IMPROVED MOTION-BASED LOCALIZED SUPER RESOLUTION TECHNIQUE USING DISCRETE WAVELET TRANSFORM FOR LOW RESOLUTION VIDEO ENHANCEMENT

*Hasan Demirel, Sara Izadpanahi, and Gholamreza Anbarjafari
Department of Electrical and Electronic Engineering, Eastern Mediterranean University, Gazimağusa, KKTC, Mersin 10, TURKEY
* phone: +(90) 392 630 1436, fax: +(90) 392 365 0240, email: hasan.demirel@emu.edu.tr
web: http://faraday.ee.emu.edu.tr/hdemirel

ABSTRACT

Super resolution is used for resolution enhancement of images or video sequences. Instead of super resolving frames globally, using localized motion based super resolution increases the quality of the enhanced frames. In this paper, we propose to use the super resolution on different subbands of localized moving regions extracted from discrete wavelet transform (DWT) and composing the super resolved subbands using inverse DWT (IDWT) to generate the respective enhanced high resolution frame. The results based on PSNR values, in comparison with the global super resolution method, show improvement in quality. The improvement is achieved by isolating different frequencies in different subbands extracted from DWT and super resolving them separately.

1. INTRODUCTION

There are several ways of performing super resolution algorithms, but most of them are variations on two main approaches. The first approach is multi-frame super resolution which is based on the combination of image information from several similar images taken from a video sequence [1]. This method consists of two main stages, firstly estimating motion parameters between two images referred as registration, and secondly projecting the low resolution image into the high resolution lattice referred as reconstruction. The second approach is called single-frame super resolution, which uses prior training data to enforce super resolution over a single low resolution input image. In this paper we are using multi frame super resolution taken from low resolution video sequences.

Tsai and Huang are the pioneers of super resolution idea [2], who used the frequency domain approach. Further work has been done by Keren, et al. [3] who described a spatial domain procedure to perform image registration using a global translation and rotation model. Lucchese and Lucchese [4] presented a method for estimating planar roto-translations that operates in the frequency domain and hence is not based on features. Another frequency domain approach was introduced by Reddy et al. [5]. Their method utilizes separately rotational and translational components of the Fourier transform similar to [4]. Irani et al. [6] has developed a motion estimation algorithm, which considers translations and rotations in spatial domain. Demirel et al. [7] introduced a method which reduced the computational cost of the multi-frame super resolution by utilizing motion-based localization and also improved the quality of the enhanced super resolved frames. In [7] the moving region has been separately super resolved by using the state-of-art methods such as the method of Irani et al. [4] and the static region has been simply interpolated using bicubic interpolation.

Image and video resolution enhancement in the wavelet domain is a relatively new research area and recently new algorithms have been proposed [8, 9].

In this work, we propose to use the super resolution on different wavelet subbands of localized moving regions and composing the super resolved subbands using inverse DWT (IDWT) to generate the respective enhanced high resolution frame. Localized moving region registration is more precise than global registration of the entire frame, due to the fact that local alignment of the changes in the local neighborhood is more efficient.

The proposed super resolution method is tested using different super resolution methods described in [5], [4], and [6] as reported in the previous work [7] to achieve enhanced high resolution video sequences.

2. STATE-OF-ART SUPER RESOLUTION METHODS

In this section, we review a brief introduction of three super resolution methods which have been used to compare the performance of the proposed super resolution technique

2.1. L. Lucchese et al. super resolution method

Lucchese et al. super resolution method operates in the frequency domain. The estimation of relative motion parameters between the reference image and each of the other input images is based on the following property: The amplitude of the Fourier transform of an image and the mirrored version of the amplitude of the Fourier transform of a rotated image have a pair of orthogonal zero-crossing lines. The angle that these lines make with the axes is identical to half the rotation angle between the two images. Thus the rotation angle will be computed by finding these two zero crossings lines. This algorithm uses a three-stage coarsest to finest procedure for rotation angle estimation with a wide range of degree accuracy. The shift is estimated afterwards using a standard phase correlation method.

2.2. Reddy et al. super resolution method

In this method a registration algorithm that uses the Fourier domain approach to align images which are translated and rotated with respect to one another, was proposed [5]. Using a log-polar transform of the magnitude of the frequency spectra, image rotation and scale can be converted into horizontal and vertical shifts. These can therefore also be estimated using a phase correlation method. Their method utilizes reparability of rotational and translational components property of the Fourier transform which is similar to [4]. According to this property, the translation only affects the phase information, whereas the rotation affects both phase and amplitude of the Fourier transform. One of the properties of the 2D Fourier Transform is that if we rotate the image, the spectrum will rotate in the same direction. Therefore, the rotational component can first be estimated. Then, after compensating for rotation, and by using phase correlation techniques, the translational component can be estimated easily.

2.3. Irani et al. super resolution method

Irani et al. [6] have developed a motion estimation algorithm. This algorithm considers translations and rotations in spatial domain. The motion parameters which are unknown can be computed from the set of approximation that can be derived from the following equation (1), where the horizontal shift a, vertical shift b, and rotation angle θ between two images g_1 and g_2 can be expressed as:

$$g_2(x,y) = g_1(x\cos\theta - y\sin\theta + a,\, y\cos\theta + x\sin\theta + b) \quad (1)$$

Finally, after determining and applying the results, the error measure between images g1 and g2 is approximated by (1) where this summation is counted over overlapping areas of both images.

$$E(a,b,\theta) = \sum \left[g_1(x,y) + \left(a - y\theta - \frac{x\theta^2}{2} \right) \frac{\partial g_1}{\partial x} + \left(a + x\theta - \frac{y\theta^2}{2} \right) \frac{\partial g_1}{\partial y} - g_2(x,y) \right]^2 \quad (2)$$

For reducing E to its minimal value and obtaining more accurate result, the linear system in (5) is applied. By solving the following matrix, the horizontal shift a, vertical shift b, and rotation angle θ will be computed as follows.

$$M = \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \quad B = \begin{bmatrix} \sum \frac{\partial g_1}{\partial x}(g_1 - g_2) \\ \sum \frac{\partial g_1}{\partial y}(g_1 - g_2) \\ \sum R(g1 - g2) \end{bmatrix} \quad (3)$$

$$A = \begin{bmatrix} \sum \left(\frac{\partial g_1}{\partial x} \right)^2 & \sum \left(\frac{\partial g_1}{\partial x} \frac{\partial g_1}{\partial y} \right) & \sum \left(R \frac{\partial g_1}{\partial x} \right) \\ \sum \left(\frac{\partial g_1}{\partial x} \frac{\partial g_1}{\partial y} \right) & \sum \left(\frac{\partial g_1}{\partial y} \right)^2 & \sum \left(R \frac{\partial g_1}{\partial y} \right) \\ \sum \left(R \frac{\partial g_1}{\partial x} \right) & \sum \left(R \frac{\partial g_1}{\partial y} \right) & \sum R^2 \end{bmatrix}_{3\times 3} \quad (4)$$

$$AM = B \Rightarrow A^{-1}AM = A^{-1}B \Rightarrow M = A^{-1}B \quad (5)$$

Fig. 1(a-d) shows the four low resolution consecutive frames, where (e), (f) and (g) shows super resolved high resolution images by using Lucchese et al. [4], Reddy et al. [5] and Irani et al. [6] methods respectively.

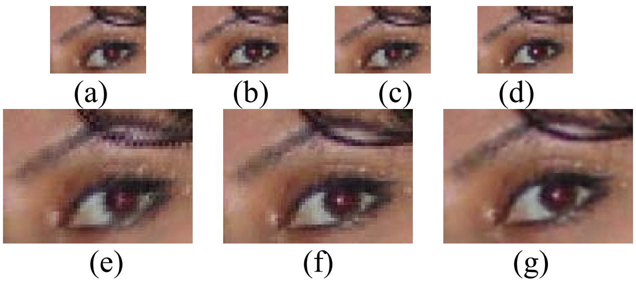

Figure 1 - (a) through (d) are four low resolution images taken from a video sequence. (e) High resolution image of (a) through (d) using Lucchese et al. (f), Reddy et al. (g) and Irani et al. methods.

3. PROPOSED DWT BASED MULTI-FRAME SUPER RESOLUTION METHOD

The main loss of an image or a video frame after being super resolved is on its high frequency components (i.e. edges), which is due to the smoothing caused by super resolution techniques. Hence, in order to increase the quality of the super resolved image, preserving the edges is essential. Hence, in this work, DWT [10] has been employed in order to preserve the high frequency components of the image by decomposing a frame into different subband images, namely Low-Low (LL), Low-High (LH), High-Low (HL), and High-High (HH).

LH, HL, and HH subband images contain the high frequency components of the input frame. The DWT process for each frame of the input video generates 4 video sequences in each subband (i.e. LL, LH, HL and HH video sequences). Then, the Irani et al. super resolution method in [6] is applied to all subband video sequences separately. This process results in 4 super resolved subband video sequences. Finally, IDWT is used to combine the super resolved subbands to produce the high resolution video sequence.

By super resolving the LL, LH, HL and HH video sequences and then by applying IDWT, the output video sequence would contain sharper edges than the super resolved video sequence obtained by any of the aforementioned super resolution techniques directly. This is due to the fact that, the super resolution of isolated high frequency components in HH, HL and LH will preserve more high frequency components after the super resolution of the respective subbands separately than super resolving the low resolution image directly.

In this work, the moving regions are extracted as in [7] to be super resolved with the proposed super resolution technique explained above. The static regions are similarly transformed into wavelet domain and each static subband sequence is interpolated by bicubic interpolation. The high resolution sequence of the static region is generated by composing the interpolated frames using the IDWT. Eventually, the super resolved sequence is achieved by combining the super resolved moving sequence and the interpolated static region sequence. The proposed method can be summarized with the following steps:

1. Acquire frames from video and extract motion region(s) using frame subtraction.
2. Determine the significant local motion region(s) by applying connected component labeling.
3. Apply DWT to decompose the static background region into different subbands.
4. Apply bicubic interpolation for enhancing resolution of each subband obtained from step 3.
5. Use IDWT to reconstruct the super resolved static background.
6. Apply DWT to decompose the moving foreground region(s) into different subbands.
7. Super resolve the extracted subbands by applying Irani et al. [6] super resolution method.
8. Use IDWT to reconstruct the super resolved moving region(s).
9. Combine the sequences obtained from steps (5) and (8) to generate the final super resolved vide sequence.

In the first step, four consecutive frames are used where each frame is subtracted from the reference frame so the differences between frames are extracted. After applying OR operation for all subtracted images local motion(s) will appear.

In the second step, the area of local motion(s) can be determined by using connected component labeling. In the third, fourth, and fifth steps the rest of the frames which does not include any motion and it is static, will be decomposed by DWT, interpolated with the help of bicubic interpolation, and composed by IDWT.

Fig. 2 shows four consecutive frames taken from a video sequence. The second frame is used as the reference frame. The rectangular part shown in each frame corresponds to the moving part. The rest of the reference frame is the static part. In every four frame the rectangular moving part will change according to the moving part in those frames.

In the sixth, seventh and eighth steps, motion parts will be decomposed into different subbands by DWT, super resolved by using Irani et al. super resolution technique, and all subbands will be composed by IDWT.

In the final step, we combine super resolved motion frames with the interpolated background to achieve the final high resolution video sequence. The algorithm is shown in Fig. 3.

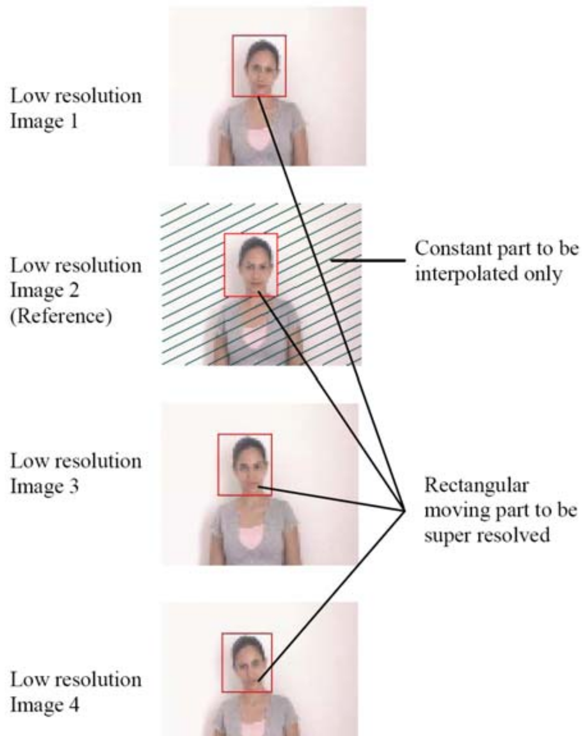

Figure 2 - First four consecutive frames taken from a video sequence with one moving region. The rectangular moving part for each four frames is changing adaptively.

5. RESULTS AND DISCUSSIONS

Super resolution method proposed in this paper is compared with the state-of-art super resolution techniques in section 2. Two low resolution test video sequences with 4 second recordings (30 fps) are used to evaluate the performance of the proposed super resolution method. One of the sequences contains a single moving object, where the second one contains two moving objects. The frames are of size 256x256 pixels. The motion regions change adaptively for each frame in comparison of next two frames and the previous frame.

The low resolution video sequences are generated by downsampling and lowpass filtering each frame of the high resolution video sequence. In this way we keep the original high resolution video sequences for comparison purposes. The PSNR of each frame is calculated according to the mean squared error (MSE) calculated between the super resolved low resolution image and the original high resolution image.

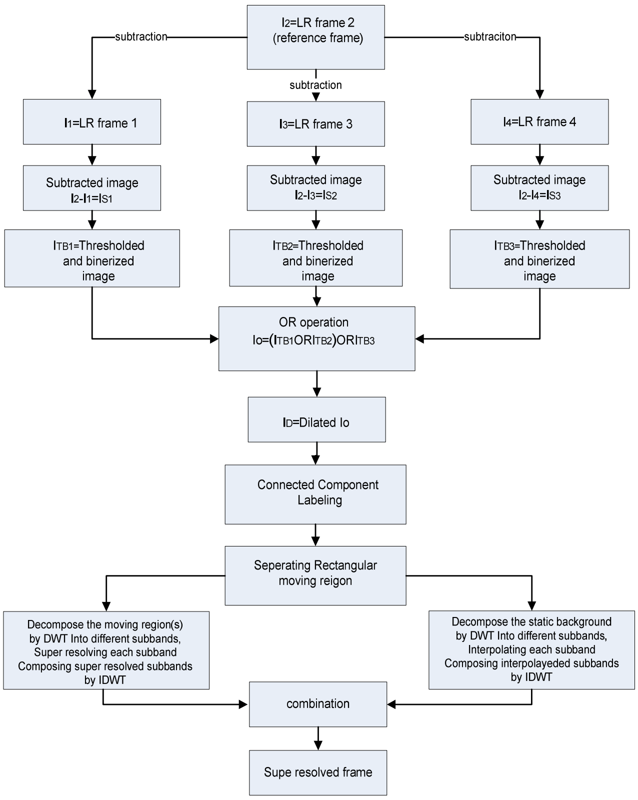

Figure 3 - The algorithm of the proposed super resolution method for video enhancement.

Table 1 and Table 2 show the result of calculated average PSNR of two different video sequences containing one and two moving object.

The proposed method generates higher PSNR than the state-of-art super resolution methods. Localized motion based super resolution method typically considers a much smaller rectangular region than the entire frame [7]. Smaller regions with local variations can be registered with higher accuracy than registering the same local variations using the entire global region. The results in Table 1 and 2 confirm the stated fact. The results show the superiority of the proposed localized super resolution method in PSNR based frame quality.

Table 1: Average PSNR of a video sequence with one moving object.

Methods used to enhance the resolution	PSNR of super resolved motion part	PSNR of global super resolution
Proposed Method	**34.19**	**33.17**
Irani et al. [6]	31.66	31.52
Reddy et al. [5]	30.68	30.36
Lucchese et al. [4]	30.07	29.86
Bicubic interpolation	29.65	29.62

Table 2: Average PSNR of a video sequence with two moving object.

Methods used to enhance the resolution	PSNR of super resolved motion part	PSNR of global super resolution
Proposed Method	**35.68**	**34.09**
Irani et.Al [6]	32.98	32.78
Reddy et.Al [5]	32.01	31.89
Lucchese et.Al [4]	31.97	31.89
Bicubic interpolation	30.42	30.29

6. CONCLUSION

In this paper, we propose to use the super resolution on different subbands of localized moving regions extracted from DWT and composing the super resolved subbands using IDWT to generate the respective enhanced high resolution frame. The results based on PSNR values show that the proposed method, in comparison with the global super resolution method, show improvement in quality. Furthermore, the performance of the proposed super resolution method outperforms the state-of-art super resolution methods. The improvement is achieved by isolating different frequencies in different subbands extracted from DWT and super resolving them separately.

7. REFRENCES

[1] Michael Elad, Arie Feuer, “Super-resolution reconstruction of image sequences”, *IEEE Trans. On Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 21, No. 9, September 1999.

[2] R. Y. Tsai and T. S. Huang, “Multiframe image restoration and registration”, *in Advances in Computer*

Vision and Image Processing, T. S. Huang, Ed. JAI Press, Vol. 1, 1984, pp. 317–339.

[3] D. Keren, S. Peleg and R. Brada, "Image sequence enhancement using subpixel displacements", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 1988, pp. 742-746.

[4] L. Cortelazzo and G. M. Lucchese , "A noise-robust frequency domain technique for estimating planar roto translations", *IEEE Transactions on Signal Processing*, Vol. 48, No. 6, June 2000, pp. 1769–1786.

[5] B. S. Reddy and B. N. Chatterji, "An fft-based technique for translation, rotation and scale-invariant image registration", *IEEE Transactions on Image Processing*, Vol. 5, No. 8, August 1996, pp. 1266–1271.

[6] M. Irani and S. Peleg, "Improving resolution by image registration", *CVGIP: Graphical Models and Image Processing*, Vol. 53, No. 3, May 1991, pp. 231–239.

[7] Hasan Demirel, Sara Izadpanahi, "Motion-Based Localized Super Resolution Technique for Low Resolution Video Enhancement", *16th European Signal Processing Conference (EUSIPCO-2008),* Lausanne, Switzerland, Aug. 2008.

[8] Y. Piao, l. Shin, and H. W. Park, "Image Resolution Enhancement using Inter-Subband Correlation in Wavelet Domain"**,** *International Conference on Image Processing (ICIP 2007),* Volume 1, pp.: I- 445-448.

[9] W. K. Carey, D.B. Chuang, and S.S. Hemami, "Regularity-preserving image interpolation", *IEEE Trans. Image Proc.*, vol.8, no.9, pp.1295- 1297, Sep. 1999.

[10] S. Mallat, "*A wavelet tour of signal processing*", Published by Academic Press, 1999, 2nd edition ISBN 012466606X, 9780124666061

[11] A. Temizel, **"**Image resolution enhancement using wavelet domain hidden Markov tree and coefficient sign estimation", *International Conference on Image Processing (ICIP2007),* Vol. 5, pp.: V-381-384.