VIDEO SUPER RESOLUTION RECONSTRUCTION FROM LOW RESOLUTION IMAGES USING SPLINE INTERPOLATION

1V. VENUBABU, 2M. V. RAGHUNADH

Department of E and CE, NIT Warangal, India
Email: venu.vemparala@gmail.com, raghunadhmv@gmail.com

Abstract—Transmitting a real time video for Military, Satellite, Cellular and Surveillance applications over a low bandwidth channel is a challenging problem. Due to limited channel bandwidth, the bit rate of high resolution image/video is decreased by decimation (down sampling) and transmitted through the channel, resulting in a low bit rate compressed image. At the output(receiving) side, the high resolution image/video is reconstructed from the set of compressed low resolution noisy and blurred images using a technique known as super resolution. In this paper we explain the problem of reconstructing a high resolution image/video using several low bit rate compressed (low resolution) noisy and blurred images is considered. We propose an efficient method for reconstruction of a high resolution image using a set of low resolution images, Spline interpolation and De-blurring. The performance of the proposed method is verified using objective and subjective quality measures for an image. Based on the experimental results, the proposed algorithm provides a better PSNR (Peak Signal to Noise Ratio), SSIM (structural similarity index measure) and SNR (Signal to Noise Ratio).

Keyword: Motion estimation, spline interpolation and super resolution (SR).

I. INTRODUCTION

Super Resolution Reconstruction (SRR) is one of the emerging technology used in Military, Satellite and Cellular applications. Due to low processing power of a system and low bandwidth capability of the channel, it is difficult to transmit a high resolution image through the channel. For efficient use of a low bandwidth channel compressed version of a high resolution images sent through the channel and at the receiving side the high resolution (HR) image can be reconstructed using a sequence of compressed low resolution (LR) images. Super Resolution (SR) is a process of increasing the number of pixels in an image for more details using a set of low resolution images. The spatial distance between the pixel decreases as number of pixels increases in an image using a low resolution images and possible to preserve a high frequency components such as edges and completely removing the degradations arises from the common imaging system [1] via down sampling, blur and noise. Each low resolution image captures a small amount of information from the scene; The main goal of super resolution is construction of a high resolution image from a sequence of low resolutions images by extracting an unique information from each low resolution image. They are different methods for constructing the HR image, these methods are based on multiple low resolution images and a single low resolution image. Almost all the SR algorithms are based on the reconstruction of multiple low resolution images using the sub-pixel motion shifts and nonuniform to uniform sampling techniques.

The rest of the paper is organized as follows: The detailed literature survey is discussed in section II, mathematical formation model of super resolution is discussed in section III, and in section IV the proposed method is explained in detail. And then, the results and conclusions were discussed in section V, VI respectively.

II. LITERATURE SURVEY

To slove the problem of limited resolution devices the super resolution was studied from decades [2]. The main goal of super resolution is to reconstruct a high resolution image from a set of compressed low resolution images. They are different methods to construct a high resolution image and they are mainly classified as frequency domain approaches [3], learning based techniques [4], iterative reconstruction and interpolation based techniques [5], [6], [7] and dynamic tree and wavelet based resolution techniques [8]. Based on literature work [2], [9] considers the different super resolution techniques based on filtering approach in spatial domain. Early works on super resolution considers [10] transformation of non-uniform samples into uniform samples using the concept of the sampling theorem [11] with out loss of data by satisfying the Nyquist theorem. Some of the super resolution techniques are considered as frequency domain techniques for registering the images using the properties of the discrete Fourier transform [3] and this approach only considers the global rotation of the image and it is sensitive to noise. In [12] this an algorithm is proposed for effectivly eliminating the blurring effect caused by sensor using the POCS (Projection of convex sets). This algorithm need an early knowledge of process constructions, it is slow convergence and the accuracy of HR image depends on initial guess of the algorithm. An Iterative Back Projection algorithm is proposed in [13] it iteratively decreases the reconstruction error between a low resolution image.
and a high resolution image by converting a high resolution into a low resolution image using back projection. The constraints and prior knowledge of IBP (Iterative Back Projection) techniques is very difficult compared to [12].

Dynamic tree structure model SR (super resolution) is proposed in [14] with a mean field approach. In this algorithm the high resolution images is a combination of series of patches with different shapes. Due to different nature of priori and measured image this algorithm is producing erroneous results for some conditions. Learning based super resolution is proposed in [8] using an Adaptive Regularization. In this the high resolution image was reconstructed using a single low resolution image with pair matching. This algorithm requires large collection of database and storage space. [15] proposed an algorithm for video super resolution using the quality coefficients for an accurate motion estimation and number of LR frames required of this algorithm is 8 to 16. Super resolution using wavelet techniques is require more memory, this memory requirement can be decreased by using wavelet lifting schemes proposed in [8] using SPHIT. In this algorithm FFT based image registration technique is used and this method is sensitive to noise and it is work for only global rotation. Based on Kernel Regression and spatial temporal orientation a super resolution algorithm is proposed in [16] according to this method kernel regression is best way to use irregular interpolation and this method is causes a loss of edges and boarder information. Super resolution reconstruction is an inverse problem is effectively eliminated by regularization of an image is proposed in [17] with bilateral total variance. This algorithm suffers due to smooth output and loss of edges.

Using the higher order motion prediction and bilateral filtering a new super resolution algorithm is proposed in [18]. According this algorithm a higher order prediction is used to find the sub pixel motion with structural similarity features. This algorithm is not suitable for aliased images due to loss of structure edges. The variable block size motion estimation algorithm for super resolution is proposed in [19] for decreasing the local and global error of an image up to sub pixel variations. The global level registration is present between a set of high resolution images. A multi-resolution analysis based super resolution approach is proposed in [20]. Multi-resolution analysis is used to approximate edges with less coefficients compared to wavelet transform and this method suffers from an unwanted frequency components due to interpolation. The high resolution image constructed from correlation function and non local means is proposed in [5]. This is invariant to blur and resolution factor and it is an iterative algorithm. The image reconstruction algorithm using spline interpolation is proposed in [6], [7] using single image. Image interpolation using single image basically produce smooth results and loss of high frequency components such as edges and human eye is more sensitive edges. The proposed algorithm reconstruct a high resolution image using a set of sub-pixel shifted low resolution images and each sub-pixel shifted image contains different information after combining each low resolution image the high resolution image provides more details than any of the interpolated single image. The high resolution image can be constructed using multiple low resolution images and spline interpolation. The proposed method efficiently eliminating the noise and smoothness of image by integrating a de-blurring into the spline interpolation.

III. MATHEMATICAL FORMATION MODEL FOR SUPER RESOLUTION

In this section, we give a higher order restoration model for super resolution using a sequence of low resolution images. Low resolution image means degraded, blurred and noisy version of a high resolution image and constructing a high resolution (HR) image from set of low resolution images is an inverse problem. Due to common imaging system [1] the spatial distance between the pixels is increasing causes the loss of high frequency components and image looks a degraded image with less details and pixel density.

Let us consider a down sampling factors w1 in horizontal and w2 in vertical direction and a low resolution (LR) image of size p1 x p2, denoted as L=[x1; x2;::;xN], where N=p1 x p2. The high resolution (HR) image size by considering the down sampling factors is w1p1 x w2p2, denoted as Z=[y1; y2;::;yK], where K=w1p1 x w2p2. By taking the high resolution image Z (s; t) and down sampling factor the low resolution image L (i; j) can be processed as

$$L(i,j) = \frac{1}{w_1w_2} \sum_{u_1}^{(w_1+1)-1} \sum_{u_2}^{(w_2+1)-1} Z(s,t)$$  \hspace{1cm} (1)

Based on down sampling factor the low resolution image is an average of pixels of window size w1 x w2 in a high resolution image. According to the process of framework let us take g compressed, blurred low resolution images [Lm]m=1;::;g of size p1 x p2 for constructing a high resolution image of size q1 x q2, where q1=w1p1, q2=w2p2. From the concept of basic imaging system [1] the low resolution image is a blurred and noise version of a high resolution image. By including the blur matrix, noise and transmission matrix the high order restoration model can be expressed as

$$L_f = H_{p_{sf}}M_fD_fZ + \eta_f$$  \hspace{1cm} (2)

This equation can also be expressed as

$$L_f = H_fZ + \eta_f$$  \hspace{1cm} (3)

Where $H_f$ = $H_{p_{sf}}M_fD_f$ and $H_{p_{sf}}$ is point spread function of optical lens of common imaging system and the size of $H_{p_{sf}}$ is $p_1p_2$ x $p_1p_2$. $M_f$ is transmission matrix of size $w_1p_1w_2p_2$ x $w_1p_1w_2p_2$. $D_f$ is decimation matrix of size $(p_1p_2)^2$ x $w_1p_1w_2p_2$ and $\eta_f$ is noise vector of size $p_1p_2$.
IV. PROPOSED METHOD FOR SUPER RESOLUTION

In this paper, we implement an algorithm for reconstruction of a high resolution image using spline interpolation and a set of low resolution images. The proposed algorithm mainly divided into three blocks. The first block is motion estimation and compensation (ME), the second one is spline interpolation and final one is de-blurring (or) sharpening filter. The adaptive full search block matching algorithm is used for motion estimation (ME) and motion compensation. Finally, HR image is reconstructed with cubic spline interpolation and de-blurring filter and the block diagram of proposed model as shown in the Fig 1. According to proposed model the working procedure can be discussed as follows.

A. Full search motion estimation

Full search motion estimation algorithm [21] is used to compensate the local motion between two consecutive frames. The motion vectors of a current frame corresponding to the reference frame can be calculated using the MAD (Mean Absolute Difference) is shown in equation 4.

$$MAD = \frac{1}{mn} \sum_{m=1}^{m-1} \sum_{n=1}^{n-1} |r(x, y) - c(x, y)|$$  \hspace{1cm} (4)

Where \( r(x, y) \) is a reference frame and \( c(x, y) \) is a current frame. Finding a motion vectors using a full search algorithm is computational inefficient. In order to increase the performance of the algorithm and decreasing the number of computations, a full search algorithm is applied on Gray scale image and corresponding motion vectors are used on three R, G and B channels to compensate the sub-pixel motion between two consecutive frames. The number of computations further decreases by using threshold value. This threshold value can be calculated based on the estimation of the maximum motion between two consecutive frames. According to full search [21] the MAD value is compared with threshold value and if MAD value is more than the threshold value the block is moved to next consecutive block. Due to threshold value some local error are occur in compensated image and these local error can be removed by finding the median image from the four compensated images. Median image is a combination of all four frames and this image is formed by finding a median of four pixel values at same location.

B. Cubic spline interpolation and deblurring

Spline interpolation is a polynomial based interpolation in that each point on the polynomial is connected smoothly. We need a \( n+1 \) coefficients for each piece of polynomial of a degree \( n \) and first order spline is approximately equal to a linear interpolation. In a cubic interpolation the continuous and smooth spline curves [22], [23] are obtained from a series of unique polynomials between two data points. Cubic spline are used to determine the rate of change and cumulative changes over an interval. Cubic spline interpolation is used to interpolate the equally spaced and non-equally spaced data points. The mathematical spline is similar in principle. Numerical data are the points in this case. In this cubic polynomials weights are the coefficients used to interpolate the data. These coefficients are used to bend the line so that it passes through each of the data points without any breaks in continuity.

The main aim of the cubic spline interpolation is forming the interpolation formula at its boundaries with in the interval [24]. The interpolation is continuous in the second derivative and smooth in the first derivative. The cubic interpolation in the interval \( x_j \) and \( x_{j+1} \) using the tabulated function \( R = R(x_j), i = 1 \ldots N \) is given in equation 5.

$$R = ER_j + FR_j + GR_j + AR_j + A'$$

where \( R_j' \) is second derivative of the function and 

$$E = \sum_{j=1}^{n} F_j = 1 - E \quad , \quad G = \frac{1}{h} \left( E^2 - E \right) (x_{j+1} - x_j)^2$$

and \( A = \frac{1}{h^3} (F_j - F_{j+1}) (x_{j+1} - x_j)^2 \)

The high resolution image constructed from a set low resolution images is shown in fig 1. Based on the method in fig 1 the motion compensated images obtained using full search motion estimation with threshold and find the median image using a three compensated frames and one original frame for decreasing the local error observed in a motion compensated images. Apply a cubic spline interpolation on median image to obtain a high resolution image. Due to spline interpolation the high resolution image loses some of the high frequency components. To recover the edges de-blurring or sharpening technique (unsharpened masking) is integrating with Spline Interpolation.

V. RESULTS AND DISCUSSIONS
In this section, we will evaluate the performance of the proposed method on different test images football, coastguard, tennis and garden using OpenCV. Fig 2 shows typical video sequence frames used to evaluate the performance of our work. The whole football sequence consists of a dynamic foreground and variable background and it is best suited video frame for complex analysis. Static background and slow local motion are present in coastguard sequence. Garden sequence has a high movement on complex background. In tennis sequence the human hands are moving away from the body, it contains rigid and non-rigid objects. The proposed algorithm is compared with four other method namely Nearest neighbourhood interpolation, Bilinear interpolation, Bicubic interpolation and finally Y Chen [17] method. For the purpose of testing we generated a low resolution frames by down sampling the high resolution image by two times. The performance of this algorithms are compared using image quality assessment techniques such as subjective and objective quality measures of an image. Human eye is more sensitive to sharp change in the intensity values in an image and it is related to subjective quality of the image. To show the performance of the proposed method from other methods from subjective point of view fig 3 is included. Fig 3 shows that image in (f) reconstructed by using the proposed technique is approximately same as input high resolution image (a) as well as image reconstructed using nearest neighbourhood interpolation (b), linear interpolation (c), cubic interpolation(d) and method proposed in [17] (e). It clearly shows that proposed method gives a sharper image. The subjective quality measure is based on human dependent and it changes independtly, the objective quality measures such as Peak Signal to Noise Ratio (PSNR), Mean Average Error (MAE), Mean Square Error (MSE), Structural Similarity Index Measure (SSIM) and Signal to Noise Ratio (SNR) are used to measure the quantitative information of an image. These equations are defined in appendix by equations 6 to 11. The performance of proposed algorithm is superior compared with other interpolation and Y Chen [17] methods based on experimental results. The objective quality measures shows that proposed method outperforms all the other methods. The method proposed in Y Chen [17] and proposed method shows approximately equal results for tennis and garden sequences. Table I shows that proposed method has a higher PSNR compared to all other method and the higher PSNR will indicates a good quality image. Most of the time human visual system is to observe structural information of the viewing scene and it is highly adapted for this purpose. In order to know the perceived image distortions we need to measure structural distortions of an image. Table II compares structural similarity index measure using original frame and reconstructed frame. Higher the SSIM indicates a better image quality with less structural distortions. Table III compares the signal to noise ratio of original and reconstructed image for all test sequences. Higher the SNR indicates the better image quality. Table IV compares the mean square error between reconstructed image and original image for all test cases. Lower the MSE indicates better quality image. The comparison of root mean square (RMS) between original and reconstructed image shown in Table V and finally Table VI compares the mean average error of original image and reconstructed image. Lower value of MAE shows image has better quality. The proposed algorithm shows a better results for video sequences by observing the PSNR comparison of different methods for football sequence is shown in fig 4. This fig 4 indicates the variation of PSNR along the number of frames for different methods. Based on fig 4 the proposed method is very effective along the video frames. However proposed method is good for restoring a human eye preserving edges.
TABLE I
PSNR: COMPARISON OF DIFFERENT SUPER RESOLUTION METHODS (dB)

<table>
<thead>
<tr>
<th>Method</th>
<th>football</th>
<th>coastguard</th>
<th>tennis</th>
<th>garden</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nearest neighbour</td>
<td>28.24</td>
<td>30.03</td>
<td>28.46</td>
<td>29.97</td>
</tr>
<tr>
<td>Bi-linear</td>
<td>29.48</td>
<td>30.48</td>
<td>29.02</td>
<td>30.05</td>
</tr>
<tr>
<td>Bi-cubic</td>
<td>30.416</td>
<td>31.21</td>
<td>30.39</td>
<td>29.93</td>
</tr>
<tr>
<td>Y Chen [17]</td>
<td>31.407</td>
<td>32.02</td>
<td>31.46</td>
<td>28.78</td>
</tr>
<tr>
<td>Proposed method</td>
<td>34.486</td>
<td>33.29</td>
<td>31.96</td>
<td>30.3</td>
</tr>
</tbody>
</table>

TABLE II
SSIM: COMPARISON OF DIFFERENT SUPER RESOLUTION METHODS

<table>
<thead>
<tr>
<th>Method</th>
<th>football</th>
<th>coastguard</th>
<th>tennis</th>
<th>garden</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nearest neighbour</td>
<td>0.401</td>
<td>0.662</td>
<td>0.553</td>
<td>0.168</td>
</tr>
<tr>
<td>Bi-linear</td>
<td>0.416</td>
<td>0.680</td>
<td>0.559</td>
<td>0.6209</td>
</tr>
<tr>
<td>Bi-cubic</td>
<td>0.407</td>
<td>0.664</td>
<td>0.567</td>
<td>0.618</td>
</tr>
<tr>
<td>Y Chen [17]</td>
<td>0.481</td>
<td>0.6254</td>
<td>0.574</td>
<td>0.717</td>
</tr>
<tr>
<td>Proposed method</td>
<td>0.718</td>
<td>0.779</td>
<td>0.632</td>
<td>0.729</td>
</tr>
</tbody>
</table>

TABLE III
SNR: COMPARISON OF DIFFERENT SUPER RESOLUTION METHODS (dB)

<table>
<thead>
<tr>
<th>Method</th>
<th>football</th>
<th>coastguard</th>
<th>tennis</th>
<th>garden</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nearest neighbour</td>
<td>26.73</td>
<td>28.23</td>
<td>27.29</td>
<td>28.46</td>
</tr>
<tr>
<td>Bi-linear</td>
<td>27.60</td>
<td>28.56</td>
<td>28.41</td>
<td>29.02</td>
</tr>
<tr>
<td>Bi-cubic</td>
<td>27.91</td>
<td>29.43</td>
<td>29.43</td>
<td>30.10</td>
</tr>
<tr>
<td>Y Chen [17]</td>
<td>28.89</td>
<td>29.97</td>
<td>29.89</td>
<td>30.89</td>
</tr>
<tr>
<td>Proposed method</td>
<td>29.97</td>
<td>31.42</td>
<td>30.43</td>
<td>31.23</td>
</tr>
</tbody>
</table>

TABLE IV
MSE: COMPARISON OF DIFFERENT SUPER RESOLUTION METHODS

<table>
<thead>
<tr>
<th>Method</th>
<th>football</th>
<th>coastguard</th>
<th>tennis</th>
<th>garden</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nearest neighbour</td>
<td>9.68</td>
<td>7.878</td>
<td>9.439</td>
<td>7.893</td>
</tr>
<tr>
<td>Bi-linear</td>
<td>8.39</td>
<td>7.49</td>
<td>9.295</td>
<td>7.86</td>
</tr>
<tr>
<td>Bi-cubic</td>
<td>7.535</td>
<td>6.877</td>
<td>7.558</td>
<td>7.97</td>
</tr>
<tr>
<td>Proposed method</td>
<td>4.716</td>
<td>5.413</td>
<td>6.31</td>
<td>7.637</td>
</tr>
</tbody>
</table>

TABLE V
RMS: COMPARISON OF DIFFERENT SUPER RESOLUTION METHODS

<table>
<thead>
<tr>
<th>Method</th>
<th>football</th>
<th>coastguard</th>
<th>tennis</th>
<th>garden</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nearest neighbour</td>
<td>20.19</td>
<td>21.23</td>
<td>25.62</td>
<td>41.89</td>
</tr>
<tr>
<td>Bi-linear</td>
<td>20.15</td>
<td>21.67</td>
<td>27.03</td>
<td>40.54</td>
</tr>
<tr>
<td>Bi-cubic</td>
<td>20.15</td>
<td>20.68</td>
<td>27.87</td>
<td>41.89</td>
</tr>
<tr>
<td>Y Chen [17]</td>
<td>26.6</td>
<td>23.4</td>
<td>39.12</td>
<td>42.36</td>
</tr>
<tr>
<td>Proposed method</td>
<td>20.7</td>
<td>20.67</td>
<td>38.76</td>
<td>42.9</td>
</tr>
</tbody>
</table>

APPENDIX

The structural similarity index measure (SSIM) between an original image x (i; j) and test image y (i; j) is given by

\[ Q = \frac{4\sigma_{xy}\cdot\bar{x}\cdot\bar{y}}{(\sigma_x^2 + \sigma_y^2) \cdot (\bar{x}^2 + \bar{y}^2)} \]  

(6)

Here, \( \bar{x}, \bar{y} \) are the mean of original and reconstructed images and \( \sigma_{xy} \) is the co-variance between original and reconstructed images and \( \sigma_x^2, \sigma_y^2 \) are the variance of original and reconstructed images. The high value of SSIM indicates a good quality image.

The mean square error (MSE) between an original image \( r(i, j) \) and test image \( t(i, j) \) is given by

\[ MSE = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (r(x, y) - t(x, y))^2 \]  

(7)

Where MN is total number of pixels in an image and less value of MSE indicates a good quality image.

The PSNR (Peak Signal to Noise Ratio) is given by

\[ PSNR = 10\log \left( \frac{\text{max} | r(x, y) |^2}{MSE} \right) \]  

(8)

The lower value of PSNR indicates poor quality of the image.

The signal to noise ratio between an original image \( r(i, j) \) and test image \( t(i, j) \) is given by

\[ SNR = 10\log \frac{\text{max} | r(x, y) |^2}{\text{E}[(r-x)^2]} \]  

(9)

The higher value of SNR indicates good quality of the image.

The MAE (mean average error) between an original image \( r(i, j) \) and test image \( t(i, j) \) is given by

\[ MAE = \frac{1}{pq} \sum_{i=0}^{p-1} \sum_{j=0}^{q-1} | r(x, y) - t(x, y) | \]  

(10)

Where pq is total number of pixels in an original image. Lower the value of MSE indicates image is good quality. The root mean square value (RMS) is...
difference between an original image \( r(i; j) \) and test image \( t(i; j) \) is given by

\[
\text{RMS} = \sqrt{\text{MSE}}
\]  

(11)

The lower value of RMS indicates a good quality image.

REFERENCES


