Conversion of Low Resolution Video into High Resolution Video using Super-resolution in Wavelet Domain

Tanuja J Salunke[1], V. B. Baru[2]

[1]PG Student, E&TC Department, Sinhgad College of Engineering, Pune, Maharashtra, India

[2]Associate Professor, E&TC Department, Sinhgad College of Engineering, Pune, Maharashtra, India

Abstract: ***Super-resolution is a class of techniques that enhance the resolution of an imaging system. The technique of image super-resolution (SR) has been extensively studied to solve the problem of limited resolution in imaging devices since many years. In this paper, image super-resolution method in wavelet domain is applied for conversion of low resolution video into high resolution video. The proposed resolution enhancement technique uses spatio-temporal wavelet transform to decompose the input image into different sub bands. The peak signal-to-noise ratio (PSNR) and visual results of the image show the superiority of the proposed method over the conventional image resolution enhancement techniques.***

Keywords: Super-resolution; Spatio-temporal wavelet transform; Interpolation; PSNR.

1. Introduction

Super-resolution of image is the most widely used and extensive area of research. The resolution is referred as an important aspect of image. The problem of limited resolution by image acquisition devices can be solved by super-resolution. Digital cinema has up to 4K horizontal pixels, which is specified in the digital cinema initiatives. The ultrahigh definition television (UHDTV) has 8K horizontal pixels, which is proposed by NHK Science & Technology Research Laboratories. If digital cinema content is displayed by the UHDTV systems such as a public viewing of sports event or future TV broadcast, it is necessary to convert a video format. Hence, video conversion method from digital cinema to UHDTV is studied.

Interpolation-based super resolution has been used for a long time, and many interpolation techniques have been developed to increase the quality of this task. There are three well-known interpolation techniques; namely, nearest neighbor interpolation, bilinear interpolation, and bicubic interpolation [3].

The spatial resolution of an image is limited by the imaging sensors or the imaging acquisition device. The sensor size or equivalently the number of sensor elements device per unit area in the image acquisition determines the spatial resolution of the image to capture. Higher the density of the sensors, higher is the spatial resolution of the imaging system. The increase in the sensor density will eventually increase the spatial resolution of an imaging system. Thus, it results to an increase in the shot noise and hardware cost of the sensor also increases with the increase in sensor density [2].

Image resolution enhancement in the wavelet domain is a relatively new research addition. Recently, many new algorithms have been proposed. Proposed methodology in this paper is very reliable. It also shows a greater overall success comparative with other methods.

The Paper is organized as follows- Section 2 explains an overview of the different state-of-art image resolution enhancement techniques used for comparison purposes in this paper. Section 3 gives an introduction to the proposed wavelet domain based image resolution enhancement technique. Section 4 discusses the qualitative results of the proposed method with the conventional and state-of-art image resolution enhancement methods. The conclusion is given in the further section.

2. Resolution Enhancement Methods

Resolution is very important factor in images. The resolution of images becomes of vital importance because increasing the resolution of the images will eventually affect the performance of the system. The field of image resolution enhancement is mainly developed from interpolation techniques. There are different interpolation techniques devised for the better interpolation of data. Most widely used interpolation techniques in spatial domain are nearest neighbor interpolation, bilinear interpolation and bicubic interpolation. In case of nearest neighbor interpolation, the block uses the value of nearby translated pixel values for the output pixel values. In case of bilinear interpolation, the weighted average of two translated pixel values is used for each output pixel value while for bicubic interpolation the weighted average of four translated pixel values is used for each output pixel value. Compared to bilinear and nearest neighbor interpolation, bicubic interpolation gives better result of image. Thus, in this paper, the proposed method results are compared to the results of bicubic interpolation.

3. Proposed Methodology

In this paper, Spatio-temporal wavelet transform is used to decompose the input video into high and low frequency subbands. The block diagram of proposed method is shown in figure 1. Hereinafter, details of block diagram of our proposed method are explained.

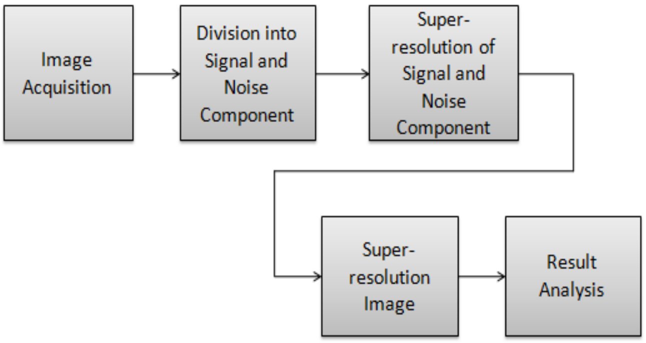

Figure 1: Block Diagram of Proposed method [1].

3.1 Division into signal and noise components

The original images are divided into signal and noise components for further processing.

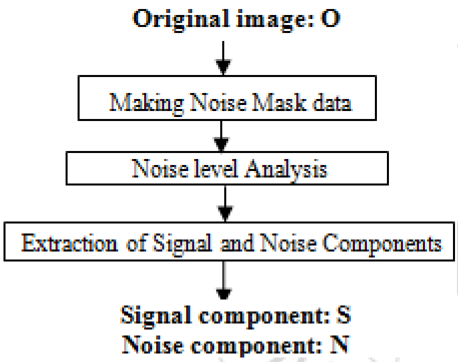

Figure 2: Division into signal and noise Component

In the "making noise mask data" block, original images O(t) are decomposed by spatio-temporal wavelet decomposition without decimation filtering. Then an isolated component in spatio-temporal high frequency H^1HH^1 band is taken and it is regarded as white noise component because it has spatio-temporally low correlation. Therefore a noise mask image M is extracted by binarization of coefficients in $|H^1HH^1|$ high frequency band using threshold value which is a median value of nonzero coefficients in $|H^1HH^1|$ band.

In the "noise level analysis" block, noise levels LVN for arbitrary band are extracted by calculating median values of nonzero coefficients in particular band as follows. For example, a noise level LV_N (H^1HH^1) is *median* $(|H^1HH^1 \cap M|)$ where *median* is the median function.

In the "extraction of signal and noise components" block, noise components of each bands are calculated by *max* $(|^{*kt*ks}(i, j) \cap M(i, j)|$, LVN $(^{*kt*ks}))$ except $L^{kt}LL^{kn}$ band where *max* is the max function and (i, j) is the coordinates of horizontal and vertical axes in ^{*kt*ks} band. Signal components of each $^{*kt*ks}S$ bands are residual components except ^{*kt*ks} N in each *kt*ks bands. Finally, signal and noise components, S and N, are extracted by (Kt, Ks)-level spatio-temporal wavelet reconstruction of S and N bands.

3.2 Super-resolution of signal component

After the separation of signal and noise components, these components are separately processed. Firstly, the signal component is super-resolved. Fig. 3 is details of the "super-resolution of signal component" block in Fig. 1.

In the "setting of super-resolved components of S" block, LL_S^1, LH_S^1, HL_S^1, and HH_S^1 are extracted by 1-levelspatial wavelet decomposition of the signal component S without decimation filtering. Then LH_S^1, HL_S^1, and HH_S^1 are set to horizontal, vertical, and diagonal high-frequency components of S.

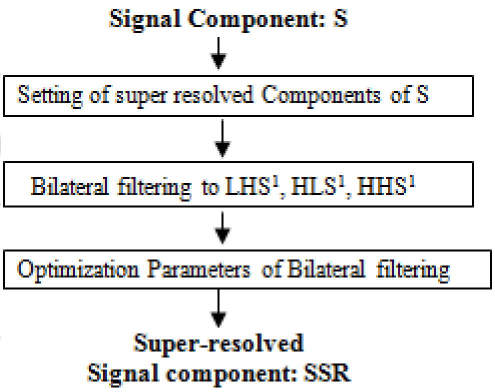

Figure 3: Super-resolution of Signal component

In the "bilateral filtering to LH_S^1, HL_S^1, HH_S^1" block, LH_S^1 is convolved by the bilateral filtering using Gaussian function with standard deviation σLH and gain γLH, and the convolved $LH_S^{1'}$ is output. $HL_S^{1'}$ and $HH_S^{1'}$ are processed similarly. Then super-resolved signal component SSR is generated by 1-level spatial wavelet reconstruction using S, $LH_S^{1'}$, $HL_S^{1'}$, and $HH_S^{1'}$.

In the "optimization parameters of bilateral filtering" block, σLH, σHL, σHH and γLH, γHL, γHH are optimized by searching the smallest differences between signal component S and four reduced size images of SSR. In the optimization, SSR is decomposed by 1-level spatial wavelet decomposition without decimation filtering at first. Next, four reduced size images are extracted by applying 4:1 pixel resampling with 1-pixel phase shift for horizontal, vertical, and diagonal directions respectively. Next, a DI-PSNR value is calculated by taking an average value of four PSNR values which are differences between S and the four reduced size images. Finally, all processes in this section are repeated using various σLH, σHL, σHH and γLH, γHL, γHH values. The super-resolution for the signal component is performed by optimum σLH, σHL, σHH and γLH, γHL, γHH values with maximum DI-PSNR values.

3.3 Super-resolution of noise component and synthesis of super-resolved signal and noise components

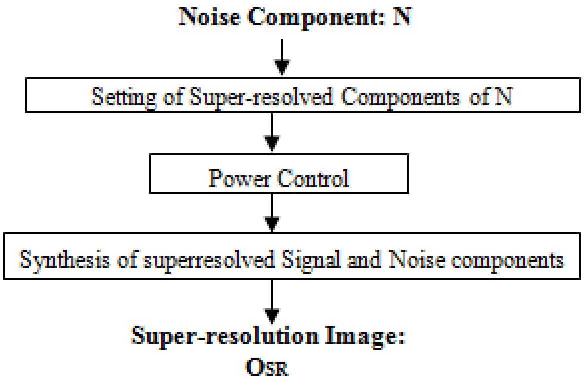

Figure 4: Super-resolution of noise component block and synthesis of super-resolved signal and noise components

Fig. 4 is details of the "super-resolution of noise component" block and "synthesis of super-resolved signal and noise components" .

In the "setting of super-resolved components of N" block, N is set to horizontal, vertical, and diagonal high frequency components of N itself because the white noise has low correlation for spatio-temporal direction. Then super-resolved noise component NR is reconstructed by 1-level spatial wavelet reconstruction using N, N, N, and N.

In the "power control" block, NR is normalized for each band by considering LVN (*1*ks). In this block, NR is decomposed by Ks-level spatial wavelet decomposition at first. Next, *ks bands are normalized by LVN (*1*ks). Finally, Super-resolved noise component NSR is reconstructed by Ks level spatial wavelet reconstruction using normalized *ks bands.

This low frequency subband image contains less information than the original image. Hence, instead of using the low frequency subband image, in the proposed method, original input image is used for interpolation process. The high frequency subband images are interpolated with interpolation factor α while the input image is interpolated with interpolation factor α/2. The proposed method uses bicubic interpolation as it gives better results than the other interpolation techniques.

The block diagram for the proposed method is shown in figure 2. The input image is interpolated by using interpolation factor α and the high frequency subbands HH, HL and LH are interpolated with interpolation factor α/2. After the interpolation process, inverse discrete wavelet transform (IDWT) is applied on the image. Thus, the output image will contain sharper edges than the interpolated image obtained by interpolation of the original input image directly. This is because the interpolation of isolated high-frequency components in HH, HL, and LH will preserve more high-frequency components after the interpolation of the respective subbands separately than interpolating input image directly.

The IDWT of the input image and the three high frequency subband images gives the super-resolved image. The proposed technique can be summarized as follows:

Proposed technique interpolates the input image as well as the high-frequency subband images obtained through the DWT process. The final high-resolution output image is generated by using the inverse DWT of the interpolated subband images and the input image. In this paper, the employed interpolation method is the same for all subband images and the input images. The interpolation technique and the wavelet function are two important factors in determining the quality of the super-resolved images.

The visual and PSNR results in the proceeding section show that the proposed technique with bicubic interpolation out performs conventional and state-of-art techniques included as references in this paper.

4. Results and Discussions

In this paper, the resolution of video is doubled. The original video is of resolution 240*320. This video is down sampled by factor 2. The proposed technique is applied on the video of resolution 120*160. and the resolution is increased to 240*320. Figure 5 shows the original downsampled video. Figure 6 (a) shows the video obtained by applying proposed method while Figure 6 (b) is the video obtained after bicubic interpolation.

Figure 5: Original Video

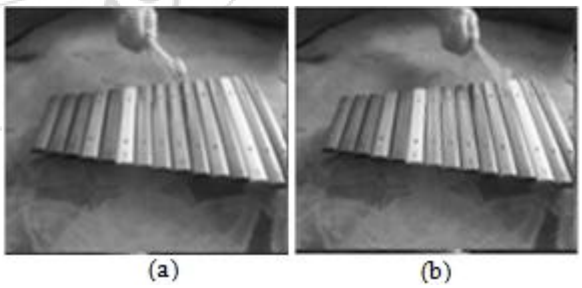

Figure 6: Super-resolved video using (a) proposed method (b) Bicubic interpolation

Table 1: PSNR values of the video for bicubic interpolation and proposed method.

Method	*PSNR (dB)*
Bicubic Interpolation	18.5
Proposed Method	22.2

The result is observed on the basis of Peak Signal to Noise Ratio (PSNR). Table 1 shows the values for the video in Figure 6. Thus, from the table, it can be seen that the proposed method gives more PSNR than that of bicubic interpolation. Also, the visual appearance of the video of proposed method is better than the video of interpolation.

5. Conclusion

The paper has proposed a method to obtain a conversion method from digital cinema to ultrahigh definition Television using Super-resolution Technique. The proposed method is much better than previous methods. The comparison of the resultant images is done based on the PSNR values. Proposed method has highest PSNR than other methods. The method involves Spatio-temporal wavelet transform. The proposed technique is used to double the resolution of input video. This gives super resolved video. The comparison of proposed method with state of the art method is shown in Table 1. Also, the obtained super-resolved video has better visual results.

References

[1] Yasutaka Matsuo, Shinya Iwasaki, Yuta Yamamura, and Jiro Katto, "Super-resolution from Digital Cinema to Ultra High Definition Television using Image Registration of Wavelet Multi-scale Components", *IEEE International Conference on Consumer Electronics*, 2013
[2] Yasutaka Matsuo, Shinya Iwasaki, Yuta Yamamura, and Jiro Katto, "Wavelet Domain Image Super-Resolution from Digital Cinema to Ultrahigh Definition Television by dividing Noise Component", *Visual Communications and Image Processing, IEEE*, 2012.
[3] Hasan Demirel, Sara Izadpanahi, and Gholamreza Anbarjafari, "Improved Motion-based localized Super-resolution Technique using Discrete Wavelet Transform for Low Resolution Video Enhancement", *European Signal Processing Conference*, 2009.
[4] Hasan Demirel, Gholamreza Anbarjafari, "Discrete Wavelet Transform-based Satellite Image Resolution Enhancement", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 49, No. 6, June 2011.
[5] Alptekin Temizel, Theo Vlachos, "Image resolution upscaling in the wavelet domain using directional cycle spinning", *Journal of Electronic Imaging*, Vol. 14, Oct–Dec 2005.
[6] A. Temizel and Theo. Vlachos, "Wavelet domain image resolution enhancement using cycle-spinning and edge modeling," *Electronics Letters*, vol. 41, no. 3, pp.119–121, Feb. 2005.
[7] O. H. Chavez-Roman, V. Ponomaryov, I. Loboda, "Resolution Enhancement Algorithm Based On Wavelet and Edge Extraction Techniques in Noise Presence", *Millimeter and Submillimeter Waves (MSMW)*, 2013.
[8] Y. Shishikui, Y. Fujita, and K. Kubota, "Super hi-vision– the star of the show!," *EBU Technical Review*, pp. 4-16, Jan. 2009.
[9] C. V. Jiji, M. V. Joshi, and S. Chaudhuri, "Single-frame image super-resolution using learned wavelet coefficients," *International Journal of Image Systems and Technology*, vol. 14, no. 3, pp. 105-112, 2004.
[10] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Total variation super resolution using a variational approach," *Proceedings of IEEE ICIP*, pp.641-644, 2008.
[11] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," *Proceedings of IEEE ICCV*, pp.839-846, 1998.
[12] Niyanta Panchal, Bhailal Limbasiya, Ankit Prajapati, "Survey On Multi-Frame Image Super-Resolution", *International Journal of Scientific & Technology Research Volume 2, Issue 11, November* 2013.
[13] C. Andrew Segall, Rafael Molina, Aggelos K. Katsaggelos and Javier Mateos, "Bayesian High-Resolution Reconstruction of Low-Resolution Compressed Video", *IEEE*, 2001.
[14] Digital Cinema Initiatives, LLC: "Digital cinema system specification, version 1.2," Mar. 2008.

Author Profile

Prof. V. B. Baru is an Associate Professor in Electronics and Telecommunication Department at Sinhgad College of Engineering, Pune. He has done BE in 1993 and completed ME in 1999 in Electronics and Telecommunications. He is pursuing Ph. D from College of Engineering, Pune. He has 20 years of Teaching Experience and published more than 50 papers in national and International level journals. He is author of two books 'Electronic Product Design' by Wiley Publication and 'Basic Electronics' by Dreamtech Publication. He has guided more than 100 UG students and about 25 PG students for their Dessertations.

Ms. Tanuja J. Salunke has done BE in Electronics and Telecommunications from Savitribai Phule Pune University, Pune. She is currently pursuing ME in Signal Processing (E&TC) from Sinhgad College of Engineering, Pune. Her areas of interest are Image Processing, Signal Processing.