Super Resolution Reconstruction Based on Motion Estimation Error and Edge Adaptive Constraints

Miao Liu, Hanqiang Cao*, Xutao Li, Sheng Yi

(Dept. of Elec. & Info. Eng., Huazhong University Science and Technology, Wuhan 430074)

ABSTRACT

In order to improve the quality of image with super-resolution reconstruction, a method based on motion estimation error and edge constraint was proposed. Under the condition of data consistency and amplitude restriction, the motion estimation error was analyzed, with its variance being calculated; meanwhile, in order to suppress the ringing artifacts, edge constraint was adopted and a method based clustering for judging the edge’s direction was proposed. The experimental results show that the performance of the this algorithm is better than the traditional linear interpolation and method without considering motion estimation error both in vision effect and peak signal to noise ratio.

Keywords: super-resolution image restoration, projection to convex set (POCS), motion estimation error, edge constraint

1. INTRODUCTION

Super-resolution restoration is a well known inverse problem and extensively studied in recent years1,2, since by utilizing potential redundancy, it has the ability to break through the resolution limit of original imaging system and recover the underlying high-resolution images from low-resolution video, thus can satisfy mental and visual need of human being, as well as enhance the performance of pattern recognition in machine vision. With its merit, it has wide application in remote sensing, frame freeze in video, military surveillance as well as medical diagnostics.

Since Tsai and Huang3 first derived a system equation that describes the relationship between LR images and HR image in the frequency domain, various methods had been proposed, to further improve the technique1, 2 And it now concentrated mainly in spatial and compress domains, both in which, the POCS algorithm of sets theory is effective and extensible. In conventional POCS algorithm, it is assumed that the motion estimation error is neglected or considered the same for all low-resolution image sequences4. However from statistical result, the standard deviation of motion estimation error is proportional to the distance between two frames. On the other hand, ringing artifact always exists along object’s edges due to overestimation of degradation parameter since all the pixels are treated identically5.

In order to overcome the above two drawbacks, we propose an algorithm based on POCS combining with motion estimation error analysis and edge restriction. In this approach, the local standard deviations of HR images are calculated as a threshold for limiting the motion estimation error, and the variance of the projection error is calculated in the first step, so that the threshold value for restoration is determined by it adaptively. Furthermore, to reduce the ringing artifact, edge constraint is adopted, in which we employ the method of clustering to judge edge’s direction, and it is noticeable that all these are feasible without increasing much calculation burden. Experimental results indicate that the proposed

*Communication author: itdnet@mail.hust.edu.cn; phone: (+86) 027-87556072
algorithm outperforms linear interpolation and conventional approach in terms of both objective measurements and visual evaluation.

This paper is organized as follows. Section 2 explains the main concept of super resolution restoration and its mathematic model; Section 3 based on the analysis of the motion estimation error and the direction judgment in edge constraint, presents an improved POCS implementation algorithm. Simulation on general video sequences is presented in Section 4.

2. FUNDAMENTALS

2.1 Problem Formulation

On the video capturing process, there exists natural lose of the space resolution in each frame, which is caused by optical distortion (defocus, diffraction), exposure-time of camera and noise in the sensor or during transmission, as well as insufficient sensor density. Let the obtained degraded low-resolution frames of size $N_1 \times N_2$, be denoted by $\{Y_n\}$, $n = 1 \cdots p$, the original high resolution images of size $aN_1 \times bN_2$ be denoted by $\{X_n\}$, $n = 1 \cdots p$, $a > 1, b > 1$. $a$ and $b$ represent down-sampling factor of horizontal and vertical direction respectively. Their relation can be expressed by the linear modal $y_n = DB_nM_nx + n_k$, where matrix $M_k$ is a motion matrix of size $L_1N_1L_2N_2 \times L_1N_1L_2N_2$, $B_k$ of size $L_1N_1L_2N_2 \times L_1N_1L_2N_2$ represents blur operation, $D$ of size $(N_1N_2)^2 \times L_1N_1L_2N_2$ and $n_k$ of size $aN_1 \times bN_2$ represent down-sampling and additive noise respectively. The observation modal relating LR frames and HR frames is illustrated in Fig. 1, and the simulation result of the degradation process is shown in Fig. 2.

![Observation model relating LR images to HR images](image1)

![Simulation result of the image degradation process](image2)

When define the LR image as the weighted sum of related HR pixels images with additive noise. We can get a general model as follows:
\[ y_k' = W_k x + n_k \]  

Where \( W_k = DB_n{M_n} \) represents, via blurring, motion, and sub-sampling, the contribution of HR pixels in \( x \) to the LR pixels in \( y_k' \).

### 2.2 Image Restoration Based on POCS

Super resolution restoration is realized through solving the inverse problem in (1). Owing to the ill-posedness of \( W_k \), it is usually difficult to get the solution of (1) directly. One feasible way to get the solution of \( x \) is, based on formula (1), to introduce various prior knowledge which can constrain the solutions.

By utilizing the POCS method, incorporating a priori knowledge into the solution can be interpreted as restricting the solution to be a member of closed convex set \( \mathcal{C}_i \) that are defined as a set of vectors which satisfy particular properties, such as positive amplitude, finite energy, etc. Thus if the constraint sets have a nonempty intersection, then by alternating projections onto these convex sets, a solution that belongs to the intersection set which is also a convex set, can be found. In fact, any solution in the intersection \( \bigcap_{i=1}^{n} \mathcal{C}_i \) set is consistent with the priori constraint, thus it is a feasible solution.

Assuming that the motion information is accurate, a data consistency constraint set, based on the observation model in (2), is represented for each pixel within the LR images

\[ C^d_{m_1,m_2} = \{ x[n_1,n_2] : r^{(x)}[m_1,m_2] \leq \delta_x[m_1,m_2] \} \]  

(2)

Where \( r^{(x)}[m_1,m_2] = y_k[m_1,m_2] - \sum_{n_1,n_2} x[n_1,n_2] W_k[m_1,m_2;n_1,n_2] \). \( \delta_x[m_1,m_2] \) is the bound reflecting the statistical confidence, which is determined by noise’s character. The actual image is a member of set \( C^d_{m_1,m_2} \). The projection of an arbitrary \( x[n_1,n_2] \) onto the convex set can be defined as:

\[ x^{(n+1)}[n_1,n_2] = x^{(n)}[n_1,n_2] + \begin{cases} 
\frac{(r^{(x)}[m_1,m_2] - \delta_x[m_1,m_2]) W_k[m_1,m_2;n_1,n_2]}{\sum_{p,q} W_k^2[m_1,m_2,p,q]} \cdot r^{(x)}[m_1,m_2] > \delta_x[m_1,m_2], \\
0 \quad r^{(x)}[m_1,m_2] \leq \delta_x[m_1,m_2], \\
\frac{(r^{(x)}[m_1,m_2] + \delta_x[m_1,m_2]) W_k[m_1,m_2;n_1,n_2]}{\sum_{p,q} W_k^2[m_1,m_2,p,q]} \cdot r^{(x)}[m_1,m_2] < -\delta_x[m_1,m_2].
\end{cases} \]  

(3)

While, in fact, the data consistency is not always satisfied due to motion estimation error, and additional constraints such as amplitude constraint and edge constraint can be utilized to improve the results. Thus we proposed a super resolution restoration algorithm based on motion estimation error and edge adaptive constraints.
2.3 Motion Estimation Error

Motion estimation is a critical part of SR restoration algorithm. It determines whether the sub-pixel information between two adjacent frames can be effectively utilized. Due to the ill-posedness of 2D motion estimation completely depends on image intensity, there always exists errors, which can be regarded as noise in the restored image.

Considering the structure of $W_k$ in formula (1), it contains the sub-pixel information between the k-th LR image and HR image, and can be denoted by

$$W_k = \hat{W}_k + \Delta W_k$$

(4)

In formula (4), $\hat{W}_k$ contains the accurate motion estimation information between HR frame and the k-th LR image, and $\Delta W_k$ is the error caused by inaccurate motion estimation. It is evident that the more the motion estimation error increase, the more variation between $W_k$ and $\hat{W}_k$ enlarges, thus by using such $W_k$ for restoration, would inevitably distorts the quality of the restored HR image, rewrite (1) as follows:

$$y_k = (\hat{W}_k + \Delta W_k)x + n_k$$

$$= \hat{W}_k x + (\Delta W_k x + n_k)$$

$$= \hat{W}_k x + n_k$$

(5)

If the additive noise $n_k$ is not considered, $\hat{W}_k = \Delta W_k x$ represents the noise caused by the motion estimation error between k-th LR frame and HR frame.

2.4 Edge Adaptive Constraints

It is often assume that the obtained LR images are the combination of the HR pixels influenced by PSF (Point Spread Function), with the following form $h(x, y) = \begin{cases} 1/(\pi R) & \text{if } x^2 + y^2 \leq R^2 \\ 0 & \text{otherwise} \end{cases}$, $R$ is the defocal radius estimated through spectrum analysis, for both simplicity and feasibility, it can be denoted by the exponential form $h(x, y) = e^{-\frac{x^2 + y^2}{2}}$, however, if used in the restoration procedure directly, it would produce ringing artifacts along edges, since all the pixels are treated identically, and the edge information are not utilized.

In order to maintain edge information and suppress the ringing artifact caused by repairing with isotropy PSF, it is reasonable to adopt Edge Constraint, which employs anisotropy PSF with the form $h(x, y) = e^{-\frac{\lambda_1 x^2 + \lambda_2 y^2}{2}}$ to repair the pixels that are locating in the unsmooth region or along edges. $\lambda_1$, $\lambda_2$ control the influence of horizontal and vertical
direction, respectively.

3. IMPLEMENTATION

3.1 Analysis of Motion Estimation Error in Video Sequences
To illustrate the character of motion estimation error, we have used the video sequence MobilAndCalendar for analysis, the result of motion vectors and the projection error are illustrated in Fig. 3 and Fig. 4 respectively. From intuitive inspection, the distribution of the error is not strictly Gaussian type (high impulse in the centre, with long tail along two sides). Albeit the result is slightly different from the result in [5], they are not conflicting, since here we are dealing with video images. As the distance between two frames increase, the background does not change greatly, while the object’s position and shape would have marked difference, which leads to the long tail distribution of motion estimation error (similar result can be obtained in the test of other video sequences).

Through the analysis of projection error’s statistical parameter, we have got the relationship between frame distance and motion estimation error, which is illustrated in Table.1. It shows explicitly that the larger the frame distance, the larger the projection error. The relationship is in accord with the result in Fig. 4. The distribution can be regard as generalized Gaussian distribution, with its standard deviation being proportional to the error’s power and the mean value approaching 0. So we can modify the data consistency set as follows:

\[ C_D^{[k]}(m_1,m_2) = \{ x[n_1,n_2]; f(x)[m_1,m_2] \leq \lambda \delta_k \} \]  

Fig. 3 Motion Vector field

Fig. 4 Distribution of the projection error of motion estimation

Table 1 relationship between frame interval and motion estimation error parameter

<table>
<thead>
<tr>
<th>frame distance</th>
<th>Mean</th>
<th>variance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.3547</td>
<td>-0.3553</td>
</tr>
<tr>
<td>variance</td>
<td>27.6866</td>
<td>35.1121</td>
</tr>
</tbody>
</table>

Proc. of SPIE Vol. 6246 62460B-5
Meanwhile to void the ill posed-ness, and extract HR images from the non Gaussian type noise, we have to limit the noise before repairing pixels, thus we define a convex sets below, to suppress the noise of motion estimation error.

\[ C_{A}[n_1, n_2] = \{ x[n_1, n_2]: |D(n_1, n_2)| \leq \sigma_k(n_1, n_2) \} \] (7)

\( \sigma_k(n_1, n_2) \) denotes local variance of reference frame, with its value reflecting the local variation of image and

\[
D(n_1, n_2) = \sum_{p=-1}^{1} \sum_{q=-1}^{1} \left| x\left( a(n_1 + p) + MV_{y}, b(n_2 + q) + MV_{x} \right) - y(n_1 + p, n_2 + q) \right| \] (8)

Then we can define a projection operator below

\[
P_{D(n_1, n_2)} = \begin{cases} 
\text{internal repair} & , |D(n_1, n_2)| < \sigma_k(n_1, n_2) \\
\text{directional interpolation} & , |D(n_1, n_2)| \geq \sigma_k(n_1, n_2) 
\end{cases} \] (9)

When the projection error \( D(n_1, n_2) \) is larger than \( \sigma_k(n_1, n_2) \), it means that the motion estimation error is larger, and few sub-pixel information are available, and to maintain image quality as much as possible, directional interpolation is adopted.

### 3.2 Method of Direction Judgment in Edge Restraint

To utilize the edge adaptive constraints, it is critical that the judgment of pixel’s directions is correct. Hence, we proposed a method based on the idea of clustering to decide edge direction, Assuming that the PSF’s range of influence is 3×3, as illustrated in Fig 6, (here we used four direction for illustration, to further improve the accuracy, other direction such as CEH, BEH, AEH, DEI etc can be considered). The detail of the procedure is listed below:

First compute the intra-class variance \( \text{var}_1, \text{var}_2, \text{var}_3, \text{var}_4 \) of four directions (dir1,dir2,dir3,dir4) around centre point E respectively;

\[
\text{var}_1 = \left\{ \left[ D - \text{mean}(D, E, F) \right]^2 + \left[ E - \text{mean}(D, E, F) \right]^2 + \left[ F - \text{mean}(D, E, F) \right]^2 \right\} / 3
\]

\[
\text{var}_2 = \left\{ \left[ C - \text{mean}(C, E, G) \right]^2 + \left[ E - \text{mean}(C, E, G) \right]^2 + \left[ G - \text{mean}(C, E, G) \right]^2 \right\} / 3
\]

\[
\text{var}_3 = \left\{ \left[ B - \text{mean}(B, E, H) \right]^2 + \left[ E - \text{mean}(B, E, H) \right]^2 + \left[ H - \text{mean}(B, E, H) \right]^2 \right\} / 3
\]

\[
\text{var}_4 = \left\{ \left[ A - \text{mean}(A, E, I) \right]^2 + \left[ E - \text{mean}(A, E, I) \right]^2 + \left[ I - \text{mean}(A, E, I) \right]^2 \right\} / 3
\]

Then find \( \text{var}_{\text{min}} = \min \{ \text{var}_1, \text{var}_2, \text{var}_3, \text{var}_4 \} \), which means the deviation of the pixels value along that direction is the smallest, thus they belongs to the same direction. The subscript of the above value indicates the central point’s direction.
Finally the repairing function is determined according to the subscript.

3.3 Algorithm Implementation
Combining with Data consistency, edge preserving, amplitude restriction and motion estimation error, we proposed a new SR algorithm based on POCS. The implementation flow is given below.

Step1 choose one frame in the video as a reference frame, and the fore-and-aft L frames as observation frames;
Step2 bilinear interpolate the reference frame, to get the initial estimated HR image $x^0$, and calculate its local variance $\sigma^2$;
Step3 estimate the motion information between HR frame and the LR frames;
Step4 detect the edge of HR image, and judge the direction.
Step5 calculate the projection error $D(n_1,n_2)$, and Judge the direction of the edges, if $D(n_1,n_2) > \sigma^2(n_1,n_2)$
go to Step6, else go to Step7;
Step 6 directional interpolate the point, which has inaccurate motion information, go to Step8
Step 7 determine the PSF according to the direction of $(n_1,n_2)$, using the projection error to repair pixels within it’s influence arrange.
Step 8 All pixels are finished, go to Step 9, else return to Step 5
Step 9 if $\frac{|r_{n+1} - r_n|}{r_n} \leq \varepsilon$, ($\varepsilon = 0.001$) or satisfied certain iterative time, stop the procedure, otherwise return step3.

4. EXPERIMENTAL RESULTS

We use video sequence QCIF-Claire and CIF-MobilAndCalender for testing. As showed in Fig. 6, the images are first blurred by a $3 \times 3$ Gaussian filter, and then decimated to get the low-resolution observations. Since the existing the of original high resolution images, the restoration result can be compared both objectively and subjectively.
Fig. 6 test images (a) Original High resolution image (the fifth frame of “mobile and calendar” sequence); (b) down-sampled LR image of (a); (c) Original High resolution image (the fifth frame of “Claire” sequence); (d) down-sampled LR image of (c)

4.1 Comparison among the Results of Different Methods
We have tested the method of bilinear interpolation, POCS without considering motion estimation error, POCS without considering edge constraint and our method respectively, and used fully sub-pixel based search method for motion estimation. After 5 times iteration, we get the results shown in Fig7.

As described from Fig. 7 (a) by using common intra-interpolation, it is inherently limited by the amount of data available with in the frame, and it is difficult to produce new frequency component, so the restored image looks blurry. While by using multi-frame POCS method, much new frequency component can be produced to recover the lost detail information. To further compare the result, (b) hasn’t considered the motion estimation error, so there are some artifacts caused by the
ill-posed-ness of 2D-motion estimation. (c) has considered the motion estimation error, but hasn’t adopted the edge restraint, so that the sharpness and the edges has been reduced and the ring artifact pricked up (the number “1” in the calendar); (d) has considered both the motion estimation error and the edge restraint. After 5 iterative times, the result from the proposed method looks more sharper and less artifacts, meanwhile the PSNR value has increased by 0.05dB (see Table 2).

<table>
<thead>
<tr>
<th>Discretization method</th>
<th>POCS Without considering motion estimation error</th>
<th>POCS Without edge constraint</th>
<th>Proposal method</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNR/dB</td>
<td>10.568059</td>
<td>12.202796</td>
<td>12.249342</td>
</tr>
<tr>
<td>PSNR/dB</td>
<td>21.79602</td>
<td>23.431638</td>
<td>23.478185</td>
</tr>
<tr>
<td>MSE/dB</td>
<td>26.333902</td>
<td>24.699165</td>
<td>24.652619</td>
</tr>
</tbody>
</table>

4.2 The Influence of Frame Number to Restoration Result
In the process of restoration, the number of frames available, has positive affection to the restoration results, the more the frame number, the more sub-pixels are available for restoring the lost information (see Fig. 8). But as the frame number increases, the consuming time rises, so we need to achieve the tradeoff between the frame numbers and restoration time according to the application.

![Fig. 8](image)

4.3 The Influence of Iterative Time to Restoration Result
Since the restoration based on POCS method is an iterative process, the iterative times will have significant impact on the result (see Fig. 9). The experiment result indicates that the improvement of image quality is proportional to the iterative time, however when exceeding certain times, the enhancement is negligible. It is also noticeable that at the same iterative times, our method is predominant in performance parameter of SPNR and MSE. Fig. 10 shows the relationship between the iterative number and SPNR, MSE and Projection error, respectively.
Fig. 9 results from different iteration number (a1) 1 time, (a2)(b2) 5 times, (a3) (b3) 10 times, (a4)(b4) 20 times

Fig. 10 relationship between iteration number and restoration parameter
(a) experimental result of “Mobile And Calendar” sequences; (b) experiment result of “Claire” sequences

5. CONCLUSION
In this paper, we have proposed a new method based on edge constraint and high-resolution image reconstruction algorithm considering inaccurate motion estimation. The motion estimation error and edge information were synthesized into the framework of POCS method, as well, the restoration of sub-sampled and blurred image is studied. The specialty of our work is that the motion estimation error information and edge information are efficiently utilized, so that the threshold value for repair can be selected adaptively. The experiment result indicates that the performance of the
proposed algorithm is superior to the conventional method both subjectively and objectively.

ACKNOWLEDGEMENTS
The work in this paper is supported by the Huawei Foundation (YJCB2004007MU) and the Innovation Foundation of Huazhong University of Science and Technology. (JJ-05-071).

REFERENCES