Super Resolution for Smartphones

Seiichi Gohshi[1], Sakae Inoue[2], Isao Masuda[2], Takashi Ichinose[2] and Yoshikatsu Tatsumi[2]
[1]*Kogakuin University, 1-24-2, Nishi-Shinjuku, Shinjuku-ku, Tokyo, Japan*
[2]*Fujitsu Connected Technologies Limited, 1-1, Kamikodanaka 4-chome, Nakahara-ku, Kawasaki, Japan*

Keywords: Smartphone, Real-time Video Processing, Nyquist Frequency, Super Resolution.

Abstract: Smartphones were developed as an advanced communication tool. Currently they are used in various applications. The display is one of the most important features in smartphones. Compared with television (TV) and cinema screens the display size of a smartphone is small. However, TV and film content is commonly enjoyed on smartphone screens. Currently, the smartphone display is one of the most used displays for various kinds of content. In the past it was thought that it would be difficult to recognize the resolution differences on small displays. However, this is no longer the case. The resolution of smartphones have been steadily improving, and high-definition television (HDTV) ($1,920 \times 1,080$ pixels) viewing resolution support is common. Signal processing is another way to improve resolution. Super resolution (SR) has become an interesting research field and is applied to images and videos. SR is a technology for improving display resolution. Consequently, SR is mainly studied for application to TV screens and computer displays. SR technology algorithms are complex and a heavy load for a smartphone's central or graphics processing unit (CPU/GPU). It is very difficult to apply SR for real-time videos on smartphones. Consequently, there have been no reports in SR for smartphones. This paper proposes a method for implementing real-time SR in smartphones. This method works for real-time videos on a smartphone GPU with the developed software.

1 INTRODUCTION

Communication environments and devices have dramatically changed in the last two decades. Smartphones have become major devices on the mobile phone market, and new models are introduced every year. Smartphones are all-in-one small computers and come equipped with various functions. The display on a smartphone is used as an input terminal as well as a conventional display for video content. Smartphones have thus become important devices for enjoying television (TV) and cinema content as well as games. Smartphone manufacturers are constantly developing new products and trying to stand out from the competition. The resolutions of smartphone monitors are increasing, and high-definition TV (HDTV) resolution ($1,920 \times 1,080$ pixels) is now common. Some of the latest smartphones now come with 4K resolution. However the display sizes are approximately five inches. If there were an obvious resolution quality difference between a small five-inch HDTV display and a 4K one, it would be worthwhile to invest further in 4K smartphone technology. However, there has been no discussion or subjective assessments of using 4K displays on smartphones. Moreover, none of the content for smartphones on the Internet has 4K resolution, and some do not have HDTV resolution. Despite the resolution of the Internet content, only the number of pixels of a display is increasing for marketing. When the resolution of the content is not sufficient for the display, as with video graphics array (VGA) or quarter VGA (QVGA), the content is specially interpolated to be fixed with the resolution of the smartphone display. Interpolated images and videos are blurry and cannot take full advantage of the performance of a smartphone display. Although the content has HDTV pixels, HDTV resolution is not always guaranteed because the focus is not always fine. Clearly, smartphone users prefer high-resolution images and videos. However, high-resolution displays (HDTV/4K) do not always provide high-resolution images and videos. Resolution of the content is more important than the number of pixels in displays.

Improving image resolution has been a highly active field of research for many years. Unsharp masking (USM) also referred to as edge enhancement, has been the only method for enhancing video in real-time systems (Schreiber, 1970)(Lee, 1980)(Pratt, 2001). Although USM is a simple and cost-effective method, it does not actually improve resolution; it provides

Gohshi, S., Inoue, S., Masuda, I., Ichinose, T. and Tatsumi, Y.
Super Resolution for Smartphones.
DOI: 10.5220/0005991301060112
In *Proceedings of the 13th International Joint Conference on e-Business and Telecommunications (ICETE 2016) - Volume 5: SIGMAP*, pages 106-112
ISBN: 978-989-758-196-0

a better image quality using either a band-pass filter (BPF) or a high-pass filter (HPF). However, USM can introduce noise and edges to images.

Super resolution (SR) technology, which has been studied for approximately two decades (Elad and Feuer, 1996), creates a high-resolution image (HRI) from low resolution images (LRIs). Various SR technologies have been proposed during the past ten years. However, most of these are proposed for still images and are difficult to apply to videos owing to because of their complex algorithms (Farsiu et al., 2004)(Park et al., 2003)(Katsaggelos et al., 2010)(van Eekeren et al., 2010)(Panda et al., 2011)(Glasner et al., 2009)(Park et al., 2003) (Sun et al., 2008)(Dong et al., 2014). Recently, super resolution with non-linear processing (SRNP) has been proposed by one of the authors. SRNP can process video in real time.This paper proposes SRNP for smartphones for real-time video processing. Our method can work with software over the central or graphic processing unit (CPU/GPU) of a smartphone. It shows good results and improves videos to fine quality on the display of a smartphone.

2 SUPER RESOLUTION FOR SMARTPHONES

SR is a technology for improving image and video resolution. As discussed in the previous section there have been many methods and proposals (Farsiu et al., 2004)(Park et al., 2003)(Katsaggelos et al., 2010)(van Eekeren et al., 2010)(Panda et al., 2011)(Glasner et al., 2009)(Park et al., 2003) (Sun et al., 2008)(Dong et al., 2014). The size of the monitor becomes an important factor in seeing SR-processed image results. This point has not been discussed in SR research. SR studies freely select their processed image sizes to recognize the resolution improvement. Personal computer (PC) monitors are used to check image resolution. Although commercial HDTV sets with SR functions are also available (Tos, 2009), the sizes of HDTV screens are 40 inches or larger. It would be difficult to recognize improvement with SR on a small smartphone monitor. If we are to implement SR technology, resolution improvement must be recognizable on smartphone displays.

Smartphones are sophisticated devices, but it is impossible to add devices to a smartphone to use SR. There are two difficulties in implementing SR on a smartphone with limited resources. The first is the complexity of SR algorithms. Many SR algorithms have been proposed (Farsiu et al., 2004)(Park et al., 2003)(Katsaggelos et al., 2010)(van Eekeren et al., 2010)(Panda et al., 2011)(Glasner et al., 2009)(Park et al., 2003) (Sun et al., 2008)(Dong et al., 2014). Super resolution image reconstruction (SRR) and learning-based super resolution (LBSR) are typical SR technologies, and many others have been proposed (Farsiu et al., 2004)(Park et al., 2003)(Katsaggelos et al., 2010)(van Eekeren et al., 2010)(Dong et al., 2014). However all SR algorithms including SRR and LBSR are difficult to use in real time for video because they require iteration to create a high-resolution image. Iteration is very time consuming and difficult to execute on the CPU/GPU of a smartphone. Although a non-iterative SRR algorithm for HDTV has been proposed (Matsumoto and Ida, 2010), its resolution is lower than that of a conventional HDTV, and an additional device is required to use the SRR algorithm.

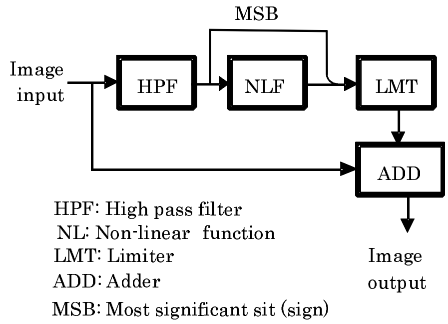
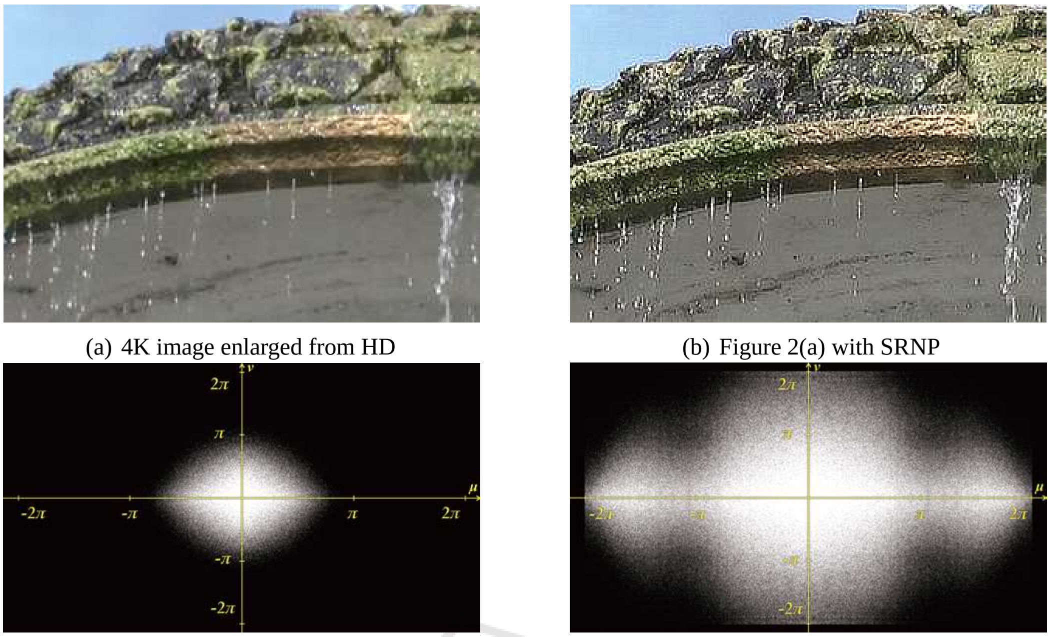


Figure 1: SRNP algorithm.

The second difficulty is that SR on smartphones must work on the CPU/GPU of a smartphone. It is difficult for smartphones to handle additional devices required for SR to work, because of the space and power consumption. There is almost no space to implement additional parts on a smartphone, and new parts shorten battery duration owing to higher power consumption. If we try to make SR work on a smartphone, SR would have to work with the CPU/GPU and its resources on the smartphone. The CPU/GPU executes many tasks, and resources such as the memory bandwidth are limited. If sufficient CPU/GPU power and resources are not provided for the SR process, video cannot be processed in real time, frame drops can occur, and in the worst case, the video will freeze. To overcome these difficulties, an SR algorithm for a smartphone must be simple and sufficiently light to work on CPU/GPU power and limited resources.

3 SRNP

SRNP was developed for upconversion from HDTV



(a) 4K image enlarged from HD

(b) Figure 2(a) with SRNP

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

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

Figure 2: Image processed with real-time SRNP hardware.

to 4K. Figure. Figure 1 shows the signal flow of the proposed method. The input video has two paths. The first path consists of a high-pass filter (HPF), a non-linear function (NLF), and a limiter (LMT). This path creates high-frequency elements that the original video does not have. The edges in the video are detected with the HPF. The detected edges are then processed with the NLF. An example of an NLF is a cubic function $y = x^3$. The NLF generates harmonic waves from the edges. It is well known that images and videos can be expanded with Fourier series. Fourier series consist of sine and cosine waves. Using the cubic function, the *sin* and *cos* functions are changed to $(sin\theta)^3$ and $(cos\theta)^3$. $(sin\theta)^3$ can be changed to $(sin3\theta)$and $(cos\theta)^3$ can be changed to $(cos3\theta)$. $(sin3\theta)$ and $(cos3\theta)$ are harmonic waves, and the harmonic waves have higher frequency elements that the original video did not have. The cubic function is just an example of a nonlinear function, and the NLF is used to create the high-frequency elements by the harmonic waves. The harmonic waves are generated only from the edges detected with the HPF. There are no harmonic waves in flat areas since there are no edges in flat areas. The LMT saturates these large values to fit the harmonic waves to the video. The second path is from the input and is directly connected to the adder (ADD). The ADD adds the LMT-processed harmonic waves to the original video. This process is conducted pixel by pixel. The output of the ADD thus has high-frequency elements that the original video did not have. This video processing method can improve the resolution and even create high-frequency elements that exceed the Nyquist frequency of the original video. It is a simple algorithm and real-time SRNP hardware for it has been developed.

Figure 2 shows an image processed with SRNP hardware. Figure 2(a) is an enlargement from HDTV to 4K. Figure 2(b) shows the SRNP processed result of Figure 2(a). Although Figure 2(a) is blurry, Figure 2(b) is clearly superior to Figure 2(a). Figures 2(c) and 2(d) are the two dimensional fast Fourier transform (2D-FFT) results of Figures 2(a) and 2(a) respectively. Figures 2(a) and 2(a) show the frequency characteristics in the frequency domain. The horizontal and vertical axes are the horizontal and vertical frequencies of the image. Note that Figure 2(d) has horizontal and vertical high-frequency elements that Figure 2(c) does not possess. This means that SRNP improves resolution.

However, this has not worked for video in real-time over CPU/GPU with software. As discussed in the previous section, the resources in a smartphone are limited, which makes it difficult for SR to work in real-time. Even though the SRNP algorithm is simple, there is no guarantee that it will work for video in real-time on a smartphone.

4 SYSTEM ARCHITECTURE OF SRNP FOR SMARTPHONES

Smartphones are very much the same as compact computers equipped with the latest technologies in order to provide users with various services. However, there are not sufficient resources and power in a smartphone to execute all of the tasks of a PC. Although SRNP is a simple algorithm, it might not work when there are insufficient resources. Constructing an experimental SRNP system on a smartphone platform and optimizing the system is the only way to make it practical.

The specifications for the experimental hardware are shown in Table 1. The smartphone has a MSM8992 CPU and an Adreno 418 GPU , The operating system is Android 5.1. SRNP signal flow over the hardware is shown in Figure 3. The video is coded with MPEG-4 H.264 and stored in the SD memory. The MPEG-4 stream is decoded using a hardware decoder. Although the most common video is HDTV (1920×1080), there are other formats available on the Internet. The SRNP process is performed with the original formats, such as HDTV, QVGA and so on. The GPU conducts the SRNP process to improve resolution depending on the input video format including HDTV. The output video format of the GPU is the same size as that of the input video. The GPU processes the video by frames, because the Enlarge/Shrink unit after SRNP requires an entire frame to change the video format.The GPU is controlled by the CPU and shares memory and other resources with other units. The GPU also controls its timing when delivering frames to the Enlarge/Shrink unit in 33.3 ms (30 Hz), in order not to freeze the display. The liquid crystal display (LCD) can display video in real time, if the all signal processing works at 30 Hz without delay.

A potential bottleneck of the system is the GPU because it must finish one frame of the SRNP processing within 33.3 ms. There are two difficulties in overcoming this bottleneck. First, the GPU calculation of the SRNP process itself must finish in 33.3 ms. Second, the memory bandwidth must be adequate. The input output video frame of the GPU is stored in 3 GB of RAM, shown in Table 1. The RAM is also used by other applications and is accessed by the CPU. Since the RAM for the GPU is limited, arbitration between the CPU and GPU is required. Two tunings are necessary to arrange the tasks for dual and quad CPU cores to access RAM and improve performance of the GPU for SRNP. The difficulty of the tuning is proportional to the higher frame rate and larger screen. The high frame rates require short periods of SRNP, which di-

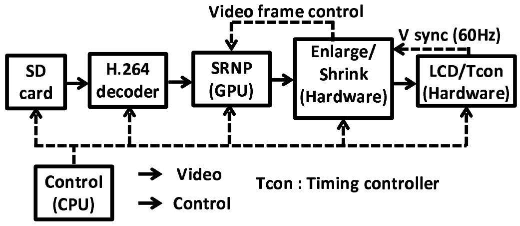


Figure 3: Signal flow of a smartphone.

rectly corresponds to the capability of the GPU. Large screens require more memory. Both of these are related to the memory bandwidth. Currently, the frame rate of smartphones is 30 Hz. The GPU must finish SRNP in 33.3 m for a frame, because smartphones display videos with 30 Hz (33.3 ms).

The bottleneck of the system is the processing time of SRNP in the GPU. The processing time of SRNP in the GPU increases in proportion to the screen size. Tuning the GPU to program SRNP can work up to the ultra-HD (UHD) $3,940 \times 2,160$ display resolution. Table 2 shows the relationship between the screen sizes and the processing times of the GPU. Currently, 4K is among the biggest practical displays available in the market. The processing time for 4K is 14.1 ms. This means that SRNP can work in real time for 4K. Moreover, SRNP can work at 60 Hz (16.7 ms) for 4K because the SRNP time for 4K is 14.1 ms, which is shorter than 16.7 ms. The simple SRNP algorithm embodies a short processing time. If other parts of the hardware, such as the H.264

Table 1: Specification of experimental smartphone.

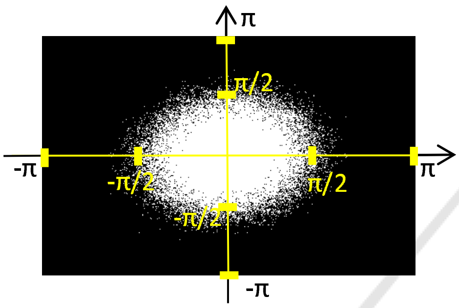
Size	154 (H) x 75 (W) x 7.9 (D) mm
Mass	174 g
CPU (system LSI)	QualcommMSM8992 1.8 GHz Dual Core1.4 GHz QuadCore
GPU (on the chip)	QualcommAdreno 418 600 MHz
Memory	RAM: 3GB, ROM: 32GB
Battery	3390 mAh
OS	Android 5.1
LCD	5.4inch WQHD1440$\times$2560

Table 2: GPU processing time (30 frame/s).

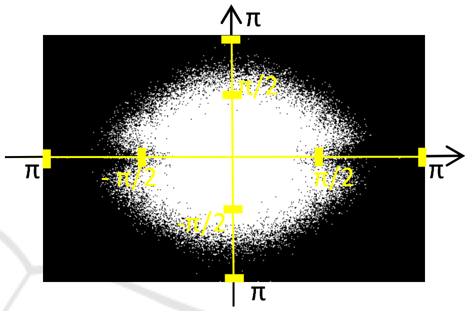
Size of image	GPU processing time (ms/frame)
UHD (3940×2160)	14.1
WQHD (2560×1440)	7.2
Full HD (1920×1080)	4.6
(1280×702)	3.9

(a) Input image

(b) Image processed with the developed smartphone



(c) 2D-FFT result of Figure 4(a)



(d) 2D-FFT result of Figure 4(b)

Figure 4: Image processed with an SRNP smartphone (software processed).

Figure 5: Developed Smartphone with NLSP.

decoder become faster, then SRNP can process even 4K 3,840× 2160;60 Hz).

5 RESULTS

Figures 6(a) to 7(b) show the resolution improvement of SRNP with a smartphone. They are not still images but frames of videos. Figures 6(a), and 7(a) are input images and Figures 6(b), and 7(b) are processed with SRNP. The resolution of the SRNP-processed images is better than that of the input images. Just comparing the input images and the SRNP-processed images is a subjective assessment.It is sometimes difficult to recognize the resolution improvement between the SR-processed images and the images processed with a conventional enhancer. An enhancer just amplifies the edges in the images. Our SRNP is completely different from the conventional enhancers. SRNP can create higher frequency elements that the original image does not possess. As discussed in Sections 3 and 4, SRNP creates higher frequency elements that the input image does not possess. Owing to space limitations, all of the 2D-FFT results cannot be shown. However, all of the SRNP-processed images are improved in their resolution. These videos have distortions caused by MPEG-4 because they were taken using a commercial video camera. However, the distortions did not affect the image quality and the resolution improvement from SRNP is conspicuous. These videos had high-frequency elements, which made MPEG-4 decoding heavy and time consuming. However, all of them were displayed at 30 Hz without any frame drops or freezing. This shows that SRNP can work in real-time on a smartphone.

Although our smartphone with SR was developed for playing real-time video, we used still images, includ-

(a) Original image (b) Image processed with the developed smartphone

Figure 6: Image No. 1.

(a) Original image (b) Image processed with the developed smartphone

Figure 7: Image No. 2.

ing documents in our assessments. Observers can control the SR switch (on/off) to compare the image with SR on and off. We conducted subjective assessments. Due to space limitations, the details of the assessments cannot be explained here because it takes several pages to explain everything, including the analysis results of the assessments. After the assessments, many observers said that the letters in the documents became easier to read when SR was on because the contrast of the LCD was enhanced. This means that we can make the backlight of the LCD lower and reduce the power consumption. SR also has the potential to extend the battery life of a device.

6 CONCLUSION

Many SR technologies have been proposed, but to date, the signal-processing load was heavy, and SR could not run on a smartphone. We proposed an SR technology that can work with smartphone hardware. It works for 60 Hz video in real-time without additional devices. The developed smartphone shows higher-frequency elements that the original video did not possess. 2D-FFT results were presented to prove that it works. Although displays of smartphones are small, the resolution improvement can be detected easily by a smartphone user. SR for the smartphone was developed for real-time video. However, it comes with a side benefit. SR enhances the contrast, and the letters in documents become easier to read. The developed smartphone is equipped with wide quad HD (WQHD), 1440×2560 pixels display. 4K displays are beginning to be implemented on smartphones. Larger displays require high loads to cope with real-time systems because the system components must process more pixels. SR for a smartphone with a 4K display is the next step.

REFERENCES

(2009). Toshiba press release.

Candes, E. J. and Fernandez-Granda, C. (2014). Towards a mathematical theory of super resolution. *Communications on Pure and Applied Mathematics*, 67.6:906–956.

Dong, C., Loy, C. C., He, K., and Tang, X. (2014). Learning a deep convolutional network for image super-resolution. *Computer Vision-ECCV*, pages 184–199.

Elad, M. and Feuer, A. (1996). Super-resolution of continuous image sequence. *IEEE Trans. on Pattern Analysis and Machine Intelligence*.

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

Glasner, D., Bagon, S., , and Irani, M. (Oct. 2009). International conference on computer vision (iccv). *Super-Resolution from a Single Image*.

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*.

Matsumoto, N. and Ida, T. (2010). Reconstruction-based super-resolution using self-congruency around image edges. *IEICE transactions on information and systems (in Japanese)*, J93-D(2)(2):118–126.

Panda, S., Prasad, R., and Jena, G. (Sept. 2011). Pocs based super-resolution image reconstruction using an adaptive regularization parameter. *IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 5, No. 2, ISSN (Online), 1694-0814*.

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. (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.*

Sun, J., Sun, J., Xu, Z., and Shum, H. (2008). Image super-resolution using gradient profile prior. *CVPR 2008*, pages 1–8.

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.