements [31]. The examples of evaluation results with PSNR are shown in Figures 1.6-1.8. Figure 1.6 is an original image and the Figure 1.7 shows a processed image of Figure 1.6 by drawing a line, and Figure 1.8 shows another processed image of Figure 1.6 by adding Gaussian noise (where the average is 0, and the standard deviation is 4). The PSNR with Figure 1.6 and Figure 1.7 is 36.5199, and the PSNR with Figure 1.6 and Figure 1.8 is 36.5062, thus the quantitative value is almost the same. However, the visual changes of Figure 1.7 and 1.8 are different and it can be observed that the Figure 1.7 is more conspicuous visually for corrupting.

Other objective assessment designed to improve the inconsistencies of the human visual per-

Figure 1.6: Original image

Figure 1.7: Image with line (PSNR= 36.5199 dB)

Figure 1.8: Image with Gaussian noise (σ =4,PSNR=36.5062 dB)

ception with PSNR [32][33] and real-time objective assessment systems [34][35] are reported. However, any objective assessment is not adaptable for assessing viewing systems with commercial displays because they are black-box systems and we cannot take any processed signals from the displays. Thus, the subjective assessment is the only way to evaluate the performance of systems embedded in the displays.

Subjective assessments are used for various purposes, and there are numerous variety of subjective assessment methods. ITU-R BT.500 [36] is one of the most common subjective assessment methods and it is used for determining the bit rate for the HDTV broadcasting [37][38]. A recommendation described in ITU-T P.910 [39] is designed for assessing video quality of multimedia applications, such as videophones. P.912 [40] is designed for assessing the quality and the ability to recognize specific targets in surveillance videos, and ITU-R BT.1788 [41] is designed for assessing quality of videos on various viewing environments such as mobile phone and personal computer. All of the ITU methodologies aim to evaluate the effect of video coding to subjective image quality, and these methods must use a single display during the assessments. Whenever consumers buy display products at a store, they compare the qualities of multiple displays. However, BT.500 cannot be used for product comparisons and there has been

no standard method for such an assessment.

TV or smartphone manufacturers are trying to develop new products competitively and many products with the signal processing technologies are released. However, the quality of the performance is usually unclear, and it requires further assessment to be conducted. However, BT.500 and other ITU methodologies cannot be used.

In Chapter 3, a subjective assessment for multiple displays, which is applicable to product comparisons, is proposed. The assessment method is presented.

1.4 Super Resolution Image Reconstruction and Super Resolution with Nonlinear Signal Processing

Many 4K TV sets are now available on the market, and some of them are advertised that they are equipped with super resolution (SR) technology as a feature. There are several types of SR methods. Super resolution image reconstruction (SRR) and learning based super resolution (LBSR) have been applied to TV sets. Recently, a novel SR with nonlinear signal processing (NLSP) has also been proposed and the additional NLSP hardware for 4K/8K displays has been released [8]. However, the actual visual capabilities of these SR performances on 4K TV sets are unclear. Although SRR has been applied to 4K TV sets, there has been no guarantee for improving resolution of TV content, as discussed in Chapter 2.

The capability of SR methods requires to be demonstrated subjectively since they are ultimately determined by users' impressions. In the related works, the subjective capability of the 4K TV set equipped with LBSR has been demonstrated [42][43]. However, the performance of NLSP and SRR has not been verified. In Chapter 4, the proposed assessment method described in Chapter 3 is applied for assessing quality of 4K TV sets with two different SR methods. The practical ability of SRR on TV sets, as discussed in Chapter 2, is demonstrated with the experiment.

1.5 Super Resolution with Nonlinear Signal Processing for Resolution Enhancement of 4K Videos

4K TV sets are currently available in the market, and some are equipped with the SR function. Recently, SR technology that uses nonlinear signal processing (NLSP) and can work in real time has been reported. Although conventional SR methods aim to improve resolution of enlarged, blurred image, NLSP is also adaptable to images/videos without enlargement. However, the performance of NLSP on 4K TV sets has not yet been prove. In Chapter 5, using the proposed assessment method from Chapter 3, the subjective assessment was performed by comparing 4K videos with and without NLSP on 4K TV sets. Assessment data was statistically analyzed, and the ability of resolution improvement with NLSP on 4K TV sets was demonstrated.

1.6 Super Resolution for Smartphones

Smartphones are used for viewing videos and they can be a substitute for TV sets. For viewing the videos, the quality of the videos is important and improving the quality is frequently required. One of the ways to improve resolution quality is super resolution (SR). In the previous work, a smartphone with super resolution using nonlinear signal processing (NLSP) has been released [44]. NLSP has been applied to TV sets and the ability of resolution improvement on 4K TV sets has been proven as described in Chapter 4 and Chapter 5, however, there has been no guarantee for the NLSP to be effective on small smartphone displays. Many smartphones equipped with image/video signal processing including SR are available; however, there is no knowing how well they actually perform. In Chapter 6, the proposed assessment method from Chapter 3 is also applied for assessing smartphone displays. The assessment was conducted by comparing the video quality of smartphones with and without NLSP or different manufacturers' smartphones. The resolution improvement of smartphone displays with NLSP is demonstrated.

Chapter 2

Super Resolution Image Reconstruction and Imaging Device

2.1 Introduction

Television (TV) set sales reached 25 million in 2010. In 2012, they decreased to eight million, because the market was almost saturated. The prices of large liquid-crystal-display (LCD) TV sets dropped significantly and TV manufacturers were at a turning point. They introduced three-dimensional (3D) TV and 4,000 pixel-resolution (4K) TV to breathe some life into the market. However, available content was insufficient. Although 3D film content is available, 3D TV is unlikely to be successful. 4K satellite broadcasting has begun, and many 4K TV sets are displayed in the best places in mass merchandise outlets. 4K TV is expected to become the successor to high-definition television (HDTV), and the Ministry of Public Management, Home Affairs, Posts and Telecommunications of Japan has released a road map illustrating the anticipated future of the industry for Tokyo Olympic Games in 2020.

HDTV sets became available for sale after HDTV broadcasting began. Currently, 4K displays are already being sold and many homes have 4K TV sets, even before 4K broadcasting has begun. Before 4K TV, TV manufacturers had actively sought broadcasting services and had tried to develop the best TV sets for broadcasting. TV manufacturers had been trying to improve broadcasting image quality to make it the best possible. However, 4K TV is different. This time TV manufacturers are selling 4K TV sets ahead of the broadcasting services. They have tried to promote the 4K TV, but 4K broadcasting service is limited. To compensate for the lack of 4K content, attempts are being made to upgrade HDTV content to 4K reso-

lution to enable it to work with 4K displays. The resolution of 4K is $3,840 \times 2,160$; thus, the size of the HDTV content has to be doubled horizontally and vertically. However, enlarging an image always causes blurring. Some TV manufacturers have announced that their 4K TV sets are equipped with super resolution (SR). However, they do not provide access to their algorithms, except for super resolution image reconstruction (SRR). There are many papers discussing SRR[2][3][4][5][6][21][22][23][24][25][26][27][28][29].

Farsiu and Robinson [2] proposed an SRR algorithm and reported its simulation results for still images. They also showed the simulation results of eight-frame video and claimed successful SRR. However, the images he used involved aliasing created from high-resolution still images [26]. They used a special infrared video with aliasing, which was created with a still camera that did not perform general camera operations such as panning and tilting, as we see in TV and film content. SRR cannot be expected to work for general content[19][20][45][46]. There is no report claiming that SRR can improve resolution for TV or film video content that has sufficient resolution without aliasing and that works with general camera operations.

In contrast, many SRR results for still images have been reported [2][3][4][5][6][23][24][26] [27][28]. These reports have increased during this century and have been associated primarily with newer imaging devices such as charge-coupled devices and complementary metal oxide semiconductors.

One paper reported that SRR does not work for images that are enlarged twice for practical applications [47], however, it provided no explanation. Baker and Kanade introduced the hallucination algorithm to overcome the enlargement limitation of SRR [21]. Lin and Shum cited their paper[21] and pointed out that no explicit limits were deduced in Baker's paper [22]. Lin and Shum demonstrated that, in general, the enlargement limitation of SRR was 1.6 and that it could be as high as 5.6 in suitable conditions[22]. However, both papers[21][22] used a still high-resolution image (HRI) as the original image and low-resolution images (LRIs) generated from the original HRI. Therefore, they started from a point where the correct answer was known to exist. In practical applications, we do not have an HRI and only the LRIs are given. Thus, we need to create an HRI from the given LRIs. We do not know if the HRI will exist or if there will be a correct answer for the SRR to converge to.

Another limitation of SRR was reported by Tanaka and Okutomi. They theoretically demonstrated the relationship between the limitations of SRR and the point spread function (PSF)[29]. In reference[29], they assumed that the shape of the aliasing in LRIs was a sinc function ($Sinc$)

and that the aliasing continued infinitely. However, they did not show evidence for these assumptions.

SRR studies increasing the enlargement limitation of SRR are reported separately; however, the relationship between them is still unclear. In reality, the performance of digital imaging devices (CMOS and CCD) and the limitations of SRR are not connected in a manner that can be explained theoretically. In this chapter the relationship between SRR and imaging devices is analyzed, and the SRR enlargement limitation is investigated. The validity of the assumptions in reference [29] is theoretically discussed.

2.2 SRR and Aliasing

2.2.1 Nyquist Sampling Theorem and Subsample

SRR has a deep relationship with aliasing, and LRIs must have aliasing. LRIs without aliasing cannot be used to reconstruct an HRI with SRR [48]. In this section, the relationship between SRR and aliasing is analyzed theoretically. To simplify the discussion, a simple one-dimensional signal shown in Figure 2.1(a) is used as an example of HRI. $f(t)$ is sampled with a period of $t = T$, which is written as follows:

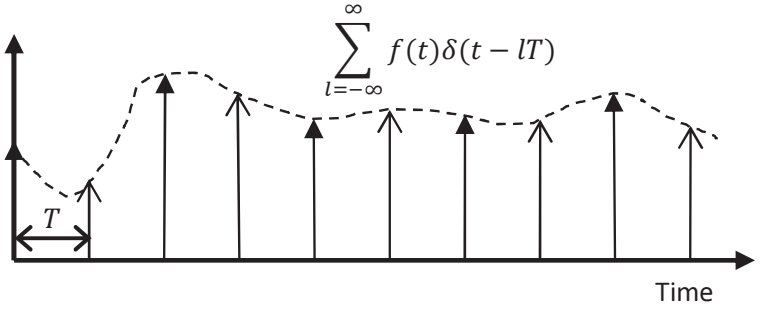
$$\sum_{l=-\infty}^{+\infty} f(t)\delta(t - lT) \tag{2.1}$$

where l is an integer. A 2:1 subsample is generally used to create LRIs. According to the Nyquist sampling theorem, the bandwidth of the image must be limited to avoid causing aliasing. For a 2:1 subsample, the bandwidth is limited by a low-pass filter (LPF) to half the bandwidth of the original image. The bandwidth limited $f(t)$ with LPF is expressed as $\tilde{f}(t)$, and Equation 2.1 becomes Equation 2.2.

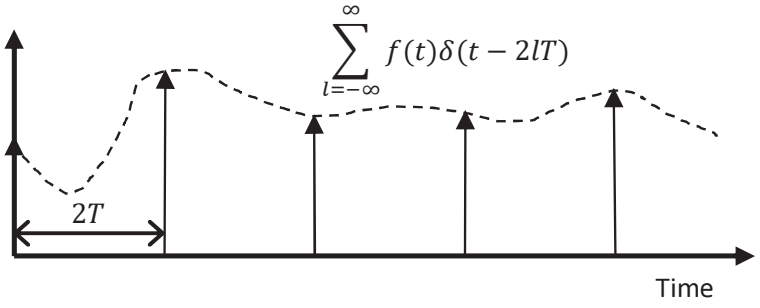
$$\sum_{l=-\infty}^{+\infty} \tilde{f}(t)\delta(t - lT) \tag{2.2}$$

After the 2:1 subsampling, Equation 2.2 becomes Equation 2.3.

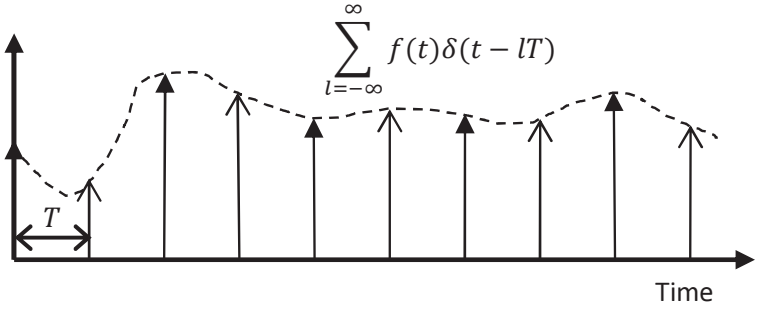
$$\sum_{l=-\infty}^{+\infty} \tilde{f}(t)\delta(t - 2lT) \tag{2.3}$$



(a) Original image



(b) 2:1 subsample



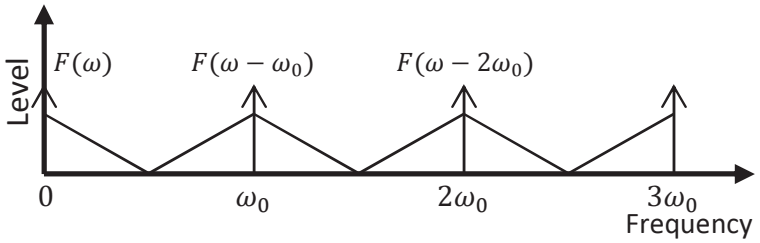
(c) Reconstructed image with SRR

Figure 2.1: Subsample and SRR（time domain）

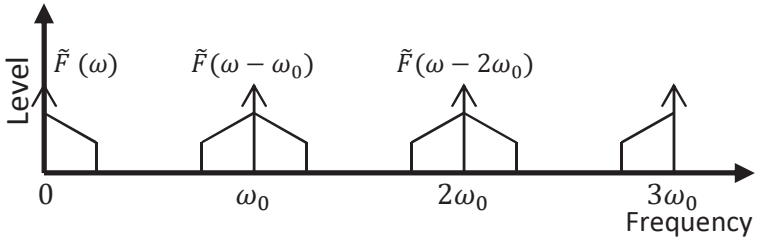
According to the SRR study [2] and others using subsampled images as LRIs, the HRI shown in Figure 2.1(c) can be reconstructed.

In this section, this signal processing process is analyzed with Figure 2.2 in the frequency domain. The sampling frequency is $\omega_0 = 2\pi/T$. The Fourier transforms of $f(t)$ and $\tilde{f}(t)$ are defined with $F(\omega)$ and $\tilde{F}(\omega)$ respectively as follows.

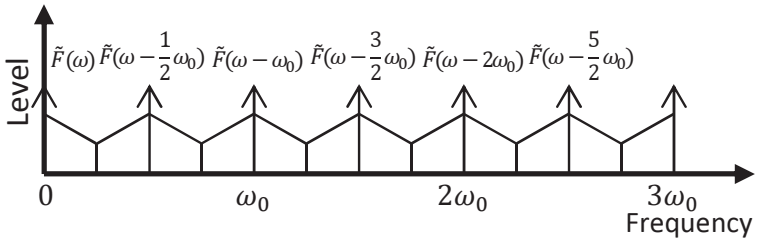
$$\begin{aligned} &\int_{t=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} f(t)\delta(t-lT)e^{-j\omega t}dt \\ &\qquad = \sum_{n=-\infty}^{+\infty} F(\omega - n\omega_0) \\ &\int_{t=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} \tilde{f}(t)\delta(t-lT)e^{-j\omega t}dt \end{aligned} \tag{2.4}$$



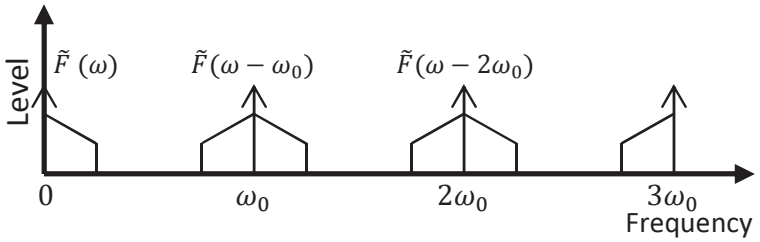
(a) Spectra of original image



(b) Spectra processed with the appropriate LPF



(c) Subsampled spectra with the appropriate LPF

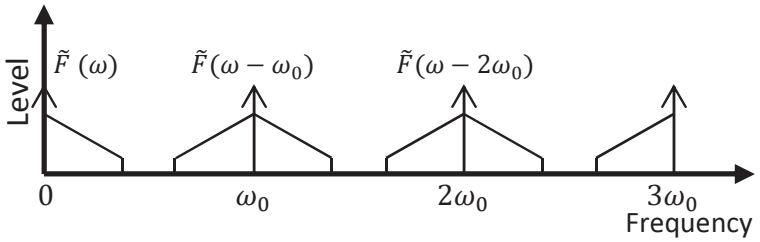


(d) Spectra of reconstructed image

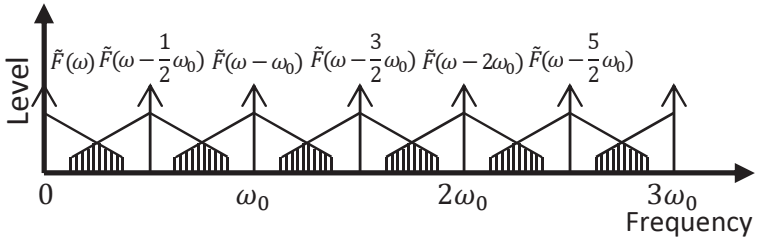
Figure 2.2: Subsample without aliasing (frequency domain)

$$= \sum_{n=-\infty}^{+\infty} \tilde{F}(\omega - n\omega_0) \tag{2.5}$$

Where j denotes imaginary number. The spectrum expressed with Equation 2.4 is shown in Figure 2.2(a). In the SRR algorithm, the spectrum shown in Figure 2.2(a) is processed with LPF. LPF limits the bandwidth of images and videos and the processed images and videos become blurry. The LPF pre-processed image is $f(t)$ (Equation 2.1) and the post-LPF processed image is $\tilde{f}(t)$ (Equation 2.2). The sampling frequency of $F(\omega)$ is ω_0 and the maximum frequency elements of $F(\omega)$ are expected to be less than $\omega_0/2$. The frequency elements of $F(\omega)$ shown in Figure 2.2(a) are expected to be less than $\omega_0/4$ for 2:1 subsampling with LPF. Figure 2.2(b)



(a) Spectra processed with the imperfect LPF



(b) Subsampled spectra with the imperfect LPF

Figure 2.3: Subsample with aliasing (frequency domain)

shows the band-limited spectrum of $\tilde{F}(\omega)$. The 2:1 subsampling result of Equation 2.5 is written as Equation 2.6.

$$\int_{t=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} \tilde{f}(t)\delta(t-2lT)e^{-j\omega t}dt = \sum_{n=-\infty}^{+\infty} \tilde{F}(\omega - n\omega_0/2) \tag{2.6}$$

The spectrum of Equation 2.6 is shown as Figure 2.2(c). Figure 2.2(c) has half the pixels of Figures 2.2(a) and (b). Although Figure 2.2(c) is 1:2 interpolated to create the same number of pixels, as in Figures 2.2(a) and (b), it actually becomes Figure 2.2(d) and has only half the bandwidth of Figure 2.2(a). This means that the pixel number does not equal the resolution but the bandwidth equals the resolution. There have been misunderstandings that the resolution of HDTV increases when we see it with 4KTV. Although 4KTV has four times as many pixels as HDTV, HDTV content on 4KTV has exactly the same resolution as HDTV. This also applies to the relationship between Figures 2.2(a) and (d).

According to SRR studies, using many LRIs shown in Figure 2.2(c) enables the HRI shown in Figure 2.2(a) to be reconstructed. However, as discussed in this section, this does not work. LRIs are just LRIs, and it is impossible to reconstruct HRIs from LRIs. If it works, then HDTV resolution can be used with blurry National Television System Committee standard videos. However, LPF must be used as a prefilter when subsampling is conducted. This is the basic

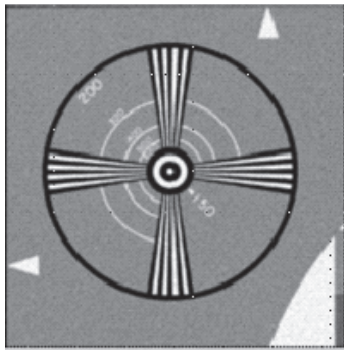


Figure 2.4: HRI

Figure 2.5: Blurry LRI

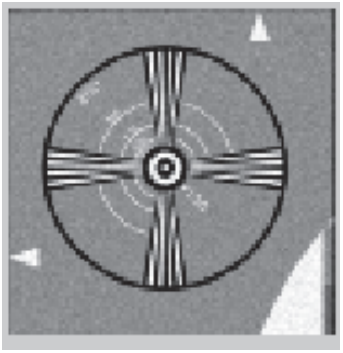


Figure 2.6: LRI with aliasing

theory in digital signal processing. Using an LPF with the Nyquist sampling theorem, we have shown that SRR does not work.

SRR research is conducted widely throughout the world with new studies appearing frequently [27][28]. The reason is very simple; these studies use LRIs with aliasing. If we begin an SR study using LRIs with aliasing, many methods can be proposed using different forms of aliasing to reduce the number of ways to combine LRIs' pixels. Since redundant LRIs are prepared for conducting SRR, many options are available. In Figure 2.2 the pre-filter limits the bandwidth and aliasing did not occur by subsampling. This is the well-known Nyquist sampling theorem. Figure 2.3 is an example that the Nyquist sampling theorem is not satisfied. Figure 2.3(a) shows the spectrum processed with a wide-band LPF, and Figure 2.3(b) shows the 2:1 subsampling result of Figure 2.3(a). The overlapped elements in Figure 2.3(b) indicate aliasing. SRR uses the aliasing to reconstruct an HRI. Figure 2.4 is an example of an HRI. A generally low-resolution for Figure 2.4 yields Figure 2.5. However, SRR cannot reconstruct HRI from this type of low-resolution image. SRR requires LRIs with aliasing, as shown in Figure 2.6. The aliasing shown in Figure 2.6 appears only when the bandwidth of the LPF prefilter does not satisfy the Nyquist sampling theorem.

2.2.2 Still Image and Video

Although SRR was studied for still images, it was applied to videos [2][24]. Video frames were used as LRIs to reconstruct an HRI. However, still images and video frames have completely different properties. Figure 2.7 was taken with a digital camera, and Figure 2.8 was taken with a video camera. The still image shown in Figure 2.7 does not have motion blur, whereas one of the video frames as shown in Figure 2.8 has it. Motion blur occurs based on the movement

Figure 2.7: Moving object (still camera)

Figure 2.8: Moving object (video camera)

of objects or cameras when it was shot. The shutter speed is an important parameter to control the time length of exposure for the imaging device to light. The slower shutter speed causes greater motion blur. General shutter speed of digital video cameras is 1/60 second, whereas that of still digital cameras could be shorter than 1/500 second. Therefore, a large part of all video content contains certain amount of motion blur, and aliasing does not naturally exist. We cannot obtain LRIs with aliasing shown in Figure 2.6 from the video frames. It means that SRR cannot increase the resolution of the general video content.

2.3 Essential Issue of SRR

In SRR studies, an HRI is prepared and processed with LPF, and LRIs are created with sub-sampling. The created LRIs are processed with SRR, and an HRI is reconstructed. In actual applications, only blurry LRIs are provided, and the HRI is not prepared. As discussed in the previous section, SRR can be used to create an HRI only when LRIs with aliasing are provided, and the reconstructed HRI is nothing more than the LRIs with the aliasing removed. In SRR

Figure 2.9: HRI

Figure 2.10: LRI

Figure 2.11: Reconstructed HRI

studies, the LPF characteristics are not shown and wide-band LPFs are used to create aliasing. The reconstructed HRI with SRR is compared with the original HRI and peak signal-to-noise ratio (PSNR) values are compared with previous SRR studies.

The original HRI appears to be prepared to calculate PSNR, but the original HRI is, in reality, prepared to create LRIs with aliasing. LRIs with aliasing are not sufficient for SRR to work efficiently. Pixels of all LRIs must have different phases. SRR detects sub-pixel phase differences in the pixels of LRIs and moves the pixels within sub-pixel precision to reduce aliasing. This requires the pixels in LRIs to have sub-pixel differences. In practical situations, it is necessary to move a camera with sub-pixel precision when we take LRIs for SRR. It is almost impossible to take more than 16 images under this condition. Although sub-pixel LRIs are prepared, there is no guarantee that an HRI will exist as a convergence image, because LRIs are not created from a single HRI.

Preparing an HRI is the best and simplest method to overcome these difficulties. Processing the prepared HRI with a wide-band LPF and creating LRIs with aliasing is sufficient. This method has been tested in SRR studies; for example, one well-known study [2] on SRR uses the well-known image Lena and a wide-band LPF to create LRIs with aliasing. Using Figure 2.9 as the HRI, the 16 LRIs shown in Figure 2.10 are created. The subsample ratio in the process is 2:1 horizontal and vertical. Although four LRIs have the same amount of information as one HRI, 16 LRIs with four times more information than the HRI are required for SRR and more than 100 iterations are necessary to create an HRI [2]. Figure 2.11 is the HRI reconstructed with SRR. If the HRI is completely reconstructed, then Figure 2.9 and Figure 2.11 with their frequency characteristics are the same. However, their frequency characteristics will be different.

Figure 2.12 and Figure 2.13 are the two-dimensional Fourier transform (2D-FT) results of

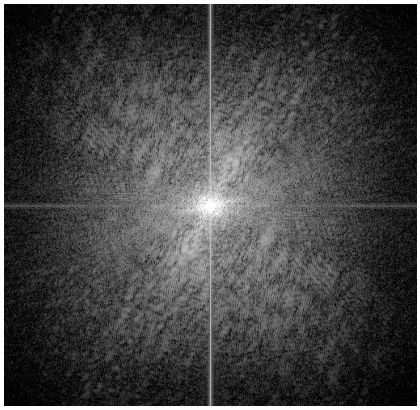


Figure 2.12: 2D-FT of Figure 2.9

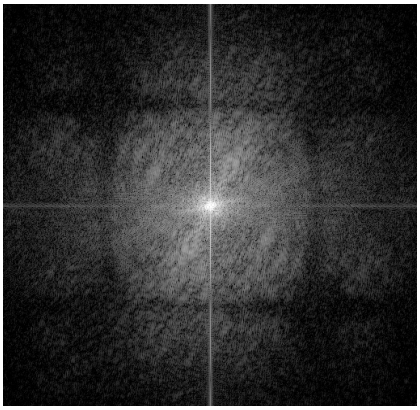


Figure 2.13: 2D-FT of Figure 2.11

Figure2.9 and Figure 2.11, respectively. Comparing Figure 2.12 and Figure 2.13, frequency characteristics are different. The spectrum in Figure 2.13 contains null areas that do not appear in Figure 2.12. These null areas are caused by the 2:1 subsampling when the HRI is created from LRIs. If the size of the LRIs is one quarter, the null areas appear in the half frequency of those areas horizontally and vertically [46] [49].

The null areas depend on the subsample ratio. The original and reconstructed HRIs appear to be identical only when we compare them visually. However, analyzing them in the frequency domain reveals that the reconstruction is not complete. This paper uses the SRR algorithm from the study [2]. However, other SRR proposals involve the same problems, because they are connected to essential SRR algorithms.

2.4 Imaging Device

As discussed in the previous section, the Nyquist sampling theorem requires LPF intended for subsampling to be designed not to cause aliasing. Subsampling must be conducted after LPF processing. However, SRR studies have used wide-band LPFs and LRIs with aliasing for processing with SRR. An HRI can be reconstructed by reducing aliasing in the LRIs and combining them.

Although film (silver salt) cameras have long been widely used, they have most recently been replaced with digital cameras. Images shot with film cameras do not have aliasing. When SRR studies use images captured with film cameras, they create LRIs with aliasing, wide-band LPFs, and subsampling [2]. Images captured with a digital camera also can have aliasing

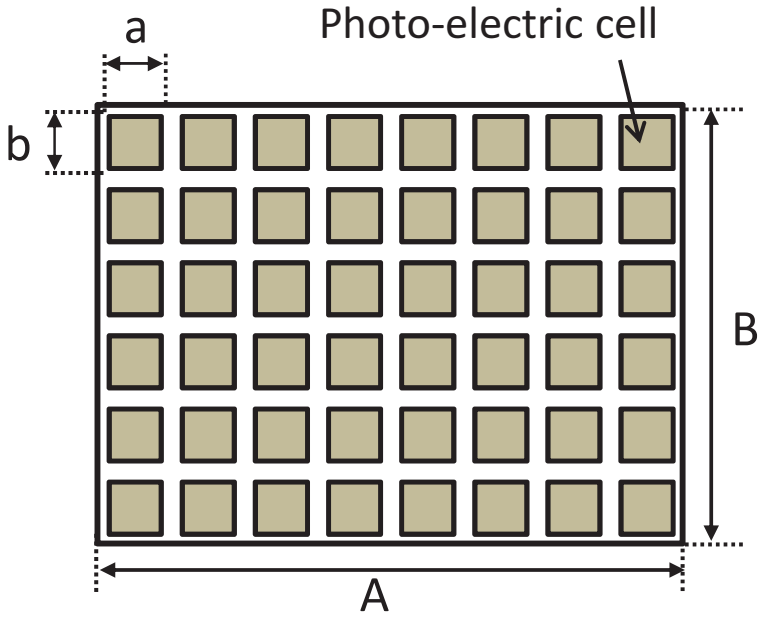


Figure 2.14: Imaging device

[50]. Imaging devices for photoelectric conversion consist of cells shown in Figure 2.14. The ideal condition of photoelectric conversion is that the photoelectric cells are small and have the same characteristics as the δ function. In engineering technology, it is impossible to develop an infinitely small photoelectric conversion cell. Photoelectric cells with finite sizes and areas have the aperture effect, and the high frequency elements are attenuated in inverse proportion to the size of the photoelectric imaging cells.

Figure 2.14 is a model of a photoelectric device that has $m \times n$ cells. The size of cells is $a \times b$, and that of an imaging sensor is $A \times B$. The horizontal sampling frequency is m and the vertical sampling frequency is n. The horizontal axis is μ and the vertical axis is ν. The horizontal and the vertical fill factors are $\frac{ma}{A}$ and $\frac{nb}{B}$ respectively. The aperture effect of the finite cell is given by Equation 2.7.

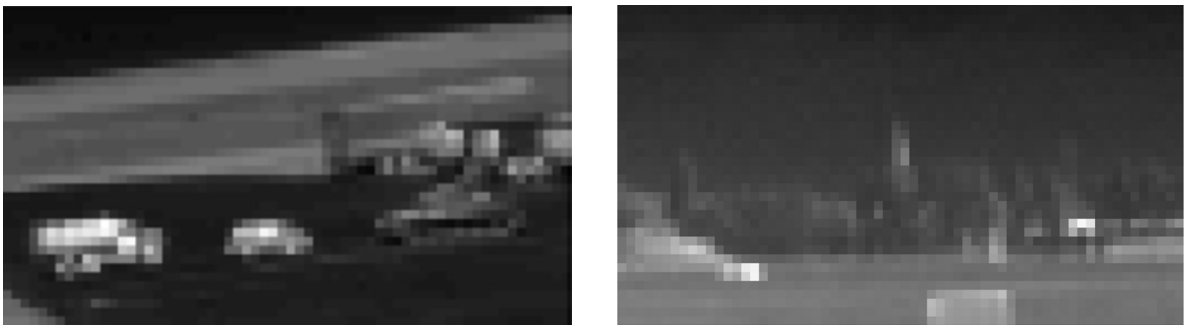
$$Sinc\left(\frac{ma}{A}(\frac{1}{m}\mu)\right) Sinc\left(\frac{nb}{B}(\frac{1}{n}\nu)\right) \tag{2.7}$$

There are spaces between cells in Figure 2.14. The fill factor becomes 100% when the cells are fully tiled and no space remains [51][52][53][54][55][56].

The amount of input to each cell increases in proportion to the fill factor, and it becomes possible to image in dark areas. Currently, the fill factor of imaging devices equipped with commercial digital cameras is almost 100%. In contrast, the aperture effect becomes notice-

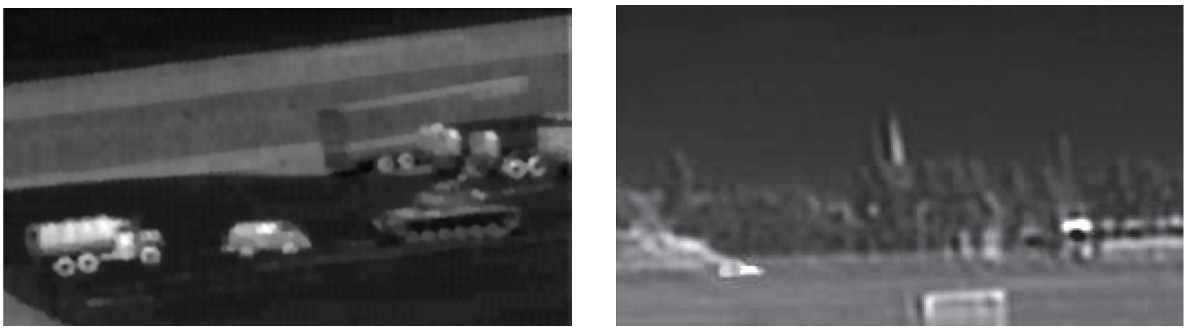
able in proportion to the increases of the fill factor, and high frequency elements are attenuated. When the fill factor becomes 100%, $A = ma, B = nb$ holds. When the sampling frequencies $\mu = m$ and $\nu = n$ are substituted into Equation 2.7, the responses are null. The high-frequency elements are also attenuated. These phenomena are caused by the aperture factor. The fill factor was 15% – 41%[51][52][53] and became 50% in 2006[54]. Recently it was improved to 100% [55][56]. The theoretical analysis given by Tanaka and Okutomi[29] used Equation 2.7. However, they did not elucidate the effect of the shape of the sinc function on the frequency characteristics of aliasing. The aliasing frequency characteristics affect the quality of the reconstructed HRI. As a result, the resolution of HRI is affected by the fill factor of the imaging device. This discussion theoretically explains the relationship between the reconstructed HRI by SRR and the fill factor of the imaging devices. This point has not been mentioned previously.

Here we assume that the fill factors of the current imaging devices are 50%-90%. When the fill factor is 50%, the frequency characteristics become zero at double the sampling frequency. When it is 90%, the frequency characteristics become zero at the sampling frequency $\times 1.1$. These considerations show that there are higher frequency elements than the sampling frequency if the fill factor is less than 100%. This means that aliasing exists as aliasing in the images created with the fill factors [57]. Recent imaging devices have a high percentage of fill factors to cope with poor lighting conditions. However, at the beginning of the century, these were lower than the current ones. If the fill factor was close to 50%, the images have the frequency elements up to double of the sampling frequency. In this case, SRR can create elements of twice higher frequency than the sampling frequency. However, the relation between the fill factor and aliasing explains that images created with low fill factor imaging devices have aliasing that can be used with high frequency elements to improve resolution with SRR. In this century, imaging devices have advanced and digital cameras are equipped with those devices. Early generation digital cameras were equipped with low fill factor imaging devices, with the result that images with aliasing became widespread. SRR research began in this environment, because there were so many images with aliasing. Efforts to increase the fill factor of photo imaging devices have continued and enabled the production of high fill factor imaging devices to cope with poor lighting conditions [58]. This narrows the application fields of SRR, because images created with high fill factor imaging devices have less aliasing.



(a) (b)

Figure 2.15: LRI for SRR [2][24]



(a) (b)

Figure 2.16: Reconstructed HRI[2] [24]

2.5 Block Shaped Aliasing

The references [2][24] reported SRR for video. Several frames of the videos are used as LRIs and HRI is reconstructed using SRR. Figure 2.15(a), (b) are the LRIs, and Fig.2.16(a), (b) show the SRR processed results. Comparing Figure 2.15 (a) with Figure 2.16 (a), and Figure 2.15 (b) with Figure 2.16 (b), respectively, shows that resolution is improved with SRR. However, the resolution of the videos used in the studies was very low, and they have block shapes as their imaging cells, as shown in Figure 2.15. These block shapes are the result of aliasing. SRR studies [2] and [24] use LRIs created with low-resolution images to include aliasing. Using several frames as LRIs, they call aliasing-reduced images HRIs. These videos are created with still cameras and do not involve camera work such as panning and tilting, which general TV and cinema content has. If videos include such camera work, it becomes difficult to reduce aliasing, because quarter-pixel accuracy motion vector detection is necessary for SRR processing. These considerations show that SRR for videos just reduces aliasing, as in the case with SRR for still images, and SRR cannot create higher-resolutions than those of the LRIs used to create them.

Despite the large amount of available video content, there have been no low-resolution videos such as those shown in Figure2.15 (a) or (b) in broadcasting, cinema, digital video disc, or Blu-ray disc content. They have much higher-resolution than Figures 2.15. Improving unusable low-resolution videos to a somewhat higher-resolution SRR is not a practical technology for general videos.

Figure 2.16 does not have block-shaped aliasing. Using figures of comparable quality as LRIs for SRR is considered to be capable of improving resolution, because aliasing in these images has been removed. Although commercial HDTV equipped with SRR is available, its resolution is not as good as that of HDTV sets without any type of SR technologies [45]. Subjective assessments for 4K TV have been conducted [59] with results reported similar to those arrived here [45]. This fact is consistent with the discussion in this chapter regarding the inability of SRR to improve the resolution of HDTV content, as a consequence of not having block-shaped aliasing. The interlace artifact is the only aliasing in HDTV content. SRR can be used to reduce interlace aliasing. However, it has been shown that interlace aliasing can be removed using inter-field signal processing, as with motion vectors [60]. Inter-field signal processing with motion vectors is a much more concise method than SRR.

2.6 Conclusions

The relationship between imaging device and aliasing was discussed, and the limitations of SRR were examined theoretically. SRR is one of the most common and popular SR technologies, and many studies of it have been proposed. However, SRR just removes aliasing in LRIs. If an HRI is prepared, it can be used without any processing. Creating LRIs from an HRI and then reconstructing the HRI is of no use. Reconstructing an HRI with SRR is similar to assembling the parts of a jigsaw puzzle. In practical applications, only one LRI is given. It is required to create an HRI from the LRI; it is not possible to do so using SRR.

Chapter 3

Subjective Assessment Method for Multiple Displays with Super Resolution

3.1 Introduction

Digital high-definition television (HDTV) broadcasting has begun, and home-use television (TV) displays have evolved from cathode-ray tubes to liquid crystal displays. In 2011, 4K TV sets, which have four times the resolution (3,840 × 2,160) of HDTV (1,920 × 1,080), were introduced in the market, and in 2014, 4K satellite broadcasting started in Japan. However, 4K video content is still not widespread, resulting in the release of 4K TV sets ahead of the 4K broadcasting. Almost all TV content available currently is in HDTV, and thus, format conversion is necessary to play conventional HDTV content on 4K TV sets. However, enlarging an image causes blurring.

To improve image/video quality, almost all TV sets are equipped with signal processing technologies such as an enhancer. However, the enhancer only enhances the edges of an image and cannot actually improve resolution. Super-resolution (SR) technology is one way to increase resolution. 4K TV sets equipped with SR have been released by some manufacturers.

A popular SR method is super-resolution image reconstruction (SRR), which uses multiple low-resolution images to reconstruct a high-resolution image [2]. Although 4K TV sets equipped with SRR are available, the inability of SRR to improve the resolution of the TV content has been discussed [61]. Note that SR is a catchphrase used in TV marketing, and the performance of SR on TV sets is not guaranteed.

Although the assessment of SR performance on TV sets is required, there is no method for

such an assessment at present. The simplest evaluation of SR is signal analysis, which is a comparison of the signals with and without SR in the frequency domain However, there is no way to measure the signals after the SR processing on the TV sets. As signal analysis cannot be used, a subjective assessment is the only way to evaluate the performance of SR embedded in video devices.

There are various TV sets equipped with signal processing technologies including SR by different manufacturers. Consumers compare these products when they purchase a TV set. Although image quality is frequently considered in the decision, there is no way for consumers to evaluate the relative merits of image quality between the products. A standardized assessment methodology for television video quality is described in BT.500 [36]. However, BT.500 is not adaptable for assessing multiple displays leading to a product comparison. In other method, a paired comparison [62], was applied to image quality assessments [63], and also applied to assess different display panels [64]; however, these assessments are for still-images. The typical use for TV sets is video appreciation. The usefulness of the method for video assessments on multiple displays has not been verified. The purpose of this study is to propose an assessment method for multiple displays enabling to obtain consumers' subjective impressions. In this chapter, the methodology of the proposed assessment method is presented.

3.2 Subjective Assessment

Subjective image quality is a psychophysical quantification of how a viewer perceives images and videos. Human perceptions vary individually. Thus, statistical analysis is essential to validate the reproducibility of psychophysical quantity measurements. Statistical analysis is based on the assumption that the evaluations for each sample follow different Gaussian distributions if there is a difference in the image qualities of the stimuli. Hypothesis tests, such as t-tests and F-tests, assess the error probability that an observation difference in samples is just the result of random noise. If the error probability is significantly low, then a significant difference between the samples is observed, and this indicates a reproducible result. Thus, significant differences must be detectable because the result without them makes no sense. Note that psychophysical quantities are susceptible to various factors, and we must carefully select the assessment method and experimental conditions to obtain reproducible measurements.

One of the most common subjective assessment tools is BT.500 [36]. BT.500 is useful in

evaluating the relationship between subjective image quality and bitrate of the image coding. However, BT.500 assessments must use a single display to present assessment videos, and it is not directly adaptable for multiple display assessments. A paired comparison method and ranking method are commonly used for sensory evaluation and it is adaptable to multiple display assessments involving simultaneous comparisons. The ranking method is a comparison of all samples, whereas the paired comparison is that of every pair of samples. The ranking method is inferior to the paired comparison method with respect to the sensitivity of the assessment [71]. In this study, the paired comparison method is combined with some of the BT.500 experimental conditions, such as the eligibilities of test sequences and observers. The proposed method copes with the inadaptability of BT.500 assessments to multiple display assessments.

3.3 Proposed Method

3.3.1 Scheffe's Paired Comparison

A paired comparison method is a round-robin paired comparison that helps in obtaining a statistical order for image quality. Several paired comparison methods have been reported [62][65]: Thurstone's method [65] assesses the superiorities of pairs, whereas Scheffe's method [62] assesses the degrees of their superiorities. In this study, Scheffe's method is selected because it can obtain more detailed information for an assessment than Thurstone's method.

The process of Scheffe's paired comparison method is as follows. Using a pair of target and reference samples, observers score their quality on a five-grade scale from -2 to +2 (+2: Excellent, +1: Good, 0: Even, -1: Poor, -2: Bad). The same assessments are repeated for all pairs of samples. Figure 3.1 shows the actual experiment using the paired comparison method. The observer compares the quality of multiple displays placed together. This situation reproduces an environment in which shoppers compare multiple items at a store.

3.3.2 Observers and Test Sequences

BT.500 specifies that observers must be non-experts who do not work in the video industry and have normal visual acuity and color vision. Moreover, the number of observers must be at least 15. The proposed method adopts these conditions.

Figure 3.1: Paired comparison

BT.500 specifies that each test sequence used in the assessment must last for 10-15 s and at least four test sequences must be used. The proposed method also adopts these specifications. Although BT.500 does not specify assessment areas, it is not easy for non-expert observers to recognize the difference in quality. To stabilize the observers' decisions, the proposed method specifies assessment areas that make it easier to assess image quality in each of the test sequences.

3.3.3 Experimental Environments

A training session is conducted in advance to explain the meaning of high- and low-quality images and the experimental method to observers. The experimental process and evaluation points are effectively explained to observers using a dummy test sequence. The test sequence is repeated for each display during the assessment. There is no time limit for the assessment. The observers can freely move to the front of each display and view the test sequences to decide on their opinion. BT.500 specifies an observation angle of ±30° from the front of the screen. The proposed method maintains this angle, and the observers are asked to view the videos from the front of the display. A viewing distance of three times the display height is specified in BT.500; however, the appropriate viewing distances vary for individuals according to their visual acuity. In the proposed method, observers can freely select their viewing distance during the assessment.

Chapter 4

Subjective Assessment of 4K TV Set with Super Resolution Image Reconstruction

4.1 Introduction

At present, 4K TV sets are available on the market. However, 4K video content is still not sufficient: almost all TV content is in high-definition television broadcasting (HDTV). Images /videos with insufficient resolution are up-converted to the resolution of the display. However, enlarging an image always causes blurring.

Enhancer (or unsharp mask) [1] is one way to improve quality of the image/video in resolution. This method has simple algorithm and it is easy to enhance images and videos in real time: however, enhancer cannot actually improve resolution because it can only enhance the edges of the image/video.

Super resolution is another way to improve resolution. There are some types of SR method. Super resolution image reconstruction (SRR) is the most popular SR method and this method has been applied to TV sets. However, actual performance of TV sets with SRR has been unclear as none of them has solid evidence.

The image quality of displays is one of the important factors affecting the consumers' purchase judgements. There are many TV sets available and they are equipped with signal processing technologies including SR. However, actual performance of SR on TV sets has not become clear.

In this chapter, subjective assessments of TV sets with different SR methods are conducted by using the proposed assessment method described in Chapter 3. In authors' related works

an SR method [7] with nonlinear signal processing (NLSP) was proposed. The performances in resolution quality of NLSP and SRR are assessed, the discussion of SRR in Chapter 2 are proved by the experiments.

4.2 Subjective Assessment Experiments

Subjective assessment experiments are conducted to verify the ability of different SR methods. The stimuli are the 4K (3,840 × 2,160) signals up-converted from a HDTV (1,920 × 1,080) signal by three methods: NLSP, SRR, and the Lanczos filter [66], which is a common interpolation algorithm.

4.2.1 Experiments

Two experiments were conducted to consider the effect of different display panels. The stimuli are the 4K signals up-converted from a 2K (1, 920 × 1, 080) signal by three methods: NLSP, SRR, and the Lanczos filter [66], which is a common interpolation algorithm.

Experiment 1 uses two consumer-grade 4K TV sets, as shown in Figure 4.1, and experiment 2 uses the 4K TV set shown in Figure 4.1 and a professional 4K display, shown in Figure 4.2. The 4K TV set shown in Figure 4.1 is equipped with SRR and implements it when the resolution of an input signal is less than that of its display resolution (4K), but it does not work with the same resolution. The system diagrams for two experiments are shown in Figures 4.4 and 4.5, respectively. They are the same experiment except for the types of display devices. The solid arrow indicates the process for presenting the NLSP or original video signal. The dashed arrow indicates the process for presenting the SRR video signal. The video player outputs a video signal with 2K resolution. For the NLSP process, the 2K signal is input to the NLSP hardware and is first up-converted to 4K using the Lanczos filter. Then, the NLSP is implemented with the SR processing on the hardware enabled (ON). If this setting is disabled (OFF), the unprocessed 4K signal is output. The output signal is either displayed through the 4K TV set or the professional 4K display. For the SRR process, the original 2K signal is directly input to the 4K TV set. The signal is then up-converted to 4K by the SRR embedded in the 4K TV set and displayed through the 4K TV set.

Table 4.1: Rating scale

Score	Description
2	Excellent
1	Good
0	Even
-1	Poor
-2	Bad

4.2.2 Assessment Method

Scheffe's paired comparison method was used for the assessment. Using a pair of 4K TV sets, observers scored the TV sets on a 5-grade scale (-2 to +2). The results of the assessment were statistically analyzed using analysis of variance. Prior to the experiment, we conducted a training session to explain resolution and the experimental method to each observer. The observers were asked to assess only resolution and ignore other factors, such as noise and color.

4.2.3 Experimental Set Up and Environments

Thirty non-expert observers participated in the experiments. The observers assessed image quality using the five-grade scale from -2 to +2. They were asked to assess resolution only. Other quality factors, such as noise and color, were not considered in the assessment. Five test sequences were used in each experiment: the 4K test sequences shown in Figure 3.2 were used in experiment 1, and the 2K test sequences shown in Figure 3.3 were used in experiment 2. These sequences do not include pan and tilt scenes. The assessment areas indicated by ovals in Figures 3.2 and 3.3 were specified. These areas have high-resolution elements and are appropriate for recognizing resolution differences.

4.3 Results and Discussions

The analysis process is described by following the results for the "Ruins" sequence. Here, the stimuli for experiments are the processed signal with NLSP (NLSP), the processed signal with SRR (SRR), and the processed signal with Lanczos method (Lanczos).

Figure 4.1: 4K TV

Figure 4.2: 4K display

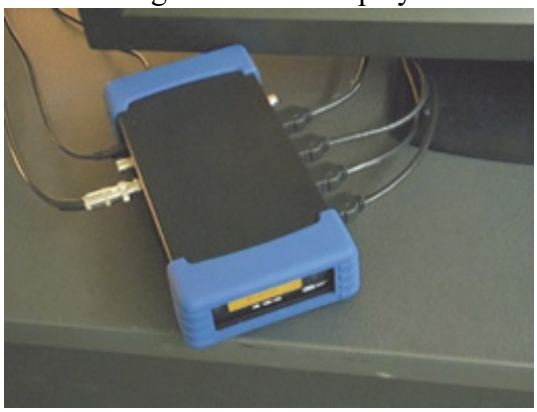


Figure 4.3: NLSP hardware

4.3.1 Experiment 1

Table 4.2 is the cross table for the "Ruins" sequence. Row i indicates the reference stimulus for comparison, and column j indicates the target stimulus. The values in Table 4.2 are the sums of the assessment scores for all observers. Further, X_i and X_j represent the sums of each row and column, $X_j - X_i$ is the difference of X_j and X_i, and $X_{...}$ represents the total of each row or column.

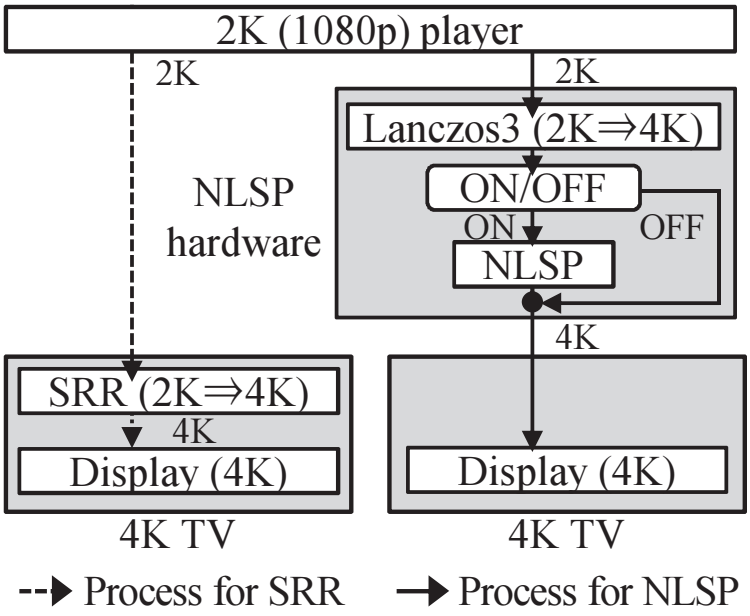


Figure 4.4: Experiment 1

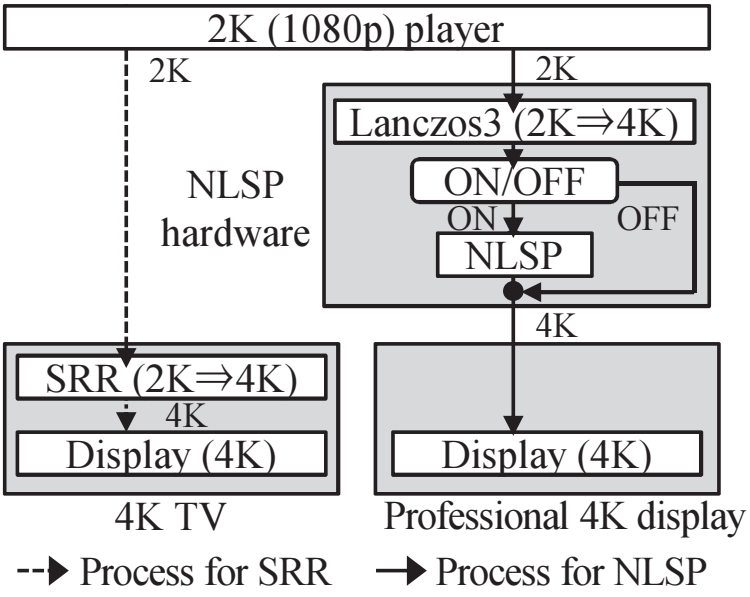


Figure 4.5: Experiment 2

Here, analysis of variance (ANOVA) was used to assess the significant differences in the assessment scores of the stimuli. The ANOVA results for the "Ruins" sequence are shown in Table 4.3. The sum of squares, degrees of freedom, and mean squares were calculated for each factor [67]. The F_0 score is a statistical value for the F-test, and it is obtained by dividing the mean square of a specific factor and that of the residual factor. Further, $F_{1\%}$ is a critical F value for the 1% significance level. If F_0 of the stimuli factor is greater than $F_{1\%}$, there is a

(1) Ruins

(2) Bricks

(3) Plaza

(4) Castle

(5) Cathedral

Figure 4.6: 2K test sequences

significant difference in the assessment scores of stimuli. Here, F_0 of the stimuli factor is $F_0 = 582.96 > F_{1\%} = 4.881$. Thus, a 1% significant difference between the stimuli is observed. All the ANOVA results for other test sequences are the same in that there are significant differences for the stimuli factor.

The meanings of the statistical differences in each factor are as bellow: The statistical difference in the "Stimuli" factor means the difference between the stimuli. The statistical difference in the "Combination" factor means the difference between specific combination of the stimuli. The statistical difference in the "Order" factor means the difference between the order of the stimuli with the reference and the target. The statistical differences in the "Stimuli/Order × Observers" factors mean the differences of "Stimuli"/"Order" factors among individuals.

Table 4.2: Cross table (Experiment 1 Ruins)

i \ j	Lanczos	NLSP	SRR	X_i
Lanczos	–	55	8	63
NLSP	-42	–	-39	-81
SRR	9	56	–	65
X_j	-33	111	-31	$X_{\cdots}$
$X_j - X_i$	-96	192	-96	47

Table 4.3: Analysis of variance (Experiment 1 Ruins)

Factor	Sum of squares	Degree of freedom	Mean square	F_0	$F_{1\%}$
Stimuli	307.20	2	153.60	582.96**	4.881
Stimuli×Observers	39.47	58	0.68	2.58**	1.746
Combination	0.05	1	0.05	0.19	6.963
Order	12.27	1	12.27	46.58**	6.963
Order×Observers	8.56	29	0.30	1.12	1.944
Residual	23.45	89	0.26	–	–
Overall result	391.00	180	2.17	–	–

**: 1% significant difference ($F_0 > F_{1\%}$)

The significant differences in each pair of stimuli were assessed because the ANOVA results guarantee the significant differences of least one of the pairs of stimuli. The yardstick values α for each stimulus are calculated by $(X_j - X_i)/(2Nn)$, where n is the number of observers (30) and N is the number of stimuli (3). The yardstick values for the "Ruins" sequence are shown in Figure 4.7. In Figure 4.7, the horizontal axis is the yardstick value, and the marks (rhombus, square, and triangle) show the values of each stimulus. Higher values indicate higher assessment. The values on the arrows show the differences between the stimuli. A critical value of the difference in yardstick values with significance level a is calculated as follows:

$$Y_a = q\sqrt{\frac{V_\varepsilon}{2nN}} \tag{4.1}$$

where V_ϵ is the mean square of the residual factor (0.26), as shown in Table 4.3. Further, q is obtained from the Student's t-distribution with the degrees of freedom for the residual factor (89) and number of stimuli N (3). Let significance level be 0.01. Then $q = 4.282$, and thus, $Y_{0.01} = 0.164$. If the difference in yardstick values is greater than $Y_{0.01}$, there is a significant difference between the yardstick values. In the results of the "Ruins" sequence, the yardstick values in Figure 4.7 are the highest for NLSP, SRR, and Lanczos, in that order. The differences

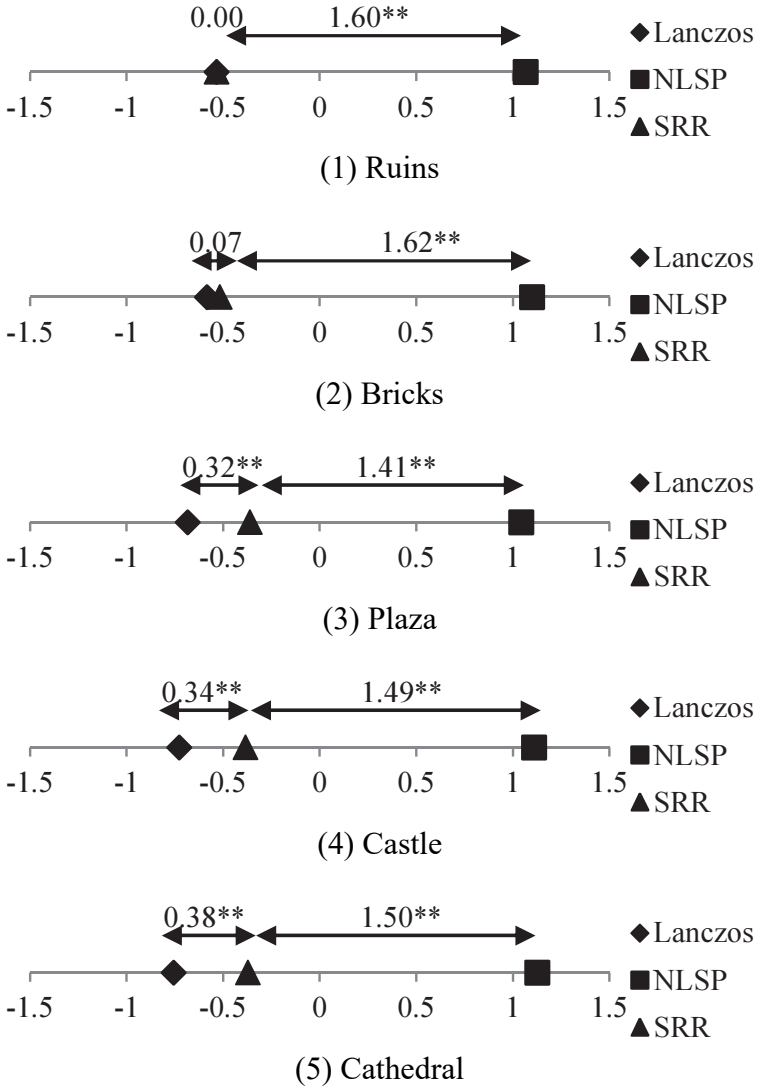


Figure 4.7: Assessment results (Experiment 1)

**: 1% significant difference

in the yardstick values of adjacent stimuli, NLSP with SRR ($\alpha_{NLSP} - \alpha_{SRR}$), and SRR with Lanczos ($\alpha_{SRR} - \alpha_{Lanczos}$) are as follows:

$$\alpha_{NLSP} - \alpha_{SRR} = 1.60 > Y_{0.01} \tag{4.2}$$

$$\alpha_{SRR} - \alpha_{Lanczos} = 0.00 < Y_{0.01} \tag{4.3}$$

Because $\alpha_{NLSP} - \alpha_{SRR}$ is greater than $Y_{0.01}$, a 1% significant difference between NLSP and SRR is observed. The value of $\alpha_{SRR} - \alpha_{Lanczos}$ is not greater than $Y_{0.01}$, and thus, a significant difference between SRR and Lanczos is not guaranteed. The asterisks (**) in Figure 4.7 indicate 1% significant differences between the stimuli. The yardstick values for other test sequences

Table 4.4: Cross table (Experiment 2 Ruins)

i \ j	Lanczos	NLSP	SRR	X_i
Lanczos	–	52	5	57
NLSP	-44	–	-38	-82
SRR	6	48	–	54
X_j	-38	100	-33	$X_{..}$
$X_j - X_i$	-95	182	-87	29

Table 4.5: Analysis of variance (Experiment 2 Ruins)

Factor	Sum of squares	Degree of freedom	Mean square	F_0	$F_{1\%}$
Stimuli	276.21	2	138.11	652.83**	4.881
Stimuli×Observers	52.12	58	0.90	4.25**	1.746
Combination	0.67	1	0.67	3.18	6.963
Order	4.67	1	4.67	22.09**	6.963
Order×Observers	10.49	29	0.36	1.71	1.944
Residual	18.83	89	0.21	–	–
Overall result	363.00	180	2.02	–	–

**: 1% significant difference ($F_0 > F_{1\%}$)

are shown in Figures 4.7 (2)-(5). All results have similar tendencies; NLSP has the highest evaluation, and there are significant differences between NLSP and SRR as well as NLSP and Lanczos in all cases. Significant differences between SRR and Lanczos are obtained for “Plaza”, shown in Figure 4.7 (3), “Castle”, shown in Figure 4.7 (4), and “Cathedral” shown in Figure 4.7 (5).

4.3.2 Experiment 2

The results of experiment 2 were analyzed in the same way as those of experiment 1. The cross table for the “Ruins” sequence is shown in Table 4.4, and the ANOVA results for the “Ruins” sequence is shown in Table 4.5. As the ANOVA results, the F_0 for the stimuli factor is $F_0 = 652.3 > F_{1\%} = 4.881$. Thus, the 1% significant difference between the stimuli is observed. The 1% significant difference in the stimuli factor is detected in all test sequences. The yardstick values for the “Ruins” sequences are shown in Figure 4.8 (1). The yardstick values are high in the order of NLSP, SRR, and Lanczos. The difference in the yardstick values

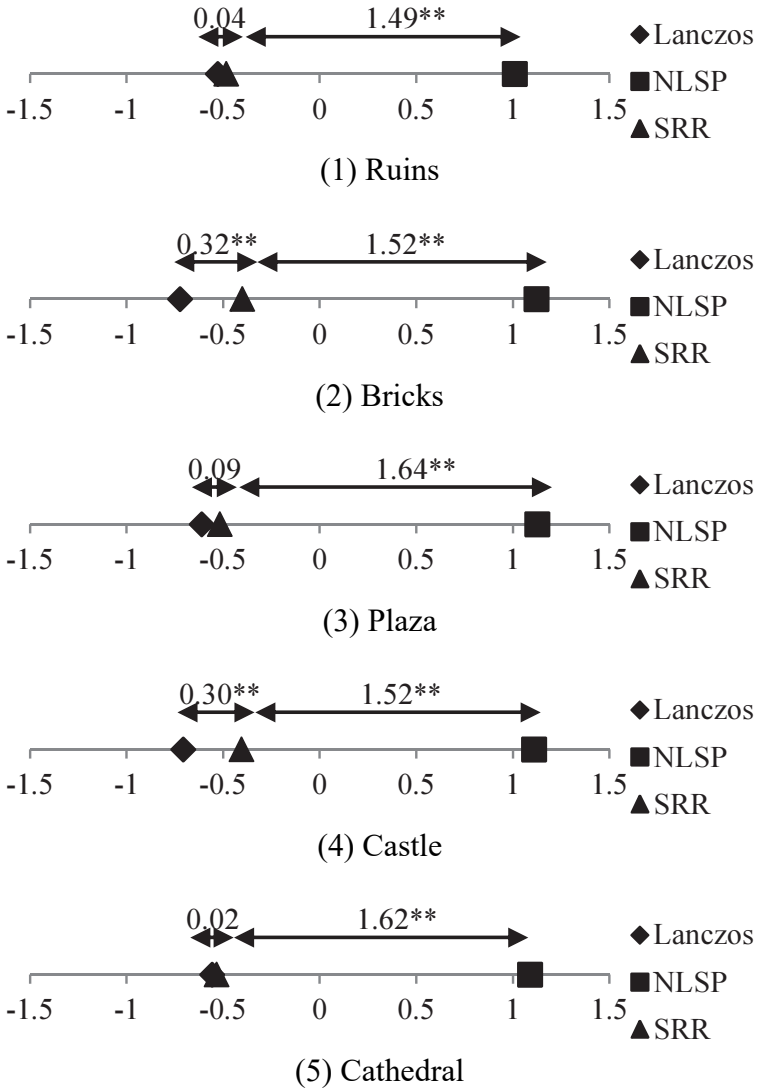


Figure 4.8: Assessment results (Experiment 2)

**: 1% significant difference

of adjacent stimuli, $\alpha_{NLSP} - \alpha_{SRR}$ and $\alpha_{SRR} - \alpha_{Lanczos}$ are as follows.

$$\alpha_{NLSP} - \alpha_{SRR} = 1.49 > Y_{0.01} \tag{4.4}$$

$$\alpha_{SRR} - \alpha_{Lanczos} = 0.04 < Y_{0.01} \tag{4.5}$$

Equitation 4.4 satisfies $(\alpha_{NLSP} - \alpha_{SRR}) > Y_{0.01}$, and Equiation 4.5 does not satisfy $(\alpha_{SRR} - \alpha_{Lanczos}) > Y_{0.01}$. There is the 1% significant difference between the NLSP and SRR. The significant difference between SRR and Lanczos is not observed. Figures 4.8 (2)-(5) show the results of yardstick values for each test sequence. All the results are similar to those of experiment 2: The yardstick values of NLSP are the highest of all stimuli, and the significant

differences are revealed between NLSP and the other two stimuli, SRR and Lanczos. Significant differences between SRR and Lanczos are obtained for “Bricks” shown in Figure 4.8 (2), and “Castle” shown in Figure 4.8 (4).

4.3.3 Discussions

As the results of experiments 1 and 2, the yardstick value of NLSP is highest. The significant differences of NLSP and other two methods, SRR, and Lanczos, are observed. The superiority of NLSP is proven from the results of two experiments with the same and different displays.

Contrarily, the quality differences between SRR and Lanczos are observed in only five results of the all ten results. The quality differences between SRR and Lanczos are too small to guarantee because they depend on the display and sequence. The essential limits of the ability of SRR to improve the resolution of the TV content were discussed in Chapter 2, and the results of the experiments are consistent with these discussions.

The yardstick values for each stimulus are asymmetrically distributed, and the statistical differences for the order factor are obtained in four cases with experiment 1, and in three cases with experiment 2. It is assumed the sensitivities how good/ bad for resolution with the Scheffe’s paired comparison are not the same: The results also show that the assessments of both, combination of stimuli with target and reference should be measured.

4.4 Conclusions

Subjective assessment experiments of 4K displays with NLSP and SRR were conducted to confirm practical performance of SRR on TV sets. As the results, the ability of SRR is inferior to NLSP in resolution quality, and the quality difference between SRR and the previous Lanczos method is obtained in some specific results only. The calculation cost for SRR is higher than NLSP or Lanczos, as it requires time-consuming iterations to reconstruct a high-resolution image. There is no particular advantage of applying SRR to TV sets.

Chapter 5

Subjective Assessment of 4K videos with super resolution with nonlinear signal processing

5.1 Introduction

4K TV sets are currently available on the market. They provide high quality images with quadruple the resolution of HD. However, although 4K TV sets are widely available, 4K video content is not; up-converted HD video content is still used with 4K TV sets.

Image enlargement is required to display low-resolution images on high-resolution displays. Note that enlarging of an image causes blurring. Thus, image resolution should be improved, and almost all consumer TV sets are equipped with a sharpness function (i.e., an enhancer or an unsharp mask). The sharpness function algorithm is simple and can work in real time; thus, it is widely used for video devices. However, the sharpness function can enhance edges but cannot actually improve resolution.

Super-resolution (SR) technology is a method for improving the resolution of images. Unlike the sharpness function, SR can reproduce high frequency spectra that the sharpness function cannot create.

While the capabilities of SR technologies for TV sets need to be improved, the evaluation of SR technology is also required. It is possible to theoretically analyze resolution by comparing SR and non-SR processed video signals. However, there is no way to take the video signals

after the SR signal process from the TV sets. Thus, a subjective assessment is the only way to evaluate the image qualities of SR technologies embedded in video devices.

Our team has proposed a new SR technology using nonlinear signal processing (NLSP), and its theoretical capability has been proven [7][8]. In addition, we have demonstrated improved subjective image quality in 4K videos up-converted from HD [59] as described in Chapter 4. NLSP can enhance images without enlargement. However, the high-resolution effect has not been tested. In this chapter, the high-resolution effect of NLSP is verified. We perform a subjective assessment wherein candidates are 4K videos with and without NLSP.

5.2 Experiment

5.2.1 Subjective Assessment Method

The assessment was conducted by comparing 4K videos with and without NLSP. Two consumer grade 4K TV sets of the same model were used. The same 4K TV sets used for the experiment in Chapter 4 (shown in Figure 4.1) were used. Observers watched the TVs and compared the image resolution qualities. This was intended to reproduce the conditions by which shoppers compare similar items at a store.

As with the experiments described in Chapter 4, Scheffe's paired comparison method was used for the assessment. A training session was conducted to explain resolution and the experimental method to observers. The observers assessed only resolution. The other evaluation factors, such as noise and color were not considered in the assessment.

5.2.2 Apparatus

A system diagram of the experiment is shown in Figure 5.1. A 4K video player was used to show video sequences to the observers, and a video signal was simultaneously distributed to each 4K TV set. In the process for original, the original video signal was output to the 4K TV set without any processing. In the process for NLSP, external NLSP hardware used in the experiment in Chapter 4 (shown in Figure 4.3 was connected between the video player and one of the 4K TV sets. Then, the output signal after NLSP was displayed on the 4K TV.

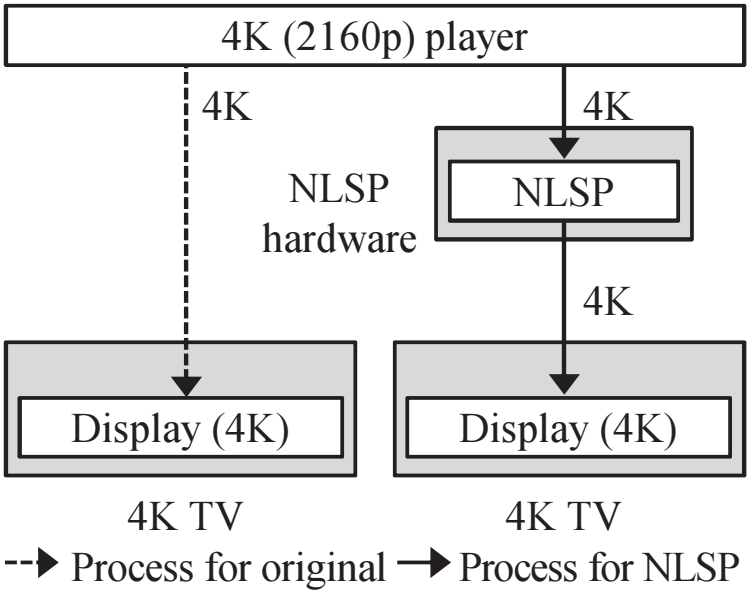


Figure 5.1: Experiment

5.2.3 Test Sequences

Test video sequences for the experiment were taken by a consumer video camera. The resolution is 4K (3,840 × 2,160), and the format is MPEG-4. Five test video sequences that are appropriate for the assessment of resolution were selected. Most did not included pan and tilt scenes. The length of each test sequence is between 10 s and 15 s in reference to BT.500. Figure 5.2 shows the test video sequences for the experiment. The ovals on the figures indicate high-resolution areas. The observers were asked to watch these areas and determine assessment scores.

5.2.4 Observers

Thirty observers participated in the experiment. The observers are non-experts, and they have normal visual acuity and color vision in reference to BT.500.

5.2.5 Experimental Conditions

Normal lighting conditions for a room were selected for the experiment to reproduce a viewing environment in which consumers choose a TV set at a store. The observers could freely change their viewing distance during the assessment.

(1) Cherry tree

(2) Bus

(3) Cruise ship 1

(4) Cruise ship 2

(5) Red brick

Figure 5.2: 4K test sequences

5.3 Results and Discussion

The assessment results for the “Cherry Tree” sequence are shown in Figure 5.3. Figure 5.3 shows a graph of the average scores and the standard deviations of each stimulus. Here, “NLSP” indicates stimulus of the process for NLSP, and “Original” indicates stimulus of the process without NLSP. The horizontal axis is the average scores. The average scores were obtained by dividing the total assessment scores by the number of observers. The bars extending from the marks are standard deviations. The graph shows that the average score for NLSP is 1.73 and that of original is -1.27. Note that the average score for NLSP is higher than that of original. Other assessment results are shown in Figures 5.4, 5.5, 5.6, and 5.7. All results demonstrate

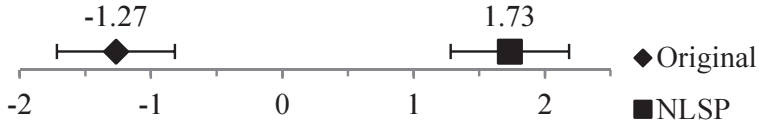


Figure 5.3: Assessment result (Cherry tree)

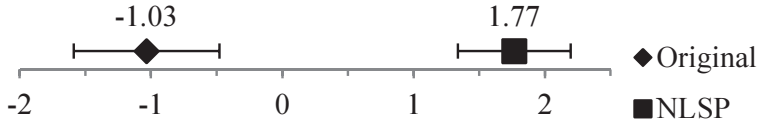


Figure 5.4: Assessment result (Bus)

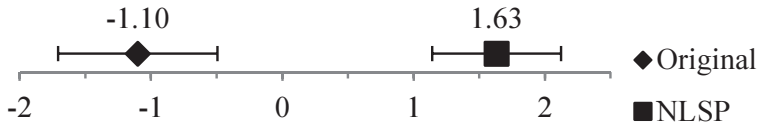


Figure 5.5: Assessment result (Cruise ship 1)

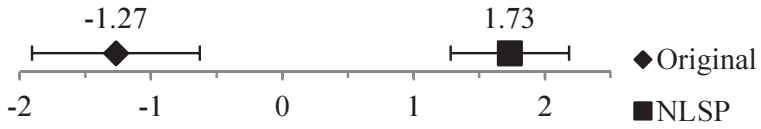


Figure 5.6: Assessment result (Cruise ship 2)

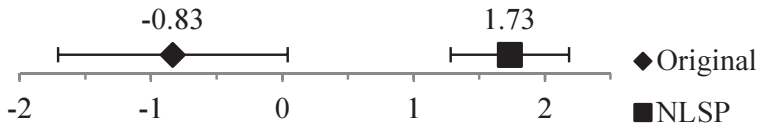


Figure 5.7: Assessment result (Red brick)

similar tendencies.

Analysis of variance was used to assess significant differences. The results of the analysis of variance for the “Cherry Tree” sequence are shown in Table 5.1. Table 5.1 shows the sum of squares, degrees of freedom, and mean square values [67]. F_0 denotes a value for an F-test. F_0 is obtained by the quotient of the mean square of the stimuli and that of the residual. Here, a critical F value for the 0.01 significance level is $F_{1\%} = 7.093$. If F_0 is greater than $F_{1\%}$, then the null hypothesis was rejected and there is significant difference between the stimuli. F_0 for the “Cherry Tree” sequence is $F_0 = 667.330 > F_{1\%}$. Similarly, other analysis of variance results for other sequences are shown in Tables 5.2, 5.3, 5.4, and 5.5.

All results satisfy $F_0 > F_{1\%}$, and significant differences between the stimuli are observed. As

Table 5.1: Analysis of variance (Cherry tree)

Factor	Sum of squares	Degree of freedom	Mean square	F_0
Stimuli	135.000	1	135.000	667.330**
Residual	11.733	58	0.202	–
Overall result	146.733	59	–	–

Table 5.2: Analysis of variance (Bus)

Factor	Sum of squares	Degree of freedom	Mean square	F_0
Stimuli	117.600	1	117.600	475.870**
Residual	14.333	58	0.247	–
Overall result	131.933	59	–	–

Table 5.3: Analysis of variance (Cruise ship 1)

Factor	Sum of squares	Degree of freedom	Mean square	F_0
Stimuli	112.067	1	112.067	367.917**
Residual	17.667	58	0.305	–
Overall result	129.733	59	–	–

Table 5.4: Analysis of variance (Cruise ship 2)

Factor	Sum of squares	Degree of freedom	Mean square	F_0
Stimuli	135.000	1	135.000	441.541**
Residual	17.733	58	0.306	–
Overall result	152.733	59	–	–

Table 5.5: Analysis of variance (Red brick)

Factor	Sum of squares	Degree of freedom	Mean square	F_0
Stimuli	98.817	1	98.817	204.448**
Residual	28.033	58	0.483	–
Overall result	126.850	59	–	–

**: 1% significant difference ($F_0 > F_{1\%}$), $F_{1\%} = 7.093$

a result, NLSP stimuli have higher average scores than original. In addition, the 0.01 significant differences were detected in all test sequences. Thus, it is statistically proven that video signals with NLSP obviously differ from the original video signals, and the image resolution quality of NLSP is superior. It is assumed that similar results would be obtained for other video sequences because the experimental method is reproducible. These results prove that NLSP can effectively enhance images.

5.4 Conclusion

A subjective assessment was performed to test the high-resolution effect of NLSP. The assessment was conducted by comparing NLSP and non-NLSP processed 4K videos on 4K TV sets. The results statistically proved that NLSP can effectively enhance images. A conventional SR function is performed for only enlarged images, whereas the proposed method is useful for enhancing images without enlargement. In this study, we have only assessed the resolution. In the future, we intend to assess other factors, such as noise and color.

Chapter 6

Subjective Assessment of Smartphone Displays with Super Resolution using Best-worst Method

6.1 Introduction

The introduction of the iPhone in 2008 increased the popularity of smartphones and revolutionized the way people communicate. Smartphones contain high-speed processors and large storage spaces, allowing for a high-performance of smartphones compared with that of the conventional mobile phones. Current smartphones are multifunctional devices that can be used as replacements for other digital devices, such as cameras, audio displays, games, and television sets (TV). The smartphone market has expanded beyond mobile phones, and numerous content services for smartphones are provided. Videos are a highly popular form of smartphone content, and there is a significant demand for the use of smartphones as video devices.

A TV set is a conventional video device for watching video content including TV programs. Image/ video quality is an important feature in video devices, and TV manufacturers have been attempting to improve image quality in TV sets to compete in the marketplace. Resolution is generally considered as an indicator of quality; hence, developing technologies for improving resolution is important. The resolution of common TV displays is the same as that of digital high-definition television (HDTV) broadcasting (1,920 $\times$ 1,080 pixels). In 2013, 4K TV sets with resolutions four times that of HDTV (3,840 $\times$ 2,160 pixels) were introduced, and in 2014,

4K broadcasting began tentatively in Japan.

When early smartphones were introduced, their major content was still images, owing to the limitations of mobile networks. In the past several years, the band limitations of mobile networks have expanded and watching video content on smartphones has become popular. TV sets have been replaced by smartphones, and using smartphones with high-resolution displays has also become a trend similar to that in the TV market. The display resolution of older smartphones was the same as that of half-size video graphics array (HVGA) (480 × 340 pixels). They have been improving steadily and HDTV resolution is now commonly used. In 2015, one smartphone manufacturer announced a smartphone with a 4K display.

Despite the high-resolution displays, such as 4K, being adapted to smartphones, the content provided is almost entirely HDTV. Furthermore, playing HDTV content on a 4K resolution display does not improve resolution, theoretically, as the content is merely enlarged to fit to the display, resulting in an interpolated image. Therefore, alternative ways of improving content resolution are required.

Signal processing is another way of improving resolution. This technology has advanced significantly over the last 20 years and resolution improvement technologies namely super resolution (SR), are used in some practical applications. Many SR methods have been proposed untill now [2][8][11], and some TV sets have also been equipped with these technologies.

In our previous works, an SR using nonlinear signal processing (NLSP) was proposed, and its image quality for 4K TV sets was also been proven [8][59][68]. We applied NLSP to a smartphone. However, the capability of SR on smartphone displays, which are smaller than TV sets, has not been proven. Display sizes of the smartphones sold in 2014 are about 4.5-5.5 inches and it means a tenth of conventional TV sets. There is no evidence that SR can improve the image quality on smartphone displays. Thus, the evaluation of NLSP on smartphone displays is necessary and we compare the image qualities of the smartphones with and without NLSP, and existing manufacturers' smartphones.

While the evaluation of the capability of SR is required, there is no standard method for the evaluation. The signal analysis in the frequency domain is a way to analyze the image resolution. However, the SR processed signal is only sent to the display of the devices, and there is no way to take the video signals from the smartphone devices. Since the signal analysis cannot be used, a subjective assessment is the only way to evaluate the performance of SR embedded in the video devices.

Subjective assessments are human psychophysical experiments. Numerous subjective assessment methods were reported in the areas of psychology and psychophysics, and they were applied to various purposes [11][12][13][62][69][70]. As described in Chapter 3, Chapter 4, and Chapter 5, we applied Scheffe's paired comparison [62] to assess an image quality of SR, and proved the usefulness of the assessment [59][68]. However, the assessment using paired comparison has an issue that becomes extremely time-consuming and fatiguing to the assessors in proportion to the number of the stimuli. In other method, a ranking method is the quicker assessment method [69] and this method was applied to an image quality assessment of CR images [71]. Nevertheless, the ranking method is inclined to cause dispersion of the assessors' opinions.

To overcome these problems, we anew introduce another assessment method, best-worst method [70]. This method can contribute a simple, quick procedure and high accuracy data collection. In this chapter, the assessment method is described and we also verify the usefulness of this method, including the image qualities of the smartphone displays.

6.2 Super Resolution with Nonlinear Signal Processing

Various SR algorithms have been proposed and some of them are equipped on TV sets.Most are very complex and require special hardware devices to work in real time. Application of them to smartphones is impossible because SR has to work over software.

Most SR methods are based on linear processing theories and pattern matching methods. Our SR method uses non-linear signal processing [8][9]. Figure 6.1 shows the signal flow of NLSP. The process is separated from the input image. First process including the high pass filter, the cubic function, and the limiter, creates the high frequency components. The high pass filter detects edges in the input image. The cubic function is one of the non-linear functions, and outputs $y = x^3$. This function generates the third harmonic waves of the edges detected by the high pass filter. The waves are generated only from the detected edges, and have high frequency components the input image does not have. The output of the cubic function becomes very large. If the input image has 8 bit depth, the range of the output of the high pass filter is from -255 to 255, and thus that of the cubic function is from $(-255)^3$ to $(255)^3$. The limiter fits the waves to the image. In the second process, the output of the limiter and the input image are added by the adder. The output image has high frequency components not existing in the

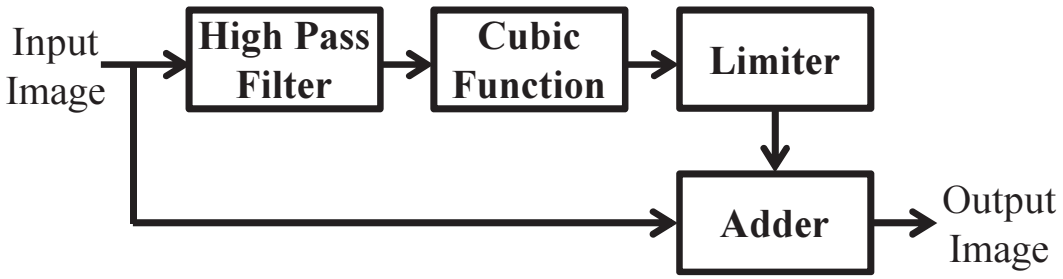


Figure 6.1: NLSP

Table 6.1: Stimuli

Smartphones	Manufacturers	NLSP
A	*a*	ON
B	*a*	OFF
C	*b*	-
D	*c*	-
E	*d*	-

original input image. The basic algorithm of NLSP is the same as Enhancer except the cubic function. This algorithm can improve the resolution Enhancer cannot, and can work in real time.

6.3 Subjective Assessment Method

The five smartphones (A, B, C, D, and E) shown in Figure 6.2 were used as stimuli for the assessment. The four smartphone models sold by different manufacturers (*a*, *b*, *c*, and *d*) were selected, and NLSP was applied to the manufacturer *a*'s model. The manufacturers and the state of NLSP are described in Table 6.1. The smartphones A and B are the same models of the manufacturer *a*. The NLSP processed video can be output onto the display by the setting of SR processing for NLSP to be enabled (ON). If the setting of it is disabled (OFF), the unprocessed video signal is output. The smartphones C, D, and E are respectively different manufacturers' models. Display resolution of the five smartphones is full-HD (1,920 × 1,080). The assessment was conducted by comparing the image qualities when full-HD videos are displayed. The smartphones were respectively ranked according to the image quality in resolution by using best-worst method.

Figure 6.2: Smartphones

6.3.1 Best-worst Method

The procedure of best-worst method is a repetition of a choice of a pair of the best and worst objects from a set. The assessment process of best-worst is as below: An observer selects the best and worst (highest and lowest resolution) smartphones from the five smartphones. The best smartphone is of the 1st rank, and the worst smartphone is of the 5th rank. The observer selects the best and worst smartphones in the same way from the remaining three smartphones. The best smartphone is of the 2nd rank, and the worst smartphone is of the 4th rank. Then the last remaining smartphone is of the 3rd rank.

6.4 Experiment

6.4.1 Apparatus

The smartphones for the assessment were flagship models that each manufacturer sold in 2014. They have 5.1- 5.5 inch displays that the resolution is full-HD (1,920 × 1,080). The same video sequence was repeated between the smartphones, and the observers compare the image qualities.

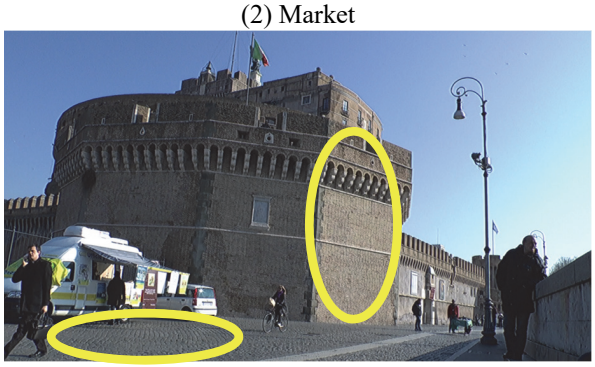
6.4.2 Test Sequences

We selected five test sequences for the assessment[36]. The length of each test sequence is 10-15 second. The test sequences were taken by a consumer digital video camera with MPEG-4 format and HDTV (1080p) resolution, and they have very high frequency elements without camera

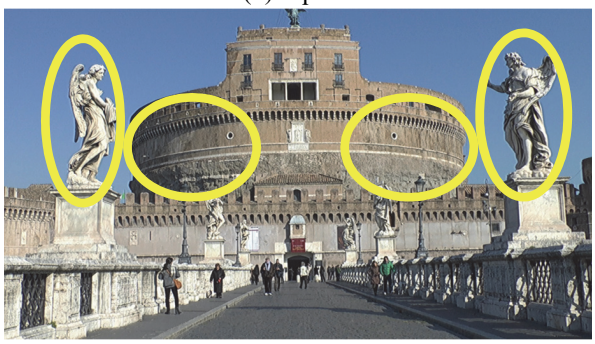
(1) Basilica

(2) Market

(3) Square



(4) Castle 1



(5) Castle 2

Figure 6.3: Test sequences

works, such as tilting and panning. The test sequences are shown in Figure 6.3 (1)-(5). The ovals on Figure 6.3 (1)-(5) indicate high-resolution areas that are easy to assess the resolution. Test sequence “Basilica” (Figure 6.3 (1)) has details in a roof, a surface of a pillar, and characters of a wall. Test sequence “Market” (Figure 6.3 (2)) has details in a stone pavement, window frames, and a surface of a house wall. Test sequence “Square” (Figure 6.3 (3)) has details in a surface of a building, a fence, and a mark of the building. Test sequence “Castel 1” (Figure 6.3 (4)) has details in a surface of a brick building, and a stone pavement. Test sequence “Castel 2” (Figure 6.3 (5)) has details in two statues and a bilateral surface of a building. These areas were presented to the observers.

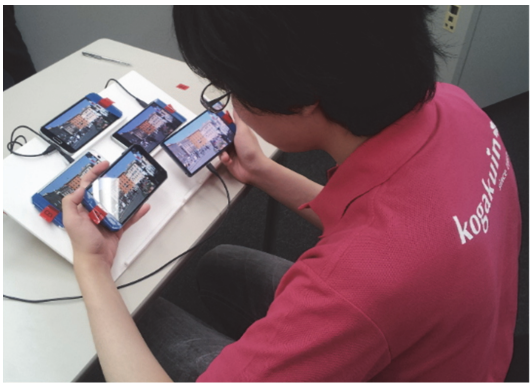


Figure 6.4: Best-worst method

6.4.3 Observers

23 observers participated in the experiment. They are non-video experts who do not work in the video industry and they have normal visual acuity and color vision [36].

6.4.4 Experimental Conditions

The experiment was conducted in a room with general lighting conditions. Prior to each assessment, a training session was held for the observers to understand the meaning of resolution. The experimental procedure and the evaluation points were properly provided to the observers using a dummy video sequence. Observers take smartphones in their hands and compare the image resolution. The assessment areas in the test sequences were not specified. The two-four highresolution areas of the ovals on Figure 6.3 (1)-(5) were presented to the observers as the examples of the assessment areas. The viewing distance during the assessment was not specified to reproduce the viewing environment because it depends on each user. The observers were asked to assess only resolution, including saturated-highlights. Saturated-highlights mean the loss of details that occurs in the white areas in the image with excessive luminance level. Other image quality factors, such as noise, luminance, and color, were not considered in the assessment. Figure 6.4 shows a photograph of the experiment.

6.5 Results

The statistical analysis process in this study is as follows. Using best-worst method, the ranks (1st-5th) were given to each stimulus (smartphones A, B, C, D, and E) as the evaluation value.

Table 6.2: Analysis result (Basilica)

	f_{kl}								
$l \setminus k$	A	B	C	D	E	r_l	P_l	ϵ_l	$K\epsilon_l$
1	19	1	3	0	0	5	90	0.1	1.28155
2	1	6	15	1	0	4	70	0.3	0.52440
3	3	12	3	4	1	3	50	0.5	0.00000
4	0	4	1	18	0	2	30	0.3	-0.52440
5	0	0	1	0	22	1	10	0.1	-1.28155
$\Sigma(f_{kl} \times K\epsilon_l)$	24.874	2.330	9.905	-8.915	-28.194				
R_k	1.081	0.101	0.431	-0.388	-1.226				
S_k^2	0.199	0.181	0.291	0.077	0.068				

(E.g. the smartphone A ranks 1st, the smartphone B ranks 3rd, the smartphone C ranks 2nd, the smartphone D ranks 4th, and the smartphone E ranks 5th.) The rank data were normalized by the distance scale since the average values of the rank do not have mathematical meanings. The normalized scores corresponding to each rank were given to the stimuli and the average scores for each stimulus were calculated [67]. The statistical differences between the average scores were evaluated by t-test.

The assessment result for "Basilica" sequence is shown as Table 6.2. The row l indicates the ranks (1- 5), and the column k indicates the stimuli (A-E). The values of f_{kl} are the numbers of the observers for the stimulus k as the rank l. (E.g. f_{C2} (15) means that 15 observers ranked the smartphone C as the 2nd.)

The normalized distances for each stimulus were calculated as follow. The column r_l in Table 6.2 represents the score corresponding to the rank l. The number of the stimuli (5) is represented as n. The column P_l indicates the mean value of the lth range, which is divided from the 0-100 range into n. The column $K_{\epsilon l}$ is the normalized quantile of P_l on the Gaussian distribution and it was used as the normalized score for the rank l. The normalized scores were given to each stimulus. The row $\Sigma(f_{kl} \times K_{\epsilon l})$ is the total of the normalized scores for the stimuli k. The row R_k means the average scores and the row S_k^2 means the variances for the stimuli k. Each average score represents the scale value of the image quality in resolution.

Figure 6.5 shows the average scores for "Basilica" sequence. The horizontal axis indicates the average score. The marks on the axis (rhombus, square, triangle, x-mark, and asterisk) show the average scores of the stimuli respectively (smartphone A, smartphone B, smartphone C, smartphone D, and smartphone E). The higher the averages indicate the higher the evaluation of resolution. The significance test is necessary because the differences in the average scores are

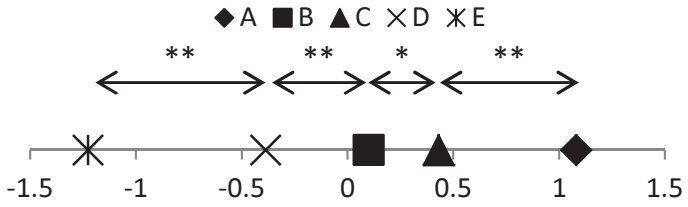


Figure 6.5: Assessment result (Basilica)

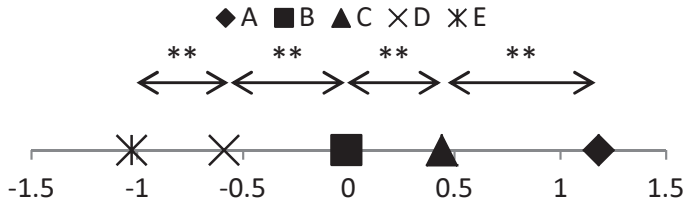


Figure 6.6: Assessment result (Market)

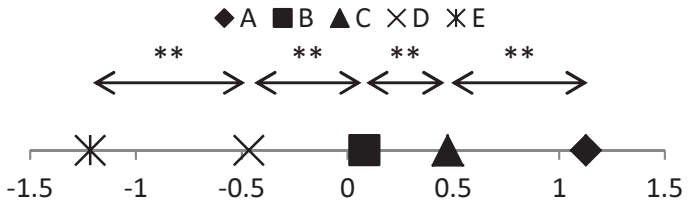


Figure 6.7: Assessment result (Square)

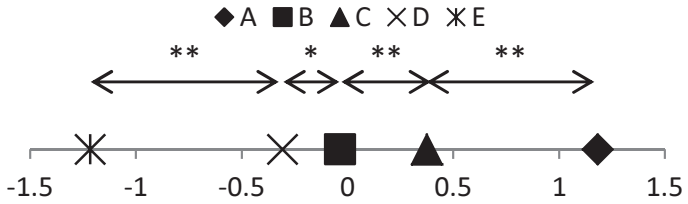


Figure 6.8: Assessment result (Castle 1)

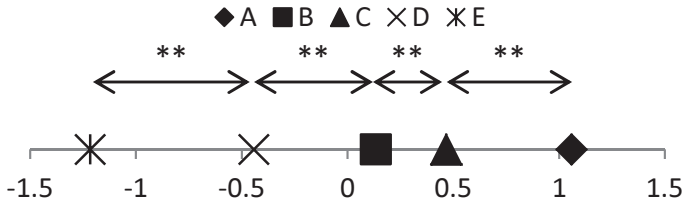


Figure 6.9: Assessment result (Castle 2)

** 1% statistical difference
* 5% statistical difference

not guaranteed. Here, t-test was used to evaluate the statistical differences between the average scores of the stimuli. A statistics quantity t_0 for t-test is as follows:

$$t_0(x,y) = \frac{(R_x - R_y)\sqrt{N-1}}{\sqrt{S_x^2 + S_y^2}} \tag{6.1}$$

where x and y are stimuli, where N is the number of observers (23). The degree of freedom of the distribution for the t_0 is $\phi = 2N - 2 = 44$. A critical t value for the 0.01 significance level with the degree of freedom ϕ is $t_{1\%} = 2.41413$, and that for the 0.05 significance level is $t_{5\%} = 1.68023$ [67]. If $t_0(x,y)$ is greater than the critical t values, the null hypothesis was rejected and there is the statistical difference between the average scores of the stimuli x and y.

In the result of "Basilica" sequence as shown Figure 6.5, the highest smartphone is A, and the second highest smartphone is C. The t_0 value for the smartphone A and the smartphone C ($t_0(A,C)$) is as follows:

$$t_0(A,C) = 4.35831 \ > \ t_{1\%} \tag{6.2}$$

Since equation 6.2 satisfies $t_0 > t_{1\%}$, the statistical difference with the 0.01 significance level between the smartphones A and C is obtained. In addition, the 3rd highest smartphone is B, the 4th highest smartphone is D, and the lowest smartphone is E. The t_0 values, for the 2nd highest smartphone and the 3rd highest smartphone ($t_0(C,B)$), the 3rd highest smartphone and the 4th highest smartphone ($t_0(B,D)$), and the 4th highest smartphone and the lowest smartphone ($t_0(D,E)$), are as follows:

$$t_{1\%} \ > \ t_0(C,B) = \ 2.24788 \ > \ t_{5\%} \tag{6.3}$$

$$t_0(B,D) = \ 4.51795 \ > \ t_{1\%} \tag{6.4}$$

$$t_0(D,E) = 10.31652 \ > \ t_{1\%} \tag{6.5}$$

Equation 6.3 does not satisfy $t_0 > t_{1\%}$, but satisfies $t_0 > t_{5\%}$. Thus, there is the statistical difference between the smartphones C and B with the 0.05 significance level. Equations 6.4 and 6.5 satisfy $t_0 > t_{1\%}$. Thus, the statistical differences between the smartphones B and D, and the smartphones D and E, are obtained with the 0.01 significance level. There are the statistical differences between all the stimuli. The asterisks ($**/*$) in Figure 6.5 represent the statistical difference between the stimuli indicated by the arrow. The 0.01 significance is shown as "$**$",

and the 0.05 significance is shown as "$*$". The results for the other test sequences are shown in figures 6.6-6.9. All results have the similar tendencies. The averages are high in the order of A, C, B, D, and E. The average score of the smartphone A equipped with NLSP is highest of the five smartphones for all cases.

6.6 Discussions

The average score of the smartphone A with NLSP is higher than the smartphone B, without NLSP, and the other manufacturers' smartphones. It implies NLSP can improve the image quality in resolution, and the image quality of the smartphone with NLSP is superior to the other manufacturers' smartphones. The assessment results prove the usefulness of the image quality of NLSP on smartphone displays. There are statistical differences between all the smartphones' image qualities. It is assumed that the differences in the image qualities were caused by technologies, such as image processing and display, which each manufacturer developed independently.

6.7 Conclusion

In this chapter, NLSP was applied to a smartphone, and the subjective assessment of smartphone displays was conducted. Best-worst method was used for the assessment. The results were statistically proven that the image quality of the smartphone with NLSP is superior in resolution to the other smartphones without NLSP.

The assessment method described in this chapter is applicable to other assessments. It is necessary to consider other factors for NLSP image quality, such as noise and color.

Chapter 7

Conclusions

In this study, the relationship between aliasing and super resolution image reconstruction was theoretically analyzed and the limitation of SRR was discussed. Aliasing in images is an important factor for SRR to reconstruct a high-resolution image. Aliasing decreases in accordance with fill factor, which is a parameter for imaging device. Current digital image and video content has almost no aliasing, and thus the application fields of SRR are extremely narrow. The results of subjective assessment experiments of a TV set with SRR were consistent with the discussions above. The experimental results show that the difference between SRR and conventional Lanczos method is obtained in some of the results only, and the performance is inferior to the other SR method. The limitation of SRR on TV sets is proved theoretically and subjectively. There has been no method for assessing displays equipped with signal processing including SR method. In this study, the assessment method for multiple displays was proposed and the reproducibility of the proposed method was proven through the experiments. The proposed assessment method is useful for product comparisons and it contributes to the satisfaction of consumers' demand for better quality of display products. The proposed assessment method is adaptable to assessments measuring other factors, such as noise, color, and texture; measurement of the overall total performance should be the focus in the future work.

Bibliography

[1] W. Burger, M. J. Burge, Principles of Digital Image Processing, Fundamental Techniques, Springer Science & Business Media, pp.150—155, 2010.

[2] S. Farsiu, M. Dirk Robinson, “Fast and Robust Multi-frame Super-resolution”, IEEE Trans. on Image Processing, Vol.13, No.10, pp.1327-1344 (2004).

[3] S. C. Park, M. K. Park, M. G. Kang, “Super-Resolution Image Reconstruction: A Technical Overview”, IEEE Signal Processing Magazine, Vol.20, No.3, pp.21-36 (2003).

[4] S. Chaudhuri, Super-Resolution Imaging, Kluwer Academic Publishers (2001).

[5] V. Bannore, Iterative-Interpolation Super-Resolution Reconstruction, Springer (2009).

[6] A. K. Katsaggelos, R. Molina, J. Mateos, Super Resolution of Images and Video, Morgan & Claypool (2007).

[7] S. Gohshi, “A New Signal Processing Method for Video: Reproduce The Frequency Spectrum Exceeding The Nyquist Frequency”, Proc. of The 3rd Annual ACM SIGMM Conference on Multimedia Systems (MMSys), pp.47-52 (2012).

[8] S. Gohshi, “Real Time Super Resolution for 4K/8K with Non-linear Signal Processing,” Journal of SMPTE, 124/7, pp. 51-56, Oct. 2014.

[9] S. Gohshi, “Real-time Super Resolution Equipment for 8K Video,” ICETE 2014 (SIGMAP), pp.149-157 Aug. 2014.

[10] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, “Learning Low-Level Vision”, International Journal of Computer Vision, 40(1), pp. 25-47, 2000.

[11] Freeman, William T., Thouis R. Jones, and Egon C. Pasztor, “Example-based super-resolution”, Computer Graphics and Applications, IEEE Vol. 22, No. 2, pp. 56-65, 2002.

[12] H. Chang, D.-Y. Yeung, Y. Xiong, “Super-resolution through neighbor embedding”, IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 1, pp. I-275-I-282 (2004).

[13] C.B. Atkins, C.A. Bouman, and J.P. Allebach, “Treebased resolution synthesis”, In Proceedings of the Conference on Image Processing, Image Quality and Image Capture Systems, pages 405—410, Savannah, GA, USA, 25—28 April 1999.

[14] R. Timofte, V. D. Smet, and L. V. Gool, “Anchored neighborhood regression for fast example-based superresolution”, IEEE International Conference on Computer Vision, pp.1920-1977, 2004.

[15] R.R. Schultz, R.L. Stevenson, “Extraction of high-resolution frames from video sequences,” IEEE Trans. on Image Processing, Vol.5, No.6, pp.996-1011, 1996.

[16] H.Stark, P.Oskoui, “High-resolution image recovery from image-plane arrays, using convex projections,” Journal of the Optical society of America, A, Optics and image science, Vol.6, No.11, pp.1715-1726, 1989.

[17] B.C. Tom, A.K. Katsaggelos, “ Reconstruction of a high-resolution image by simultaneous registration, restoration, and interpolation of low-resolution images,” Proceedings of International Conference on Image Processing, Vol.2, pp.539-542, 1995.

[18] D. Keren, S. Peleg, R. Brada, “Image sequence enhancement using sub-pixel displacements,” Proceedings CVPR ’88., Computer Society Conference on Computer Vision and Pattern Recognition, pp.742-746, 1988.

[19] S. Gohshi, “Limitation of Super Resolution Image Reconstruction for Video”, Fifth International Conference on Computational Intelligence, Communication Systems and Networks (CICSyN 2013), pp.217-221 (2013).

[20] S. Gohshi, “Limitation of Super Resolution Image Reconstruction”, Journal of the Institute of Image Information and Television Engineers, Vol.68, No.4, pp.169-173 (2014). (in Japanese)

[21] S. Baker, T. Kanade, “Limits on Super-Resolution and How to Break Them”, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.24, No.9, pp.1167-1183 (2002).

[22] Z. Lin, H. Y. Shum, “Fundamental Limits of Reconstruction-Based Superresolution Algorithms under Local Translation”, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.26, No.1, pp.83-97 (2004).

[23] S. S. Panda, M. S. R. S. Prasad, G. Jena, “POCS Based Super-Resolution Image Reconstruction Using an Adaptive Regularization Parameter”, International Journal of Computer Science Issues, Vol.8, Issue 5, No.2, ISSN (Online):1694-0814 (2011).

[24] A. W. M. Eekeren, K. Schutte, L. J. Vliet, “Multiframe Super-Resolution Reconstruction of Small Moving Objects”, IEEE Trans. on Image Processing, Vol.19, No.11, pp.2901-2912 (2010).

[25] N. Matsumoto, T. Ida, “Reconstruction-Based Super-Resolution Using Self-Congruency around Image Edges”, IEICE trans. on Information and Systems, Vol.J93, No.2, pp.118-126 (2010). (in Japanese)

[26] T. Gotoh, M. Okutomi, “High Resolution Color Image Reconstruction Using Raw Data of a Single Imaging Chip”, IPSJ Trans. on Computer Vision and Image Media, Vol.45, No.SIG 8, pp.15-25 (2004). (in Japanese)

[27] A. G. Devi, T. Madhu, K. L. Kishore, “An Improved Super Resolution Image Reconstruction using SVD based Fusion and Blind Deconvolution techniques”, International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol.7, No.1, pp.283-298 (2014).

[28] J. Zhu, C. Zhou, D. Fan, J. Zhou, “A New Method for Superresolution Image Reconstruction Based on Surveying Adjustment”, Journal of Nanomaterials, Vol.2014 (2014).

[29] M. Tanaka, M. Okutomi, “Theoretical Analysis about Limitations on Reconstruction-Based Super-Resolution”, IPSJ Trans. on Computer Vision and Image Media, Vol.47, No.SIG 5, pp.80-89 (2006). (in Japanese)

[30] D. Salomon, Data Compression, The Complete Reference, Springer, pp.240-242 (2012).

[31] B. Girod: “What’s Wrong with Mean-squared Error?”, in Digital Images and Human Vision, MIT Press Cambridge, pp.207-220 (1993).

[32] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, “Image quality assessment: from error visibility to structural similarity”, IEEE Trans. IP, vol. 13, pp. 600-612, 2004.

[33] L. Zhang, L. Zhang, X. Mou, and D. Zhang, “FSIM: a feature similarity index for image quality assessment”, IEEE Trans.IP, vol. 20, pp. 2378-2386, 2011.

[34] S. Miyaji, T. Hamada, S. Matsumoto, “Development of Digital Compressed Picture Quality Assessment System Considering Human Visual Perception”, The transactions of the Institute of Electronics, Information and Communication Engineers, Vol. J81-D-2, No. 6, pp.1084-1094, 1998. (in Japanese)

[35] Y. Nishida, S. Gohshi, Y. Itoh, H. Ohnishi, Y. Utsumi, K. Asai, H. Nishikawa, S. Kuroda, Y. Yasuda, “Real-time System for Assessing Picture Quality”, The Institute of Image Information and Television Engineers, Vol.54, No.11, pp.1623-1630, 2000. (in Japanese)

[36] Rec. ITU-R BT.500-11, “Methodology for The Subjective Assessment of The Quality of Television Pictures”, ITU-R (2002).

[37] K. Aoki, O. Mizuno, K. Kanda, E. Nakasu, K. Kubota, “A Study of Performance for MPEG-2 MP@ML and SP@ML : Correlation between Bit Rate and Picture Quality”, ITEJ Technical Report, Vol.19, No. 26, pp. 13-18, 1995. (in Japanese)

[38] E. Nakasu, K. Aoki,Y. Nishida, K. Kubota, “A Statistical Analysis of Picture Quality for Digital Broadcasting”, ITEJ Technical Report, Vol. 19, No.26, pp. 19-24, 1995. (in Japanese)

[39] Rec. ITU-T P.910-0,“Subjective video quality assessment methods for multimedia applications”, ITU-T, 2008.

[40] Rec. ITU-T P.912, “Subjective video quality assessment methods for recognition tasks”, ITU-T, 2008.

[41] Rec. ITU-R BT.1788-0, “Methodology for the subjective assessment of video quality in multimedia applications”, ITU-R, 2007.

[42] Hiroki Shoji, Seiichi Gohshi, “Performance of Learning-Based Super-Resolution on 4K-TV”, The 47th ISCIE International Symposium on Stochastic Systems Theory and Its Applications (SSS’15), pp.79-80, Dec.2015.

[43] Hiroki Shoji, Seiichi Gohshi, “ Subjective Assessment of Super Resolution for Remastering on 4K TVs”, Proc. of the Eleventh International Multi-Conference on Computing in the Global Information Technology (ICCGI) 2016, pp.10-15, November, 2016.

[44] S. Gohshi, S. Inoue, I. Masuda, T. Ichinose, Y. Tatsumi, “Super Resolution for Smartphones”, Proc. of the 13th International Conference on e-Business and Telecommunications - Volume 5: SIGMAP, pp. 106-112, 2016.

[45] S. Gohshi, T. Hiroi, I. Echizen, “Subjective Assessment of HDTV with Super Resolution Function”, EURASIP Journal on Image and Video Processing 2014 (2014).

[46] S. Gohshi, I. Echizen, “Limitations of Super Resolution Image Reconstruction and How to Overcome Them for a Single Image”, International Conference on Signal Processing and Multimedia Applications (SIGMAP 2013), pp.71-78 (2013).

[47] D. Glasner, S. Bagon, M. Irani, “Super-Resolution from a Single Image”, IEEE 12th International Conference on Computer Vision (ICCV 2009), pp.349-356 (2009).

[48] Visual Information Research Institute, http://www.viri.osakac.ac.jp/meeting07/viri07_6.pdf. (in Japanese)

[49] M. Sugie, S. Gohshi, “Performance Verification of Super-Resolution Image Reconstruction”, International Symposium on Intelligent Signal Processing and Communications Systems (ISPACS 2013), pp.547-552 (2013).

[50] T. Fukinuki, “Conditions of Super Resolution: the Roles of Solid-State lmage Sensors”, ITE Technical Report, Vol.34, No.16, pp.51-54 (2010). (in Japanese)

[51] K. Aizawa, T. Hamamoto, CMOS Image Censor, Corona Publishing, pp.26-33 (2012). (in Japanese)

[52] S. Yoshihara, Y. Nitta, M. Kikuchi, K. Koseki, Y. Ito, Y. Inada, S. Kuramochi, H. Wakabayashi, M. Okano, H. Kuriyama, J. Inutsuka, A. Tajima, T. Nakajima, Y. Kudoh, F. Koga, Y. Kasagi, S. Watanabe, T. Nomoto, “A 1/1.8-inch 6.4 MPixel 60 frames/s CMOS Image Sensor With Seamless Mode Change”, IEEE Journal of Solid-State Circuits, Vol.41, No.12, pp.2998-3006 (2006).

[53] Y. Nitta, Y. Muramatsu, K. Amano, T. Toyama, J. Yamamoto, K. Mishina, A. Suzuki, T. Taura, A. Kato, M. Kikuchi, Y. Yasui, H. Nomura, N. Fukushima, “High-Speed Digital Double Sampling with Analog CDS on Column Parallel ADC Architecture for Low-Noise Active Pixel Sensor”, IEEE International Solid State Circuits Conference -Digest of Technical Papers, pp.2024-2031 (2006).

[54] S. Iwabuchi, Y. Maruyama, Y. Ohgishi, M. Muramatsu, N. Karasawa, T. Hirayama, “A Back-Illuminated High-Sensitivity Small-Pixel Color CMOS Image Sensor with Flexible Layout of Metal Wiring”, 2006 IEEE International Solid State Circuits Conference - Digest of Technical Papers, pp.1171-1178 (2006).

[55] T. Arai, T. Hayashida, J. Yonai, H. Ohtake, “A 16.7 Mfps 312 kpixel Backside illuminated Ultrahigh-speed CCD Image Sensor”, NHK R& D, No.141, pp.30-41 (2013). (in Japanese)

[56] T. Arai, J. Yonai, T. Hayashida, H. Ohtake, H. V. Kuijk, T. G. Etoh, “A 16.7 Mfps, 312 kPixel Backside-Illuminated Ultrahigh-Speed CCD Image Sensor with a Sensitivity 12.7 Times Higher than That with Front-Side Illuminated Structure”, IEICE Trans. on Electronics, Vol.J96-C, No.7, pp.180-190 (2013). (in Japanese)

[57] Y. Matsuo, S. Iwasaki, Y. Yamamura, J. Katto, “Image Super-resolution from Digital Cinema to Ultra high Definition Video using Registration of Wavelet Multi-scale Components”, Forum on Information Technology (FIT 2012), Vol.11, No.3, pp.1-6 (2012). (in Japanese)

[58] T. Kuroda, “Fundamentals of Image Input Device Technologies (3): Introduction to CMOS Image Sensors”, Journal of the Institute of Image Information and Television Engineers, Vol.68, No.3, pp.216-222 (2014). (in Japanese)

[59] M. Sugie, S. Gohshi, H. Takeshita, C. Mori, “Subjective Assessment of Super-resolution 4K Video Using Paired Comparison”, Proc. of the International Symposium on Intelligent Signal Processing and Communications Systems (ISPACS 2014), pp.42-47 (2014).

[60] T. Kurita, Y. Sugiura, “Consideration on Motion Compensated Deinterlacing and its Converter”, IEICE technical report, Vol.17, No.81, pp.25-32 (1993). (in Japanese)

[61] C. Mori, K. Tanioka, S. Gohshi, “Relationship between Super Resolution Image Reconstruction and Image Device”, IIEEJ Transactions on Image Electronics and Visual Computing, Vol.4, No.1, pp.12-19, 2016.

[62] H. Scheffe, “an Analysis of Variance for Paired Comparisons”, Journal of the American Statistical Association, Vol.47, No.259, pp.381-400 (1952).

[63] M. Nakamae, Y. Tabata, Y. Ohga, M. Kakuta, F. Uto, T. Okunishi, T. Ochi, K. Maeda, “Method of Subjective Evaluation by Scheffe’s Method of Paired Comparisons”, Japanese Journal of Radiological Technology, Vol.52, No.11, pp.1561-1565, 1996. (in Japanese)

[64] S. Kubota, “Evaluation of Image Quality of Organic Light-emitting Diode Displays”, The Journal of The Institute of Image Information and Television Engineers Vol. 62, No.1, pp.122-125, 2008. (in Japanese)

[65] L. L. Thurstone, “A Law of Comparative Judgment”, Psychological Review, Vol.34, No.4, pp.273-286 (1927).

[66] W. Burger, M. J. Burge, Principles of Digital Image Processing: Core Algorithms, Springer Science & Business Media, pp.223-225 (2010).

[67] Tadahiko Fukuda, Ryoko Fukuda, Ergonomics handbook, Scientist press co.ltd, Tokyo, 2009. (in Japanese)

[68] C. Mori, M. Sugie, H. Takeshita, S. Gohshi, “Subjective Assessment of Super-Resolution: High-Resolution Effect of Nonlinear Signal Processing”, Proc. of The 10th Asia-Pacific Symposium on Information and Telecommunication Technologies (APSITT 2015), pp.46-48 (2015).

[69] E. K. Strong, “Application of the “Order of Merit Method” to Advertising”, The Journal of Philosophy, Psychology and Scientific Methods 8, pp.600-606 (1911).

[70] Finn, Adam and Jordan J. Louviere, “Determining the appropriate response to evidence of public concern: The case of food safety,” Journal of Public Policy and Marketing, 11:1, pp.19-25, 1992.

[71] M. Nakamae, “Study of the Reliability of Visual Evaluation by the Ranking Method: Analysis of Ordinal Scale and Psychological Scaling Using the Normalized-rank Approach”, Japanese Journal of Radiological Technology 56.5, pp.725-730, 2000. (in Japanese)

Publications

Journals (with Review)

[1] **Chinatsu Mori**, Kenkichi Tanioka, Seiichi Gohshi, “Relationship between Super Resolution Image Reconstruction and Image Device”, IIEEJ Transactions on Image Electronics and Visual Computing, Vol.4, No.1, pp.12-19, June, 2016.

[2] **Chinatsu Mori**, Masaki Sugie, Hirohisa Takeshita, Seiichi Gohshi, “Method of Subjective Image Quality Assessment of Multiple Displays”, The Journal of the Institute of Image Electronics Engineers of Japan, Vol.46 No.2, April, 2017. (in Japanese)

International Conferences (with Review)

[1] **Chinatsu Mori**, Seiichi Gohshi, “Specifying a Person on the Basis of Body Characteristics with a Low-Resolution Image”, IIEEJ International Workshop on Image Electronics and Visual Computing (IEVC 2014), October, 2014

[2] Masaki Sugie, Seiichi Gohshi, Hirohisa Takeshita, **Chinatsu Mori**, “Subjective assessment of super-resolution 4K video using paired comparison”, Intelligent Signal Processing and Communication Systems (ISPACS 2014), December, 2014.

[3] **Chinatsu Mori**, Masaki Sugie, Hirohisa Takeshita, Seiichi Gohshi, “Subjective Assessment of Super-Resolution: High-Resolution Effect of Nonlinear Signal Processing”, 10th Asia-Pacific Symposium on Information and Telecommunication Technologies (APSITT 2015), August, 2015.

[4] **Chinatsu Mori**, Seiichi Gohshi, “Image Quality of a Smartphone Display with Super-Resolution”, The 47th ISCIE International Symposium on Stochastic Systems Theory

and Its Applications (SSS 2015), December, 2015.

[5] **Chinatsu Mori**, Kenkichi Tanioka, Seiichi Gohshi, “Super Resolution Image Reconstruction and Imaging Device”, 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC 2016), October, 2016.

[6] **Chinatsu Mori**, Seiichi Gohshi, “Subjective Assessment Method for Multiple Displays with and without Super Resolution”, 12th International Conference on Computer Vision Theory and Applications (VISAPP 2017), March, 2017.

Other Presentations

[1] **Chinatsu Mori**, Keiko Kubota, Seiichi Gohshi, “Quality assessment of 3D television”, Proceedings of the ITE annual convention (2013), “16-9-1”-“16-9-2”, August, 2013. (in Japanese)

[2] Masaki Sugie, Seiichi Gohshi, Hirohisa Takeshita, **Chinatsu Mori**, “Subjective assessment of Super-resolution -Paired comparison with 4K TV-”, IEICE Technical Report, Vol. 114, No. 233, IE2014-48, pp.51-56, October, 2014. (in Japanese)

[3] **Chinatsu Mori**, Masaki Sugie, Hirohisa Takeshita, Seiichi Gohshi, “Subjective Assessment of 4K Videos : Application of Super-Resolution with Nonlinear Signal Processing to 4K Displays”, IEICE Technical Report, Vol. 115, No. 96, IE2015-41, pp.51-55, June, 2015. (in Japanese)

[4] **Chinatsu Mori**, Aya Kubota, Seiichi Gohshi, “Subjective Assessment of Video Noise Reduction with Non-Recursive Temporal Filter”, IEICE technical report, Vol.117, No. 431, pp.71-76, February, 2018. (in Japanese)