MULTIFRAME SUPER-RESOLUTION RECONSTRUCTION BASED ON CYCLE-SPINNING

Yan Cheng, Xiangzhong Fang, Jun Hou, Songyu Yu

Institute of Image Communication and Information Processing
Shanghai Jiao Tong University, Shanghai

ABSTRACT

A multiframe super-resolution (SR) reconstruction algorithm based on cycle-spinning (CS) is proposed. We utilize the relative motion information of sequential images to construct a CS-based framework for the resolution enhancement. The unique feature of the proposed algorithm is that it is effective for low-resolution (LR) images with various point spread function (PSF) and noise characteristics, even if the degradation models are unknown for the imaging system. Moreover, the computational complexity is inexpensive. Experiments demonstrate the effectiveness of the proposed method and show the superiority to previous methods in objective and subjective qualities.

Index Terms— super-resolution, cycle-spinning, multiframe

1. INTRODUCTION

Super-resolution (SR) reconstruction is to produce high quality, high-resolution (HR) images from a set of degraded, nonidentical, low-resolution (LR) images. The frequency domain approach first addressed in [1], which was extended by Bose [2], is simple but restricted to only global translational motion and linear space invariant (LSI) blur. The interpolation-restoration approach is the most basic and intuitive method for SR reconstruction. Ur and Gross [3] performed a nonuniform interpolation of an ensemble of spatially shifted LR images by utilizing the generalized multi-channels sampling theorem. Nguyen [4] exploited the interlacing structure of the sampling grid in SR and derived a wavelet interpolation for interlaced two-dimensional (2D) data. Then, wiener filtering was applied to the interpolated values to restore the degraded images. This kind of methods suffers from the limited degraded models, which are only applicable when the PSF and noise characteristics are identical across all LR images [5]. Others methods fall into the spatial domain. In [6], the iterative back-projection (IBP) method was formulated to estimate the HR image by back projecting the error between simulated LR images, of which back projecting operator is difficult to choose optimally. The authors in [7] introduced maximum a posteriori probability (MAP) method to utilize the prior information of the spatial observation model, which method is assumed on the additive Gaussian noise model. Tekalp *et al.* extended the projection onto convex sets (POCS) formulation to include sensor noise [8]. The iterative spatial domain methods discussed so far are generally computationally expensive with low convergence.

Cycle-spinning (CS) methodology, introduced in [9], was shown to be an effective method against ringing for denoising problems. It has been a useful tool for image processing. Nosratinia applied CS as a post-processing operation to improve the decompression results in the framework of JPEG and JPEG2000 [10-11]. Furthermore, it was also proven effective towards image resolution upscaling in the wavelet domain [12-13].

In this paper, we build a CS-based framework for SR reconstruction from multiple images or sequential video frames. The proposed method is a powerful alternative to enhance the spatial resolution when applied to the LR frames with various point spread function (PSF) and noise characteristics, even if the degradation model is unknown for the imaging system.

The article is organized as follows. Section 2 explains the theoretical basis of CS and CS-based upscaling (CSU) method. In section 3, a detailed description of our method is presented. Experimental results are explained in Section 4. Section 5 concludes the paper.

2. THEORETICAL BASIS

2.1 CS methodology

CS is usually applied to average over shifts to suppress artifacts for denoising problems. For a range R of shifts, the CS methodology is expressed as

$$\overline{T}(x;(S_r)_{r\in R}) = Ave_{r\in R} S_{-r}(T(S_r(x))) \qquad (1)$$

where $S_r(\cdot)$ represents the shift with *r*. $S_{-r}(\cdot)$ means the inverse shift with *r*. *T* is the denoising operator. In other word, (1) can be described as "Average[unshift-denoising-shift] ".

2.2 CSU algorithm

The CSU method [12] offers a low-complexity yet powerful technique producing good quality image reconstructions for a useful range of enlargement factors. It consists of four stages presented in Fig.1 to perform successively:

i) initial approximation ($\widehat{f}_0$) by wavelet-domain zero padding (WZP) from one LR image,

ii) spatial shift and downsampling to generate a number of LR images, which yield as $\widehat{g}_{i,j} = LWS_{i,j}\widehat{f}_0$. where L represents preserving the low frequency coefficients with discarding of high frequency ones, W denotes wavelet transform, and $S_{i,j}$ is a shift operator of (i, j) in the range $i, j \in \{-k, -k+1, \cdots k-1, k\}, (k \in Z)$,

iii) WZP processing to all $\widehat{g}_{i,j}$ yielding $\widehat{f}_{i,j}$,

iv) realignment and average of the intermediate HR image $\widehat{f}_{i,j}$ to generate the final HR image by $\widehat{f} = \dfrac{1}{(2k+1)^2}\sum_{i=-k}^{k}\sum_{j=-k}^{k} S_{i,j}^{-1}\widehat{f}_{i,j}$.

Though CSU method enhances the size of the LR image, the enlarged image has the same resolution as the original one because they both have the same information content. The additional pixels of the larger image are interpolated directly from the neighbors of the smaller one. From this point, it is impossible to obtain a HR image from a single image due to lack of additional information content. Furthermore, the reconstruction is heavily dependent on the choice of the LR image, especially when different blurring kernels and additive noises corrupt the LR images

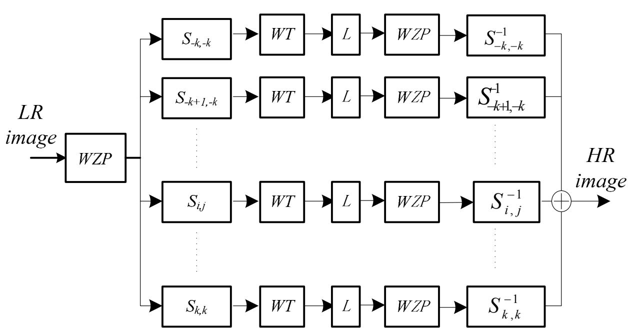


Fig.1 Block diagram of CSU method

3. THE PROPOSED ALGORITHM

3.1 Algorithm description

The observation model that relates the HR frames to the observed LR frames can be mathematically expressed as

$$\mathbf{g}_t = \mathbf{H}_t\mathbf{f}_t + \mathbf{n}_t \qquad (1 \le t \le p) \tag{1}$$

where $\mathbf{f}_t$ is a HR image of size $MNP_1P_2\times1$ written in lexicographical notation as a vector. P_1 and P_2 are the downsampling factor in the horizontal and vertical directions, respectively. Thus, $\mathbf{g}_t$ is a LR image of size $MN\times1$. The matrix $\mathbf{H}_t$ represents the contribution of HR pixels in $\mathbf{f}_t$ to the LR pixels in $\mathbf{g}_t$ through blurring, warping and downsampling operators. Since the HR frames within the HR sequence are related temporally, we formulate the relationship of the HR frames as

$$f_t(x, y) = f_{t-1}(x + s^x_{t-1,t}(x, y), y + s^y_{t-1,t}(x, y)) \tag{2}$$

where $f_{t-1}(x, y)$ and $f_t(x, y)$ are two successive frames at the time of t-1 and t. $s^x_{t-1,t}(x, y)$ and $s^y_{t-1,t}(x, y)$ denote the x and y components of the displacements between the pixels within two successive HR frames. Since the displacements of the sequential frames are intrinsic, we build the multiframe SR reconstruction framework that was analogized to CS methodology in Fig.2.

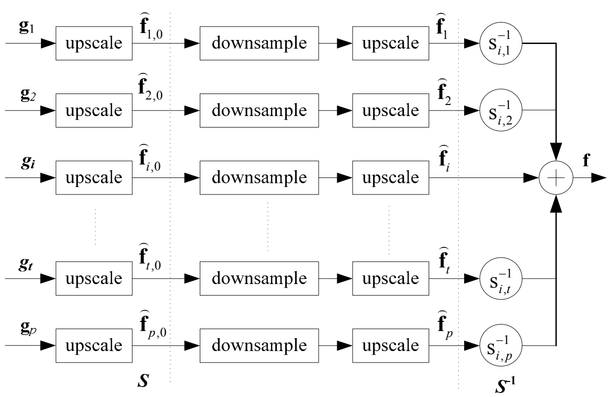


Fig.2 Block diagram of the proposed method

The proposed algorithm consists of the following steps.

(1) Initial estimations on a HR grid: Each LR image $\mathbf{g}_t$ is interpolated to the values $\widehat{\mathbf{f}}_{t,0}$ on a HR grid by an interpolation method, which implies the intrinsic shift information between the sequential images.

(2) Downsampling and upscaling the enlarged images: Each interpolated image $\widehat{\mathbf{f}}_{t,0}$ is downsampled and then upscaled to produce $\widehat{\mathbf{f}}_t$ as

$$\widehat{\mathbf{f}}_t = H^{-1}D\widehat{\mathbf{f}}_{t,0} \tag{3}$$

where $\mathbf{D}$ represents the downsampling operator by wavelet transforming and discarding the high frequency coefficients. $\mathbf{H}^{-1}$ represents the upscaling operator from the LR frame to the HR grid with the same interpolation method in the step (1).

(3) Displacement information estimation: Choose one estimated HR frame ($\widehat{\mathbf{f}}_i$) as a reference frame and then estimate the intrinsic motion information ($\mathbf{s}_{i,t}(i \neq t)$) between the reference and the other HR frames by the optical flow algorithm proposed by Gautama [14].

(4) Final HR frame estimation: These intermediate HR frames $\widehat{\mathbf{f}}_t(t \neq i)$ are warped back to the reference frame $\widehat{\mathbf{f}}_i$ according to the shift information estimated from the step (3), and then averaged to generate the final HR reconstructed frame as follows

$$\mathbf{f} = \frac{1}{p}\sum_{t=1}^{p}\mathbf{s}_{i,t}^{-1}\widehat{\mathbf{f}}_t \tag{4}$$

where p is the total number of the LR frames.

3.2 Computational complexity analysis

A straightforward implementation of this technique is very simple and requires little additional information about the imaging system. The computational burden for the proposed method comprises of two main components. The first component is the cost of upscaling the LR frames by the interpolation algorithm. The second component is the shift information estimation between the intermediate HR frames. For example, to the scale k in the CSU algorithm, the cost of the proposed method is $o(2p \times (2k+1)^2)$. It is more robust and computationally inexpensive than the previous spatial domain method that is repeated for all LR frames by every pixel, i.e. to the MAP and POCS methods, the cost for the entire cycle is $o(n \times p \times M \times N)$ for the iterative number n of the reconstructed procedure.

4. EXPERIMENTS

The performance of the proposed method is evaluated in this section. In the first set of experiments, we use several standard monochrome test images of the size 512×512 for the synthetic experiments. The multiple LR images are created in the three cases listed in Table 1 by translating, blurring and downsampling with the factor 2 in horizontal and vertical directions, respectively. For each case, the global shift M^T belongs to the set generated from the Cartesian product of the horizontal shift and the vertical shift, i.e. $\mathrm{M}^T = \{0,1\} \times \{0,1\} = \{(0,0),(0,1),(1,0),(1,1)\}$. The motion vectors are saved for use in reconstruction. In the proposed method, the initial estimations on the HR grid in the step (1) and the final upscale process in the step (2) are realized by the CSU method. The wavelet transform is implemented by Daubechies 9/7 filter and the maximum shift k is experimentally set to 2. We compare the performance of the proposed method to that of the other algorithms such as the bilinear interpolation, CSU [12], IBP [6] and POCS [8] algorithm. The initial reference frames in the POCS and IBP algorithms are interpolated bilinearly. PSNR values for various standard test images are tabulated in Table 2 to 4 for each case. The tables indicate that the proposed method is applicable in the general imaging system affected by various PSF and noises. Moreover, it is superior to others method in PSNR values.

Table 1 Three cases of synthetic test for the images

Case	M^T	PSF	Additive noise
1	{0,0.5}	3×3Gaussian with the variance 1	Gaussian noise with variances as 40,30,20,10
2	{0,1}	Wavelet low filter	Random noise
3	{0,1}	2 frames of Gaussian low filter; 2 frames of wavelet low filter	Gaussian noise with variances 20

Table 2 PSNR comparisons for Case 1

Image	Lena	Baboon	Boat
Bilinear	27.05	21.732	27.596
CSU	29.242	22.917	29.932
POCS	29.532	22.682	31.325
IBP	28.92	22.292	30.012
Proposal	31.782	23.196	32.93

Table 3 PSNR comparisons for Case 2

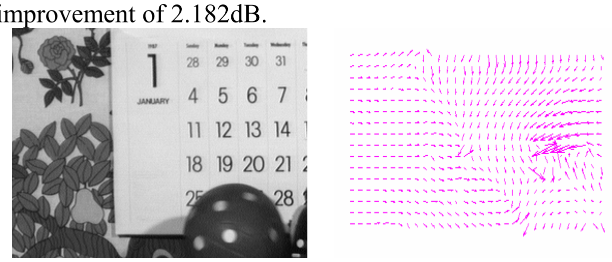
Image	Lena	Baboon	Boat
Bilinear	27.718	22.299	28.26
CSU	28.437	23.118	28.848
POCS	28.754	22.872	29.994
IBP	28.467	22.522	29.794
Proposal	31.226	23.236	32.464

Table 4 PSNR comparisons for Case 3

Image	Lena	Baboon	Boat
Bilinear	27.708	21.901	28.332
CSU	31.329	23.143	32.433
POCS	29.954	22.842	31.187
IBP	28.252	22.129	29.784
Proposal	31.616	23.308	33.362

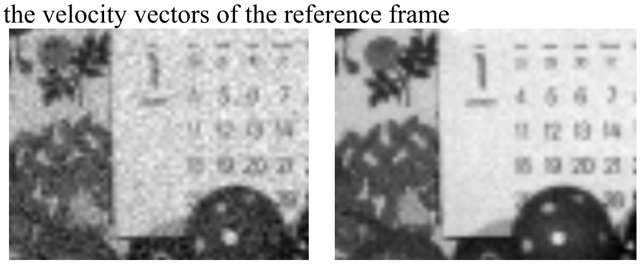
In the second experiment, the entire "mobile" sequence of the size 352×288 shown in Fig. 3(a) is used to perform nonsynthetic resolution enhancement. This experiment differs from the synthetic set in subpixel displacements between the frames, which are no longer defined explicitly and correspond to inherent motion within the scene. Sixteen successive frames are blurred with a Gaussian low-pass filter having a support of three pixels and a variance of 1.

Then the filtered images are decimated by four in horizontal and vertical directions, respectively, and added with random noise to generate a LR sequence. We choose the first frame of the sequence as the reference and estimate the motion information between the intermediate upscaled HR frames. One of the velocity vectors is shown in Fig. 3(b). The visual reconstructions are shown in Fig. 4. Fig. 4(a) shows the reconstruction of the bilinear interpolation, which is terrible to be observed. Fig.4(b) is the reconstruction of the proposed algorithm. It fuses the nonidentical LR images to enhance the image resolution while suppressing noise, especially at the numbers of the calendar. Quantifying the improvement, the PSNR of the proposed algorithm is 25.656dB and betters the bilinear interpolation with the improvement of 2.182dB.



(a) (b)

Fig. 3 Mobile sequence: (a) one of the original frames (b) the velocity vectors of the reference frame



(a) (b)

Fig. 4 Results of "Mobile" sequence: The reconstruction of (a) bilinear interpolation (PSNR=23.474dB) (b) Our algorithm (PSNR=25.656dB)

5. CONCLUSION

In this paper, the intrinsic displacement information among the sequential images or video frames is exploited to build a novel CS-based framework to enhance the image resolution. Compared to the previous methods, the proposed method has three advantages. First, the proposed algorithm solves the nonuniform interpolation and restoration problems simultaneously without any prior knowledge of the degradation model. Second, the proposed algorithm can be used to reconstruct the HR image when the PSF and noise characteristics are not identical across all LR images, which is the limitation of the nonuniform interpolation algorithms. Finally, the computational complex of the proposed algorithm is less than the popular spatial domain methods that can be used to reconstruct towards various degraded processes.

6. REFERENCES

[1] T.S.Huang and R.Y.Tsai, "Multi-frame image restoration and registration," Adv. Comput. Vis. Image Process., vol. 1, pp. 317-339, 1984.

[2] N.K.Bose, H.C.Kim, and H.M.Valenzuela, "Recursive implementation of total least squares algorithm for image reconstruction from noisy, undersampled multiframes," in Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, vol. 5, pp. 269-272, Apr. 1993.

[3] H.Ur and D.Gross, "Improved resolution from subpixel shifted pictures," CVGIP: Graphical Models and Image Processing, vol. 54 (2), pp. 181-186, 1992.

[4] N.Nguyen, and P.Milanfar, "A wavelet-based interpolation-restoration method for superresolution (wavelet superresolution)", Cirsuits Systems Signal Processing, vol. 9(4), pp. 321-338, 2000.

[5] S.C.Park, M.K.Park, and M.G.Kang, "Super-resolution Image Reconstruction: a Technical Overview," IEEE Signal Processing Magazine, vol. 20, pp. 21-36, 2003.

[6] M.Irani and S.Peleg, "Improving resolution by image registration," CVGIP:Graph. Models Image Process., vol.53, pp.231-239, May 1991.

[7] R.C.Hardie, K.J.Barnard, and E.E.Armstrong, "Joint MAP registration and high-resolution image estimation using a sequence of undersampled images," IEEE Trans. Image Processing, vol.6, pp.1621-1633, Dec.1997.

[8] A.Tekalp, M.Ozkan, and M.Sezan, "High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration," In Proc. ICASSP, vol.3, pp. 169-172, 1992.

[9] R.R.Coifman, and D.L.Donoho, "Translation-invariant denoising", Lect. Notes Stat., pp.125-150, 1995.

[10] A.Nosratinia, "Postprocessing of JPEG-2000 images to remove compression artifacts", IEEE Signal Process. Lett., vol. 10(10), pp. 296-299, 2003.

[11] A.Nosratinia, "Denoising JPEG images by re-application of JPEG", *J.VLSI Signal Process*., vol. 27 (1), pp. 69-79, 2001.

[12] A.Temizel, and T.Vlachos, "Wavelet domain image resolution enhancement using cycle-spinning", Electron. Lett., vol. 41(3), pp.119-121, 2005.

[13] A.Temizel, and T.Vlachos, "Image resolution upscaling in the wavelet domain using directional cycle spinning", Journal of Electronic Imaging, vol. 14(4), 040501(3 pages), 2005.

[14] T.Gautama, and M.M.Vhulle, "A phase-based approach to the estimation of the optical flow field using spatial filtering", IEEE Trans. Neural Networks, vol. 13(5), pp.1127-1136, 2002.