Focus-aid Signal for Ultra High Definition Cameras

Seiichi Gohshi[1] and Hidetoshi Ito[2]

[1]*Kogakuin University, 1-24-2, Nishi-Shinjuku, Shinjuku0ku, Tokyo, Japan*

[2]*Leader Corporation, 2-6-33, Tsunasima Higashi, Kouhoku-Ku, Yokohama-Shi, Kanagawa, Japan*

Keywords: Focus, Super Resolution, Nonlinear Signal Processing, Edge Detection.

Abstract: 4K and 8K systems are very promising media and offer highly realistic images. Such high-resolution video systems provide completely different impressions than HDTVs. However, it is difficult, even for a professional cameraman, to adjust the 4K/8K camera focus using only the small viewfinder on a camera. Indeed, it is sometimes difficult even to focus an HDTV camera with such a small viewfinder, and since 4K has four times higher resolution than HDTV, it is almost impossible to adjust a small viewfinder with the same size as that of an HDTV camera using only human eyes. Therefore, in content-creating fields, large monitors are generally used to adjust the focus; however, large monitors are bulky and do not fit practical requirements, which means that technical assistance is required. A possible solution to this problem is to detect the sharp edges created by high-frequency elements in fine-focus images and superimpose those edges on the image; the cameraman can then adjust the focus with additional information gained from maximizing the superimposed edges. However, conventional edge detection technologies are vulnerable against noise, which means that practical situations using this technique are limited to environments with good lighting conditions. This paper introduces a novel signal processing method that enables cameramen to adjust a 4k camera focus using their eyes.

1 INTRODUCTION

4K and 8K video systems provide highly realistic experiences and are said to be the ultimate in TV systems. 4K TVs and 4K video cameras are currently sold in stores, and experimental 4K broadcasting has begun in Japan; moreover, an 8K (SHV) broadcasting service is planned for 2016. However, 4K/8K content is still not sufficient and current professional 4K/8K equipment is bulky and heavy. To create more 4K/8K content, the size of the equipment needs to be reduced. However, even if their size were reduced, there would still be the problem of focusing the 4K/8K camera. 4K and 8K has four and sixteen times higher resolution than full HDTV, respectively. However, the viewfinders on these cameras are small. Professional cameras do not have auto-focus functions because professional camera persons are capable of adjusting the fine focus and complex focus controls. Producers sometimes use blurry scenes and gradual temporal focus in and out of scenes. There are many focus techniques in the content production. The professional camera persons were able to cope with these difficult requirements until HDTV content production. However it is very difficult to manually adjust the focus using only the viewfinder found on 4K/8K cameras, and if the focus is off, 4K/8K cannot live up to their full potential.

To solve this problem, large 4K liquid crystal displays (LCDs) are used in the field to adjust the focus. These displays are bulky and sometimes impossible to use in small places. Although small monitors are handy, they are insufficient for adjusting high-resolution 4K/8K. Thus, a new technology that can help human eyes adjust the focus would be much appreciated. Edges in a frame can be an indicator of the focus. Edges have their highest frequency elements when the focus is adjusted. It is not difficult to detect edges with a digital high-pass filter (HPF). Thus, one possible solution would be superimposing the edges of the image on the viewfinder. Adjusting the focus on small viewfinders would be easier if the edges were pronounced, and adjusting the focus to maximize edges is not difficult with small viewfinders; therefore, this method can provide a better focus point than just concentrating on the viewfinder's regular image. However, this method has limitations. Video always has noise, which creates false edges; these false edges interfere with adjusting the focus. Furthermore, the edges detected by conventional, digital HPF are thick and have low levels that can be compared with noise. The thick and low edges re-

Gohshi, S. and Ito, H.
Focus-aid Signal for Ultra High Definition Cameras.
DOI: 10.5220/0005750901740179
In *Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2016) - Volume 3: VISAPP*, pages 176-181
ISBN: 978-989-758-175-5

gardless off the image being out of focus. Sharp, high edges are required to adjust the focus, and these appear at the fine focus position. Thin edges can be generated if the edges have high-frequency elements. Super resolution (SR) is one method for creating high-frequency elements. However, most SR technologies cannot work in real time (Farsiu et al., 2004)(Park et al., 2003)(Katsaggelos et al., 2010)(van Eekeren et al., 2010)(Freeman et al., 2000)(Freeman and Liu., 2011)(Zhu et al., 2014).

SR with nonlinear signal processing (NLSP) has been recently proposed. NLSP can work in real time and can create higher-frequency elements than the original image. These high-frequency elements can create thin edges, and their quantity rapidly decreases except for the fine focus point. NLSP also has high tolerance against noise. This paper discusses focus adjustment for 4K using SR with NLSP.

2 FOCUS AND EDGE

A fine focus produces a crisp video frame. The difference between a crisp frame and a blurry frame is obtained by determining how fine the focus is. A crisp frame has high-frequency elements; a blurry frame does not. Therefore, the amount of high-frequency elements in a frame can be a barometer for the state of the focus. High-frequency elements are the edges in a frame. A focused frame has more edges than an unfocused frame. Edges in a frame appear when a camera shoots a scene that has detailed textures. When the focus is turned on, the maximum amount of edges appears. This type of focus aid system has been developed (Funatsu et al., 2013). However, it is difficult to judge whether the focus is obtained on the basis of the edges detected with the conventional HPF method because of low level edges and noise.

Figures 1, 2, and 3 provide an example. Figure 1 is the original video frame; Figures 2 and 3 show the frame with the absolute value of the edges superimposed on it. The luminance level is modified to check the edges clearly, and HPF is used to detect the edges. Figure 2 shows an out-of-focus result, and Figure 3 shows a fine focus result. Although the edges in Figure 2 are stronger than those in Figure 3, both look similar; it is difficult to see the difference. Moreover, noise is visible in both, which makes it more difficult to adjust the focus with the edges. This is the limitation of using conventional edge detection technology: noise in a frame has high-frequency elements just like the edges. Thus, even if the characteristics of HPF change, the result will be similar because of the presence of noise.

Figure 1: Original image.

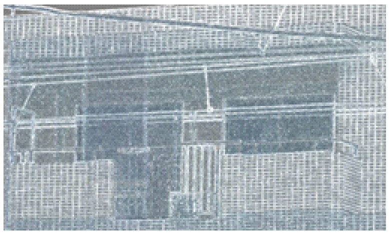

Figure 2: Out of focus image.

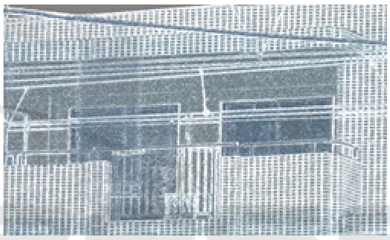

Figure 3: Fine focus image.

3 NLSP

The only practical method for improving resolution in real time has been Enhancer (unsharp mask), which is discussed here to clarify the real-time edge detection method (Schreiber, 1970) (Lee, 1980) (Pratt, 2001). Figure 4 shows the block diagram of a typical enhancer. It uses HPF to detect edges, which are then processed with a limiter (LMT) until achieving appropriate levels for the image. The edges are added to the image with the adder block (ADD). Enhancer merely amplifies the high-frequency elements in an image, which means that noise is also amplified and added to the image; this is exactly what happens in Figures 2 and 3. Edges and noise are both amplified and added. This issue will be discussed again in the section 4. Noise always appears in an image except for sufficient lighting conditions such as in sunny places in a fine day. The current technology (Funatsu et al., 2013) can be usable only in good lighting limited conditions. A new technology is necessary to distinguish valuable edges from noise.

NLSP has been recently proposed to improve the resolution of videos in real time. NLSP can distin-

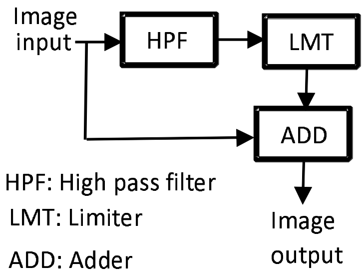

Figure 4: Enhancer.

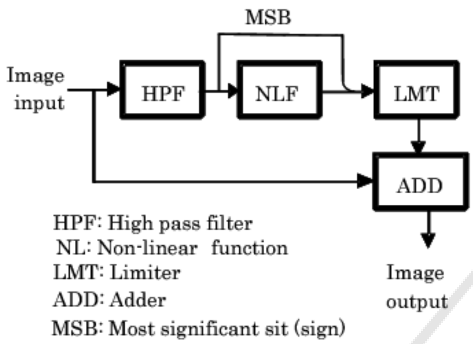

Figure 5: NLSP.

guish edges from noise. Figure 5 shows the signal flow of NLSP. The upper path comprises an HPF, a nonlinear function (NLF), and the LMT. The edges in the video are detected with the HPF and then processed with the NLF, which generates harmonic waves from the edges. These harmonic waves have higher-frequency elements than what the original video has and are generated only from the edges detected with the HPF. There are no harmonic waves in flat areas, because there are no edges in flat areas. An example of an NLF is a cubic function. The range of the input of the NLF is from 255 to 255 if the depth of the video is 8 bits. The output of the NLF becomes very large, because the cubic function generates the pixels from 255^3 to 255^3. The LMT saturates these large values to fit the harmonic waves to the video. The lower path is from the input and is directly connected to the ADD. The ADD adds the LMT-processed harmonic waves to the original video. Thus, the output of the ADD has high-frequency elements that the original video does not have. This video processing method can improve resolution and create high-frequency elements that exceed even the Nyquist frequency.

Figure 6 shows a simulation result of using NLSP. Figure 6(a) is a still image that has been enlarged horizontally and vertically by a factor of two relative to the original. Figure 6(c) shows the two-dimensional fast Fourier transform (FFT) of (a). It does not have high-frequency elements, because it is an enlargement image. Figure 6(b) shows the output of the image after it has been processed with NLSP according to the flow shown in Figure 5. Figure 6(d) shows the two-dimensional FFT of Figure 6(b). Comparing Figure 6(d) with Figure 6(c), we can see that (d) has high-frequency elements that (c) does not have. Because Figure 6(a) is horizontally and vertically enlarged by a factor of two, the horizontal Nyquist frequency of the original image is /2 in Figure 6(c). There are no spectra beyond /2 in Figure 6(c), because the enlarged image has the same spectra as the original image. In contrast, the spectra in Figure 6(d) have elements that Figure 6(c) does not have; it exceeds the Nyquist frequency (/2) of the original image. This proves that the proposed method can create higher frequency elements than the Nyquist frequency of the original image. It is also important to note that Figure 6(b) does not have visible noise. Noise in NLSP is discussed in the next section.

4 EDGE DETECTION

Figure 7(a) shows the graphical image of edges created by conventional HPF and Figure 7(b) shows the edges created by NLSP. The edges created by NLSP are stronger and sharper than those created using conventional HPF. Moreover, the edges are also thinner and higher because harmonic high-frequency elements with NLF are added to the original edges. The edges created with NLSP show the focus point with strong edges, which help in focus adjustment.

Here we discuss the edge detection method with conventional HPF (Funatsu et al., 2013) and with NLSP. In generally, noise is smaller than the edges in images and videos. Figures 8(a) shows an example of edges detected from video by the conventional method and Figure 8(b) shows an example of edges detected from video by NLSP. A threshold level is selected to detect the true edges in a video from the edges by noise. However, in Figure 8(a) the true edges detected by HPF are not sufficiently larger than the edges with noise. The edges detected by conventional HPF have levels similar to the noise. If we define a threshold level to discriminate the edges in an image from noise, the allowance for the level is narrow, as shown in Figure 8(a). The detected edges are very small and it is very difficult to separate the true edges in the image from those detected by noise. If we detect the true edges in the video, we have to lower the threshold level.

In this case as shown in Figure 8(b) the appropriate threshold level does not exist and edges cannot be separated from the edge. The edges caused with noise

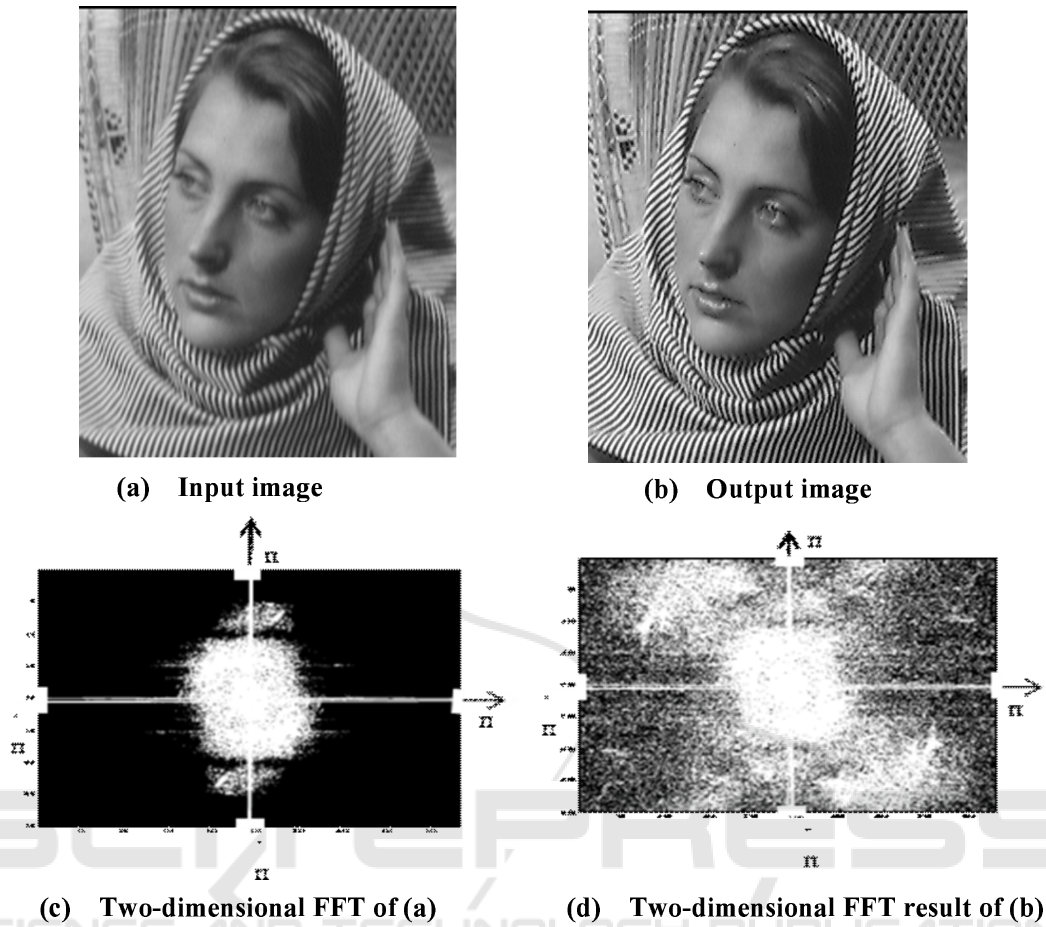

Figure 6: Image processed NLSP.

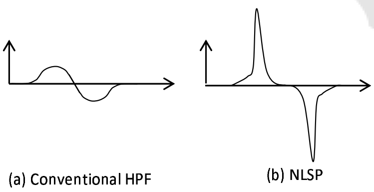

Figure 7: Created edges.

are also detected in video and it is impossible to adjust the focus. This is exactly what happened in Figures 2 and 3. Conversely if we higher the threshold level to remove the edges caused by noise, the true edges in video are also removed.

Conversely, the edges detected by NLSP are larger than those detected with HPF and noise can be suppressed by the nonlinear function shown in Figure 9(a). Edges in the image are amplified by NLSP and it becomes the high level edge(HLE) in Figure 9(a). Conversely edges detected by noise (EN) are small. NLSP makes the edge level differences bigger between HLE and EN. As shown in Figure 9, the levels of HLE and EN can be separated with an appropri-

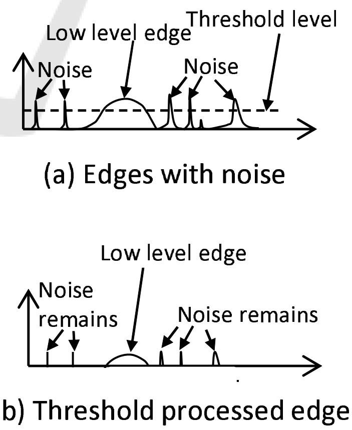

Figure 8: Threshold process for HPF.

ate threshold level. In Figures 8 it is very difficult to find the threshold level to separate the true edges from noise because their level are similar. In Figure 9 it is easy to select the threshold level than that in Figure 8. By controlling the threshold level, noise can be suppressed so that only edges are detected, as shown in

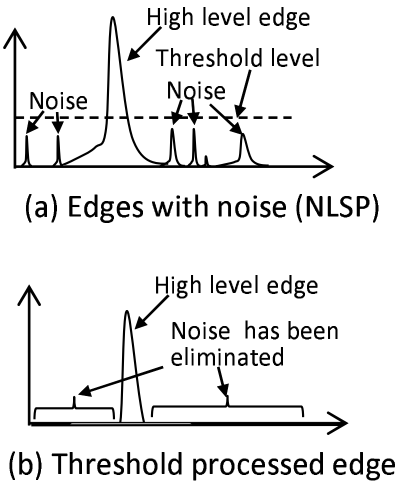

Figure 9: Threshold process for NLSP.

Figure 9(b). The edge shown in Figure 9(b) is thinner and larger than that shown in Figure 8(b). The edges amplified by the nonlinear function are much more visible than are the edges detected by conventional HPF, even though the characteristics are deeply deliberated. The edges detected by NLSP are sufficiently large to adjust the focus.

When we decide the parameters of enhancers, noise is always problem. If we try to amplify the edges very strong, noise becomes visible in flat areas in the image; Figure 10 is an example of such noise. Compared Figure 10 with Figure 6(b), noise in Figure 10 is visible in the flat areas such as in the forehead and cheeks. This happens in Figures 2 and 3. NLSP creates harmonics to make edges thinner and higher. Conversely although noise is also processed by NLSP, the energy of noise is small. Turning the parameters of NLSP, it is possible for NLSP not to make small edges unnecessarily higher; Figure 6(b) is an example. Noise is not visible in the forehead and cheeks.

Real-time functionality is a fundamental requirement for video equipment, and adjusting the lens focus is no different: it must be able to be done in real time. Furthermore, the focus adjusting system should be installed in a small device. If the system is in a bulky device, it will be impractical for daily use. The proposed NLSP algorithm meets both of these requirements because NLSP has been successfully implemented in an field programmable gate array and has worked as SR equipment for both 4K and 8K.

5 EXPERIMENT WITH REAL-TIME HARDWARE

The real-time hardware shown in Figure 11 was developed to prove the validity and practicality of the proposed method. Stuffed animals are set and shot with a 4K camera. The 4K camera is connected to the hardware and an LCD shows the stuffed animals and the edges. The LCD displays a stuffed bear with white dots appearing only on its face and hat. These dots are the image's edges and appear at very limited distances from the 4K camera; this means that the focus is adjusted on the bear's face. These white dots appear only on focused areas and can assist in adjusting the focus. Although the LCD is 9 inches, the edges are sufficiently visible. A cameraman adjusts the focus to maximize the edges he wants to focus on. Bulky, 55-inch monitors are not necessary to adjust the focus for 4K/8K.

Figure 10: Enhancer processed image.

Figure 11: Real time hardware.

As discussed in Section III, NLSP is valuable to use against noise; the conventional HPF edge detection method does not have such noise tolerance. Noise is especially a problem for dark, low-luminance scenes. Here we discuss this issue with the image shown in Figure 12. Figure 12 is shot with a 4K camera and Figures 13 and 14 were the same ones taken in a dark room. They show the difference between the NLSP and conventional edge detection methods in low-luminance conditions. In these figures, boxes with ribbons are set in the frame and taken with a 4K camera. A white rectangle in Figure 13 is drawn to in-

Figure 12: Original image.

Figure 13: Edge with the proposed method.

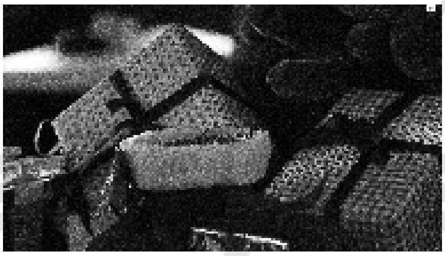

Figure 14: Edge with the conventional method.

dicate the area where the white dots exist; they appear within the rectangle only at limited distances from the 4K camera, which means that the 4K camera focused on the rectangle area. Figure 14 is shot with the conventional edge detection method. In that figure, the white dots appear all over the image because of noise and strong edges that are less visible than those of Figure 14. In Figure 13, it is impossible to determine what the camera focused on.

6 CONCLUSION

A real-time focus adjustment algorithm for 4K, which will aid human visual systems, is proposed. It produces edges only in small, on-focus areas, and these edges are easily detected by human eyes even when small LCDs are used. It also has high tolerance against noise under low-light conditions. The system is suitable for practical use in creating 4K content, unlike large monitors.

REFERENCES

Farsiu, S., Robinson, M. D., Elad, M., and Milanfar, P. (2004). Fast and robust multi frame super resolution. *IEEE TRANSACTIONS on Image Processing*, 13(10):1327–1344.

Freeman, W. T. and Liu., C. (2011). Markov random fields for super-resolution and texture synthesis. *MIT Press, Advances in Markov Random Fields for Vision and Image Processing*.

Freeman, W. T., Pasztor, E. C., and Carmichael, O. T. (2000). Learning low-level vision. *International Journal of Computer Vision*, 1(40):25–47.

Funatsu, R., Yamashita, Y., Mitani, K., and Nojiri, Y. (2013). Focus-aid signal for super hi-vision cameras. Technical Report 53.

Katsaggelos, A., Molina, R., and Mateos, J. (2010). *Super Resolution of Images and Video: Synthesis Lectures on Images, Video and Multimedia Processing*. Morgan and Clayppo Publishers, La Vergne TN USA.

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

Park, S. C., Park, M. K., and Kang, M. G. (2003). Super-resolution image reconstruction: A technical overview. *IEEE Signal Processing Magazine*, 1053-5888/03:21–36.

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

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

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*, 19(11):2901–2912.

Zhu, Z., Guo, F., Yu, H., and Chen, C. (2014). Fast single image super-resolution via self-example learning and sparse representation. *Multimedia, IEEE Transactions*, 16(8).