Development of Real-time HDTV-to-8K TV Upconverter

Seiichi Gohshi[1], Shinichiro Nakamura[2] and Hiroyuki Tabata[2]
[1]*Kogakuin University, 1-24-2, Nishi-Shinjuku, Shinjuku-Ku, Tokyo, Japan*
[2]*Keisoku Giken Co., Ltd. 2-12-2 Chigasaki-minami, Tsuzuki-ku, Yokohama-city, Kanagawa-pref., Japan*

Keywords: 8KTV, 4KTV, HDTV, Up-convert, Super Resolution with Non-linear Processing, Super Resolution Image Reconstruction, Learning based Super Resolution, Non-linear Signal Processing.

Abstract: Recent reports show that 4K and 8K TV systems are expected to replace HDTV in the near future. 4K TV broadcasting has begun commercially and the same for 8K TV is projected to begin by 2018. However, the availability of content for 8K TV is still insufficient, a situation similar to that of HDTV in the 1990s. Upconverting analogue content to HDTV content was important to supplement the insufficient HDTV content. This upconverted content was also important for news coverage as HDTV equipment was heavy and bulky. The current situation for 4K and 8K TV is similar wherein covering news with 8K TV equipment is very difficult as this equipment is much heavier and bulkier than that required for HDTV in the 1990s. The HDTV content available currently is sufficient, and the equipment has also evolved to facilitate news coverage; therefore, an HDTV-to-8K TV upconverter can be a solution to the problems described above . However, upconversion from interlaced HDTV to 8K TV results in an enlargement of the images by a factor of 32, thus making the upconverted images very blurry. An upconverter with super resolution has been proposed in this study in order to fix this issue.

1 INTRODUCTION

Research about HDTV started in the 1960s, and its practical usage began in the late 1990s. The broadcasting service began in 2000 for digital satellite HDTV and in 2003 for terrestrial HDTV, and now both services are offered in multiple countries. More than 30 years of research were required for HDTV to become a practical service, and only 18 years have passed since these services began. However, 4K TV services have been made available via satellite broadcasting and Internet services. The horizontal and vertical resolutions of HDTV are 1,980 pixels and 1,080 pixels, respectively (ITU-R-HDTV, 2015), and that of 4K TV are 3,860 pixels and 2,160 pixels, respectively (ITU-R-UDTV, 2015). The 4K TV system has evolved over time and is used for multiple applications such as sports content, cinema and personal videos. 8K TV has a horizontal resolution of 7,680 pixels and a vertical resolution of 4,320 pixels (ITU-R-UDTV, 2015), which is four times greater than that of 4K TV and 16 times greater than that of full HDTV (progressive HDTV). Broadcasting HDTV adopts an interlaced video system which contains half the information as that contained by full HDTV. This means that 8K TV content has a resolution 32 times higher than that of the broadcasting HDTV content. The system clock frequencies for broadcasting HDTV, 4K TV and 8K TV are set at 74.25 MHz, 594 MHz and 2,376 MHz (2.376 GHz), respectively. The HD equipment used currently has evolved both in terms of technology and cost effectiveness, and a majority of the video content available, including films, is made in HD. Although the use of commercial 4K TV is practical, its equipment is not commonly available, especially that used professionally, for example, 4K TV professional video cameras. Sony began to release its professional studio cameras in 2014, which are still expensive. Other 4K TV equipment such as professional editing systems, transmission systems and outside broadcasting cars are both technically immature and expensive. All types of 8K TV equipment are currently under development or are being researched; therefore, its practical use is much more difficult to begin than that of 4K TV. However, 8K TV broadcasting is projected for 2018 and is expected to be a highlight of the 2020 Olympic Games. There are a couple of problems that 8K TV services are faced with. First, 8K TV content for broadcasting is crucial but rather insufficient. Second, using 8K TV equipment in news gathering systems such as outside broadcasting cars and helicopters is not currently practical because of

Gohshi S., Nakamura S. and Tabata H.
Development of Real-time HDTV-to-8K TV Upconverter.
DOI: 10.5220/0006116000520059
In *Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2017)*, pages 52-59
ISBN: 978-989-758-225-7

the reasons described earlier. The same situation existed for HDTV in the late 1990s, wherein creating HDTV content and gathering news was difficult because of the expensive and bulky equipment. In contrast, analogue TV content was sufficient and the professional equipment required was less expensive and of small size and low weight. Therefore, analogue TV content was upconverted to HDTV content to resolve the problems of insufficient HDTV content and expensive equipment. The HD equipment available currently is sufficiently small to be used for news gathering and outside broadcasting; however, analogue content is still used for HDTV broadcasting with analogue TV-to-HDTV upconversion as much excellent analogue content has been stored and accumulated over time. However, upconverted content is blurry because the images are interpolated. The highest resolution of the original image and the interpolated image remains the same despite using an ideal interpolation filter. The upconverted HDTV content can be immediately recognised as it appears blurry. The resolution ratio of HDTV to analogue TV is 5:1. The same issue will occur if HDTV content is used for upconversion to 8K TV. The accumulated HDTV content is interlaced as the professional equipment used for it is an interlaced system; hence, we need to upconvert this interlaced content to 8K TV content, including that for news gathering and outside broadcasting.

As discussed earlier, the resolution ratio of HDTV to 8K TV is 1:32, whereas that for analogue TV-to-HDTV conversion is 1:5. This shows that HDTV-to-8K TV conversion produces blurrier content than that produced by analogue TV-to-HDTV conversion. Currently the upconversion from the interlaced HDTV to progressive HDTV (full HD) is not so difficult and the full HD equipment such as cameras, recorders and other studio equipment are available. However, the upconversion from full HD to 8K is still 1:16 and it is higher magnifying scale than that of analogue TV to HDTV. Such blurry content does not take advantages of the high resolution screen, which is the most important sales point of 8K TV. Enhancers are generally used to improve the resolution of images and videos (Schreiber, 1970)(Lee, 1980)(Pratt, 2001). These enhancers use a simple algorithm to cope with real-time signal processing for videos and are provided in most digital HDTVs and 4K TVs. However, these enhancers cannot create high frequency elements as they only amplify the edges in an image. Therefore, it is necessary to develop a new technology which can cope with creating such elements that are not available for the current upconverted images.

2 SUPER RESOLUTION (SR)

Super Resolution (SR) is a technology that creates a high-resolution image from low-resolution ones (Park et al., 2003)(Farsiu et al., 2004) (van Eekeren et al., 2010) (Houa and Liu, 2011) (Protter et al., 2009) (Panda et al., 2011). The keyword phrase ”Super resolution” gets about 160 million hits on Google. Indeed, there are many SR proposals, but most of them are complex algorithms involving many iterations. If the iterations are conducted for video signals, frame memories, of the same number as the iterations, are required. Such algorithms are almost impossible to work with real-time hardware for the upconverted 8K content. Although non-iterative SR was proposed (Sanchez-Beato and Pajares, 2008), it only reduces aliasing artifact for a couple of images with B-Splines. It is not sufficient to improve HDTV-to-8K upconverted blurry videos because the upconverted videos do not have aliasing at all. SR for TV should have low delay. Especially in live news broadcasts, conversations between announcers in the TV studio and persons at the reporting point tend to be affected by delays. For viewers, the superimposed time is not accurate on a TV screen if the delay is longer than 60 seconds. For these reasons, complex SR algorithms with iterations cannot be used in TV systems. Although a real-time SR technology for HDTV was proposed (Toshiba, 2016)(Matsumoto and Ida, 2010), its resolution is worse than that of HDTV without SR (Gohshi et al., 2014).

SR with non-linear signal processing (NLSP) has been proposed as an alternative to the conventional image enhancement methods (Authors related), and it has several advantages compared with conventional SR technologies. Since it does not use iterations or frame memories, it is sufficiently lightweight to be installed in an FPGA (Field Programmable Gate Array) for real-time video processing. Furthermore, it can create frequency elements that are higher than those of the original image, as has been proven by performing two-dimensional fast Fourier transform (2D-FFT) results (Gohshi and Echizen, 2013) . However, it has not been used for 8K content because the system clock of 8K is 2.3376 GHz. In this paper, we present real-time HD/8K upconverter with NLSP to improve actual resolution of the content upconverted to 8K from HDTV.

3 SR WITH NON-LINEAR SIGNAL PROCESSING

The basic idea of NLSP is like that of the one-dimensional signal processing shown in Figure 1 (Gohshi and Echizen, 2013). The input is distributed to two blocks. The upper path creates high-frequency elements that the original image does not have as follows. The original image is processed with a high pass filter (HPF) to detect edges. The output of the HPF is edge information that has a sign, i.e., plus or minus, for each pixel. After the HPF, the edges are processed with a non-linear function (NLF). If an even function such as x^2 is used as the NLF, the sign information is lost. To stop this from happening, the most significant bit (MSB) is taken from the edge information before the NLF and restored after the NLF. Non-linear functions generate harmonics that can create frequency elements that are higher than those of the original image. NLSP using a number of non-linear functions should be able to create high-frequency elements. Here, we propose $y = x^2$ for plus edges and $y = -x^2$ for minus edges.

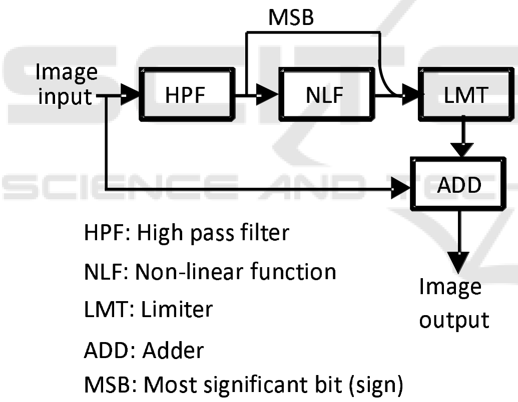


Figure 1: NLSP algorithm.

It is well known that images are expanded in a Fourier series (Mertz and Gray, 1934). Here, we take a one-dimensional image $f(x)$ to make the explanation simple. $f(x)$ is expanded as follows.

$$f(x) = \sum_{n=-N}^{+N} a_n \cos(n\omega_0) + b_n \sin(n\omega_0) \qquad (1)$$

ω_0 is the fundamental frequency and N means a positive integer. The HPF attenuates low-frequency elements including the zero frequency element (DC). We denote the output of the HPF by $g(x)$ and it becomes as follows.

$$\begin{aligned} g(x) = & \sum_{n=-N}^{-M} a_n \cos(n\omega_0) + b_n \sin(n\omega_0) \\ & + \sum_{n=M}^{N} a_n \cos(n\omega_0) + b_n \sin(n\omega_0) \end{aligned} \qquad (2)$$

M is also a positive integer and $N > M$. The frequency elements from $-M$ to M are eliminated with the HPF. DC has the largest energy in the images, and it sometimes causes saturation whereby the images become either all white or all black. The square function does not cause saturation by eliminating DC, and it has the following effect. Edges are represented with $\sin(n\omega_0)$ and $\cos(n\omega_0)$ functions. The square function generates $\sin^2(n\omega_0)$ and $\cos^2(n\omega_0)$ from $\sin(n\omega_0)$ and $\cos(n\omega_0)$. $\sin^2(n\omega_0)$ and $\cos^2(n\omega_0)$ generate $\sin 2(n\omega_0)$ and $\cos 2(n\omega_0)$. Theoretically it can be explained as follows. Since the most significant bit (MSB) of the $g(x)$ is protected, the input of the LMT for $g(x) > 0$ becomes the Equation 3 and that of the LMT for $g(x) < 0$ becomes the Equation 4.

$$\begin{aligned} (g(x))^2 = & \sum_{n=-2N}^{-M} c_n \cos(n\omega_0) + d_n \sin(n\omega_0) \\ & + \sum_{n=M}^{2N} c_n \cos(n\omega_0) + d_n \sin(n\omega_0) \end{aligned} \qquad (3)$$

$$\begin{aligned} -(g(x))^2 = & - \sum_{n=-2N}^{-M} c_n \cos(n\omega_0) + d_n \sin(n\omega_0) \\ & - \sum_{n=M}^{2N} c_n \cos(n\omega_0) + d_n \sin(n\omega_0) \end{aligned} \qquad (4)$$

Here, c_n and d_n are coefficients of the expansion of Equation 2. Although Equations 3 and 4 have the high frequency elements from $(N+1)\omega_0$) to $2N\omega_0$), they do not exist in the input image, Equation 1. Since these high frequency elements are created with the non-linear function, some of them are too large and need to be processed with LMT. After LMT processing, the created high frequency elements are added to the input with ADD. These NLFs create frequency elements that are two times higher than the input, and they can be used to double the size of the images horizontally and vertically, such as in the upconversion from HD to 4K.

It is necessary to apply NLSP horizontally and vertically, since images and videos are two-dimensional signals. Figure 2 is a block diagram of the real-time video processing. The input is distributed to two paths. The output of the upper line, the delay path, is the same as the input. The signal is

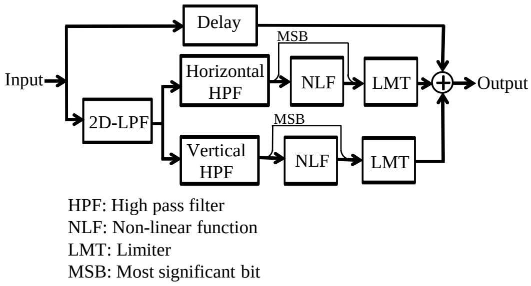


Figure 2: Block diagram of real-time hardware.

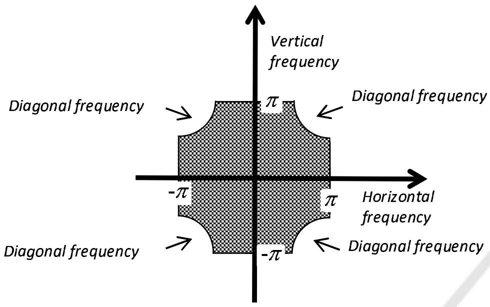


Figure 3: Characteristics of 2D-LPF.

delayed until the signal processing on the other paths ends. The bottom line includes a two-dimensional low pass filter (2D-LPF) and a parallel NLSP part. The 2D-LPF block decreases noise in video because noise has horizontal and vertical high frequency elements. Figure 3 shows the two-dimensional frequency characteristics of the 2D-LPF. 2D-LPF passes the checker marked area and eliminates the diagonal frequency elements, i.e., the four corners shown in Figure 3. NLSP creates horizontal high frequency elements and vertical high frequency elements. Both horizontal and vertical high frequency, diagonal, elements are processed with horizontal NLSP and vertical NLSP separately. If these frequency elements are processed with NLSP, the diagonal frequency elements are emphasized to excess.

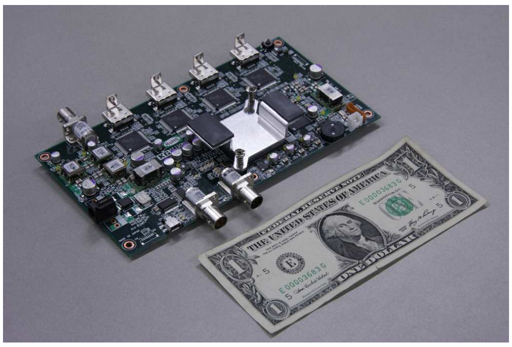


Figure 4: Appearance of real-time hardware.

The human visual system is not so sensitive to the horizontal and vertical high-frequency elements, i.e. the four corners shown in Figure 3 (Sakata, 1980). This means these frequency elements in the NLSP video do not affect the perceived resolution. Thus, to maintain the original diagonal resolution, the original diagonal frequency elements are sent through the delay line and added to the output. After the 2D-LPF the signal is provided into two paths. The upper path is the horizontal NLSP, and the lower path is the vertical NLSP. The three video paths are added together at the end to create the NLSP video. This parallel signal processing is fast. It reduces the delay from input to output, as discussed in section 1, and it can work at 60 Hz. Figure 4 shows the NLSP hardware. It upconverts full to 4K , and it processes the up-converted 4K video with NLSP to increase the resolution at 60 Hz. The NLSP algorithm is installed in the FPGA, which is located under the heat sink. Although there are many parts on the circuit board, most of them are input and output interface devices and electric power devices.

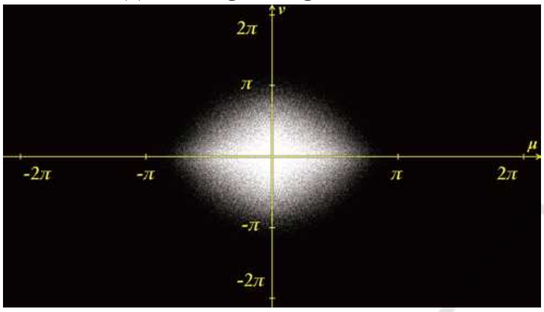
Figure 5 shows an image processed with the NLSP hardware shown in Figure 4. Figure 5(a) is just an enlargement from HD to 4K, and it looks blurry. Figure 5(b) shows the image processed with NLSP after the enlargement. Its resolution is clearly better than that of Figure 5(a). Figures 5(c) and 5(d) are the 2D-FFT results of Figures 5(a) and 5(b) respectively. In Figures 5(c) and 5(d), the horizontal axis the the horizontal frequency and the vertical axis the vertical frequency. HD Figure 5(d) has horizontal and vertical high-frequency elements that Figure 5(c) does not have. This shows that real-time hardware works and it produced the high frequency elements that the original image does not possess.

4 NLSP FOCUSING EFFECT

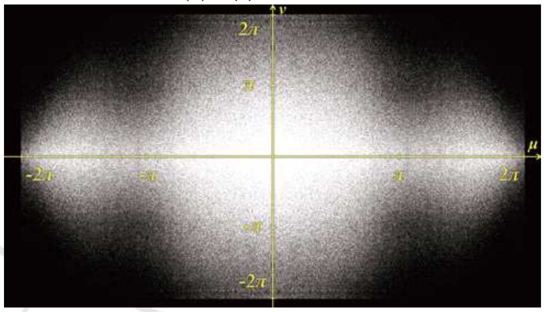
Focus is an important factor for creating finely detailed content. Professional cameras do not have autofocus functions because professional camera persons have the ability to adjust the fine focus and use complex focus controls on the HD cameras. It is very difficult to manually adjust the focus of 8K cameras using only a small viewfinder, and if the focus is off, the 8K format cannot live up to its full potential. Because 8K is developed for broadcasting, 8K cameras are equipped with zoom lenses as well as HD cameras. The focus of zoom lenses is less accurate than that of single-focus lenses. A zoom lens makes it more difficult to accurately adjust the focus. HD-to-8K upconverted videos are blurry and their character-

(a) 4K image enlarged from HD

(b) 5(a) with NLSP



(c) 2D-FFT result of Figure 5(a)



(d) 2D-FFT result of Figure 5(b)

Figure 5: Image processed with real-time NLSP.

(a) Blurry image

(b) Focused image with NLSP

Figure 6: Focusing effect.

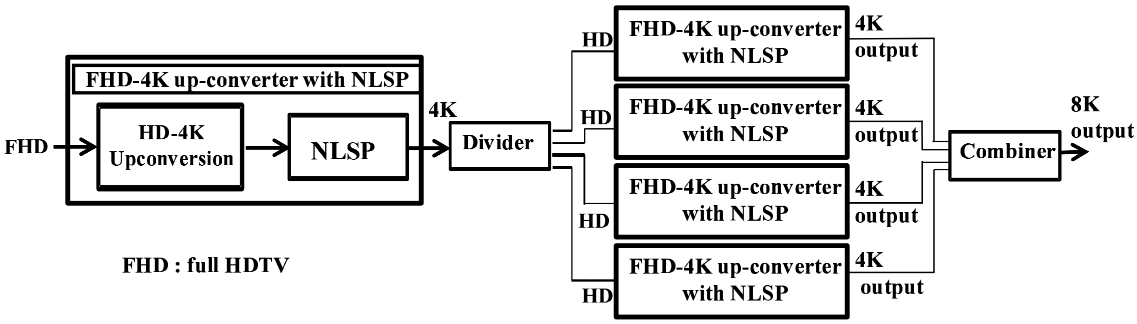


Figure 7: Block diagram of full HD to 8K upconverter with NLSP.

istics are similar to those of out-of-focus videos. It is important to note that NLSP has a focusing effect.

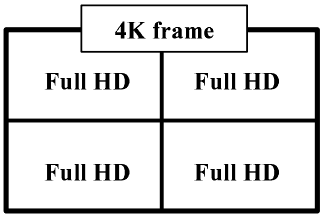


Figure 8: Divided 4K frame with full HD.

Figure 9: Real-time full HD to 4K upconverter with NLSP.

Figure 6(a) shows a blurry image. The original image is crisp (it is part of a test pattern), and a low pass filter (LPF) is used to blur the image. Figure 6(b) shows the result of processing the image in Figure 6(a) with the NLSP hardware. Comparing these figures, we observe that the resolution of the image in Figure 6(b) is better than that of Figure 6(a) and the focus looks adjusted. This effect is owing to the characteristics of NLSP. NLSP can generate high-frequency elements that the original image does not have, and these high-frequency elements have a focusing effect. The focusing effect works for the unconverted blurry 8K to improve the resolution.

5 REAL-TIME HD-TO-8K UPCONVERTER

5.1 HD to 8K Upconverter with NLSP

Figure 7 shows a block diagram of the HD-to-8K upconverter. The input, which is full HD, is shown on the left side and the 8K output is shown on the right side of Figure 7. The HD-to-8K upconversion is processed in two steps: full HD-to-4K and 4K-to-8K upconversion. The left half of Figure 7 shows the block diagram of the upconverter from full HD to 4K, which uses two dimensional Lanczos interpolation (Duchon, 1979). The upconverted 4K video from HD is blurry and is processed with NLSP to improve resolution. The real-time hardware for full HD to 4K upconversion with NSLP is shown in Figure 9.

The latter signal processing of the image in Figure 7 achieved via upconversion from 4K to 8K using NLSP. The upconverted 4K frame comprises four full HD frames, which are divided as shown in Figure 8. Each HD frame is processed with the same unit shown in Figure 9, and the four 4K frames with NLSP are created.These 4K frames are combined to create an 8K frame with NLSP. The real-time hardware for the 4K-to-8K upconversion with NLSP is shown in Figure 11, which includes four of the full HD-to-4K upconverter units shown in Figure 9. The other units are the divider and combiner units shown in Figures 7.

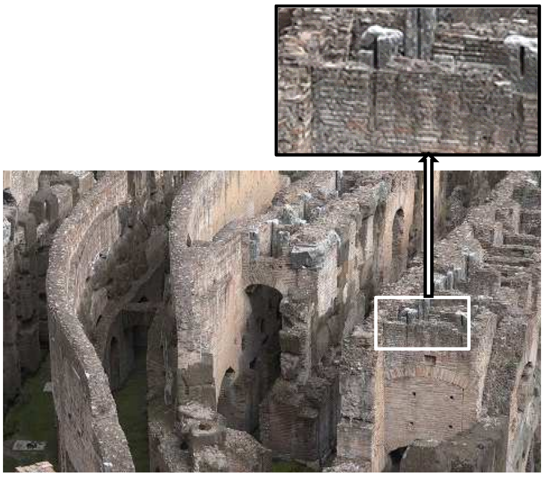
5.2 Resolution of the 8K Upconverted Image

Figure 10 shows parts of example 8K images upconverted using the real-time hardware shown in Figures 9 and 11. The images shown in Figures 10(a) and 10(c) are blurry because the NLSP option is off. The images in Figures 10(b) and 10(d) are created with the NLSP option on. The resolution of these images is better than of the ones in Figures 10(a) and 10(c). The only difference between them is whether the NLSP was used or not. Note that NLSP improves the resolution of upconverted 8K content. The focus effect discussed in Section 4 works, and it improves the resolution of the blurry image. The discussed hardware can upconvert images from full HD to 8K in real-time and will be useful for 8K broadcasting. It can address the problems of insufficient 8K content and is capable of upconverting HDTV content of varying quality to 8K.

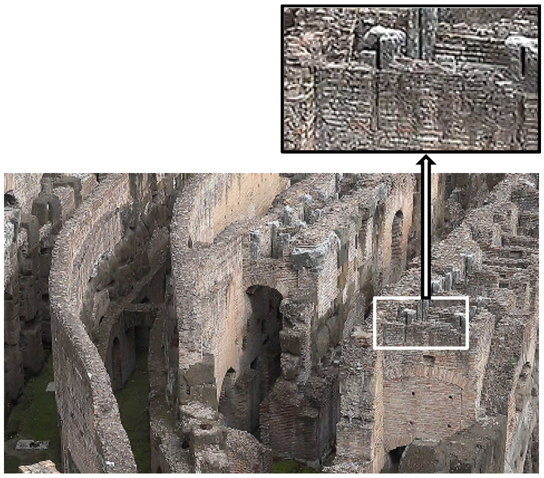
Currently the square function is applied to create the high frequency elements and it works. However, we should continue to try and find a better nonlinear function than the square function to improve image quality. 8K is a broadcasting system and there are various kinds of content such as news, drama, variety shows, sports and others. HDTV-to-8K upconverter have to process this content. Upconversion tests for the various content should be done before 8K broadcasting becomes operational.

5.3 Result

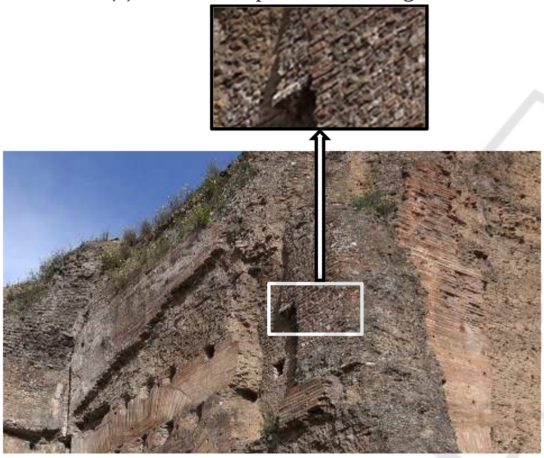
Upconverted videos are blurry, and this is a serious issue for HD-to-8K conversion. Although SR algorithms have been proposed, they are complex and cannot cope with real-time videos, particularly high-speed 8K videos. An algorithm for an HD-to-8K upconverter with NLSP was proposed and real-time hardware was developed. The converter creates high-frequency elements that the upconverted blurry video does not possess and produces high-resolution 8K content. This converter will be helpful for fixing the problems of insufficient 8K content and mobile news gathering for 8K broadcasting. Searching for a better



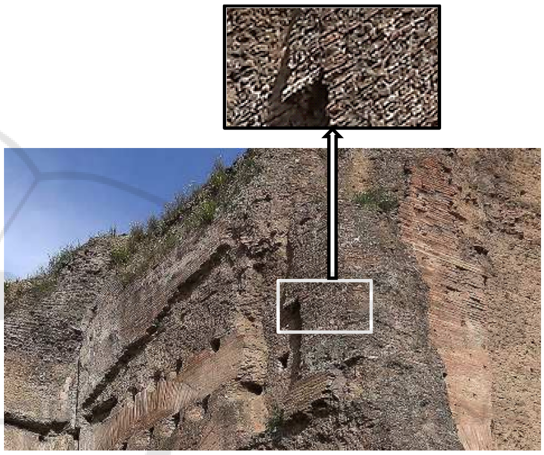
(a) HD to 8K upconverted image 1



(b) HD to 8K upconverted image 1 with NLSP



(c) HD to 8K upconverted image 2



(d) HD to 8K upconverted image 2 with NLSP

Figure 10: Upconverted images with real-time HD-to-8K upconverter.

Figure 11: Real-time 4K to 8K upconverter with NLSP.

nonlinear function and testing it with various content should be the primary focus forn the near future.

REFERENCES

Duchon, C. E. (1979). Lanczos filtering in one and two dimensions. *Journal of Applied Meteorology, Vol. 18, pp. 1016-1022*.

Farsiu, S., Robinson, D., Elad, M., and Milanfar, P. (2004). Fast and robust multi-frame super-resolution. *IEEE Transactions on Image Processing*.

Gohshi, S. and Echizen, I. (2013). Limitations of super resolution image reconstruction and how to overcome them for a single image. *ICETE2013 (SIGMAP), Reykjavik, Iceland*.

Gohshi, S., Hiroi, T., and Echizen, I. (2014). Subjective assessment of hdtv with super resolution function. *EURASIP Journal on Image and Video Processing*.

Houa, X. and Liu, H. (2011). Super-resolution image reconstruction for video sequence. *IEEE Transactions on Image Processing*.

ITU-R-HDTV (2015). https://www.itu.int/rec/r-rec-bt.709/en.

ITU-R-UDTV (2015). https://www.itu.int/rec/r-rec-bt.2020/en.

Lee, J. S. (1980). Ieee trans. on pattern analysis and machine intelligence 2:165-168. *Digital Image Enhancement and Noise Filtering by Use of Local Statistics*.

Matsumoto, N. and Ida, T. (2010). Reconstruction based super-resolution using self-congruency around image edges (in japanese). *Journal of IEICE*.

Mertz, P. and Gray, F. (1934). A theory of scanning and its relation to the characteristics of the transmitted signal in telephotography and television. *IEEE Transactions on Image Processing*.

Panda, S., Prasad, R., and Jena, G. (2011). Pocs based super-resolution image reconstruction using an adaptive regularization parameter. *IEEE Transactions on Image Processing*.

Park, S. C., Park, M. K., and Kang, M. G. (2003). Super-resolution image reconstruction: A technical overview. *IEEE Signal Processing Magazine*.

Pratt, W. K. (2001). *Digital Image Processing (3rd Ed): New York*. John Wiley and Sons.

Protter, M., Elad, M., Takeda, H., and Milanfar, P. (2009). Generalizing the nonlocal-means to super-resolution reconstruction. *IEEE Transactions on Image Processing*.

Sakata, H. (1980). Assessment of tv noise and frequency characteristics. *Journal of ITE*.

Sanchez-Beato, A. and Pajares, G. (2008). Noniterative interpolation-based super-resolution minimizing aliasing in the reconstructed image. *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 17(10):1817–1826.

Schreiber, W. F. (1970). Wirephoto quality improvement by unsharp masking. *J. Pattern Recognition, 2:111-121*.

Toshiba (Accessed 12 Sep 2016). https://www.toshiba.co.jp/regza/function/10b/function09.html (in japanese).

van Eekeren, A. W. M., Schutte, K., and van Vliet, L. J. (2010). Multiframe super-resolution reconstruction of small moving objects. *IEEE Transactions on Image Processing*.