Acta Technica *61*, No. 4B/2016, 165–174

Single image super resolution algorithm based on local energy and anisotropic filter in NSCT domain[1]

Zhang Wei[2], Chen Wei[2]

Abstract. A novel super resolution algorithm is proposed for preserving the texture and edge information and improving the space resolution of single image (this technology is widely used, for example, in the field of military reconnaissance or space observation). The sub-band coefficients of the original image are obtained through Nonsubsampled Contourlet Transform (NSCT). Then, the noise of high sub-band is removed a new high frequency coefficient together with coefficient of variation weighted interpolation are obtained. The low sub-band coefficients are enhanced and estimated low frequency coefficients. Finally, the high resolution image is obtained by inverse NSCT. Experiments on both synthetic images and natural images demonstrate the effectiveness of the proposed method. Compared with traditional super resolution reconstruction algorithms, the proposed method can preserve the structure of image and resist Gaussian noise.

Key words. Nonsubsampled Contourlet Transform (NSCT), image interpolation, coefficient of variation, super resolution reconstruction..

1. Introduction

Super resolution is the process of combining a sequence of low-resolution (LR) noisy blurred images to produce a higher resolution image or sequence [1]. Super resolution from image sequences obtain better results than from single image since the former use these similar but not identical characters among the image sequences. Many effective methods have been proposed for super resolution based on multiple low-resolution images of the same scene such as reconstruction-based resolution [2], [3] or Learning-based resolution [4], [5]. In fact, it is very difficult to acquire sequence images for the same scene in the practical application such as military investigation and space observation. Therefore, the research of single image super resolution

[1]The research work is fully supported by the Science and Technology Program of Jilin (Grant No. 20150441003SC), the Science and Technology Program of Jilin (Grant No. 20140101197JC), and the Science and Technology Program of Jilin City (Grant No. 2015554003).

[2]Department of Computer Science and Technology, Beihua University, Jilin, 132021, China

possesses important significance in practical applications. Traditional single image super resolution methods always caused the image edge information blurred that cannot perform well for image characteristics since there is no redundant information in single image. The main purpose of this paper is to improve the space resolution of single image without losing the edge and texture features.

A single image super resolution algorithm is proposed in this paper. The basic idea is to decompose the LR image through NSCT and use coefficient of variation weighted interpolation along the edge direction. A nonlinear function is used to enhancement the dim-target. Then a super-resolution image can be obtained by inverse NSCT. Several experimental results using simulated images demonstrate that the new method is efficient for single image super resolution. Compared with traditional single super resolution algorithms, this algorithm can provide smooth edges and recovery the texture features, and also resists noise.

2. Proposed single image super resolution algorithm

2.1. Nonsubsampled Contourlet Transform

The contourlet transform is an efficient directional multi-resolution image representation. It is defined and implemented by the filter banks in the discrete domain. By a redundant pyramidal multi-resolution decomposition, the image was decomposed into a detail sub-band and approximate sub-band. Then a directional filter banks was applied on each of the detail sub-band to capture the singularity of direction. Different from the wavelet transform, the multi-directionality and the property that the support of basis functions has a variety of elongated shapes with different aspect ratios make it effectively capture the geometric structure features of the image information. More importantly, the contourlet transform can flexibly and effectively combine the multi-scale with the multi-directionality for image representation, therefore it can accurately and optimally describe an image [6]. The nonsubsampled contourlet transform(NSCT) is based on a nonsubsampled pyramid structure and nonsubsampled directional filter banks. It is a shift-invariant image decomposition that can be efficiently implemented and can well depress the Gibbs impact [7]. Compared to the contourlet transform, the NSCT has a better frequency selectivity and is sensitive to texture.

After NSCT, the energy of an image is distributed in various scales and directions. The coefficients have the feature of localization in space domain and exhibit dependencies across all of scales, space, and orientations. Figs. 1a–1e show the 'zoneplate' image, an exemplary frequency partition of the NSCT, low frequency coefficients and four directional sub-bands of the NSCT decomposition in two scales, respectively. It can be seen that the most energy concentrates in low frequency coefficients, so the low frequency coefficients reflect the main information of the image. The high frequency coefficients reflect the detail of image. Most coefficients are so small that they are invisible in high frequency sub-bands.

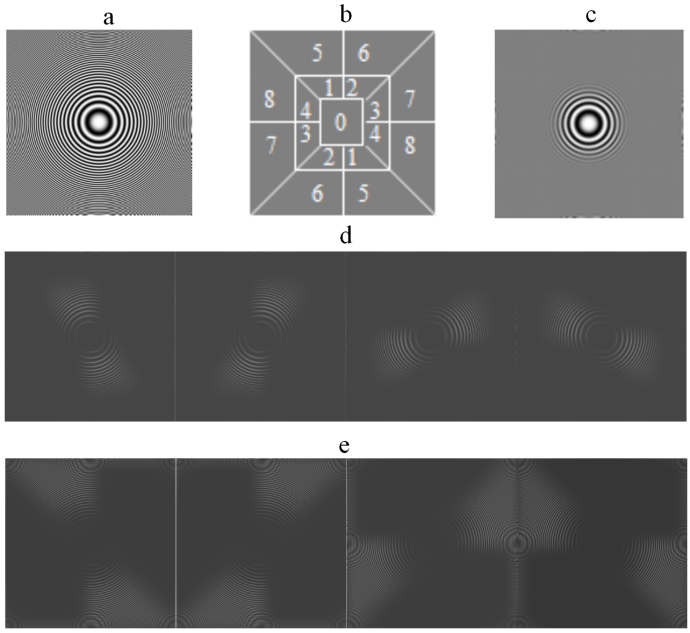

Fig. 1. Example of frequency partition and four directional sub-bands of NSCT coefficients: a–zoneplate image, b–frequency partition, c–low frequency coefficients, d–sub-bands 1–4 of NSCT coefficients, e–sub-bands 5–6 of NSCT coefficients

2.2. Coefficient of variation weighted interpolation in high frequency bands

To improve the resolution, the pixels in low resolution (LR) image should be mapped to the high resolution (HR) image. But there are many "holes" in the HR image to be filled. Interpolation is the simplest method to fill these holes. The traditional interpolation methods (bilinear interpolation, Newton interpolation) estimate interpolation point by neighborhood pixels, so blur is introduced in edge and texture, especially in weak edge and noise.

The NSCT coefficients of noise and weak edge are both small. Different from weak edge, most noises have not obvious geometrical structure, so they can be distinguished by the degree of correlation in different directions of the same scale. Reference [8] introduces a simple method that can remove the noise but preserves the edges as weak edges. First, it is necessary to calculate the high frequency bands threshold T_{B} by the Bayesian shrinkage and the coefficients that are lower than threshold are decided whether it is a noise by the correlation among the directional sub-bands in the same scale. The threshold is determined from the formulae

$$\overline{\delta}_{l,k} = \sqrt{\max\left(\frac{1}{m \cdot n}\sum_{i=1}^{m}\sum_{j=1}^{n} f_{l,k}^2(i,j) - \delta_{l,i}^2\right)}, \tag{1}$$

$$T_{\mathrm{B}} = \frac{\delta_{l,i}^2}{\overline{\overline{\delta}}_{l,i}}\,, \tag{2}$$

where $f_{l,k}(i,j)$ is the NSCT coefficient with coordinates i, j in the l-th scale and k-th direction. Symbol $\overline{\delta}_{l,k}$ is the standard deviation of the sub-band, $\delta_{l,i}$ is the standard deviation of noise and m, n are the sizes of sub-band image.

If the pixel value is lower than a threshold, its coefficient of variation c_{v} can be computed as

$$c_{\mathrm{v}} = \frac{\sqrt{\frac{1}{m\ldots n}\sum_{i=1}^{m}\sum_{j=1}^{n} f_{l,k}^2(i,j) - \delta_{l,i}^2}}{\frac{1}{m\cdot n}\sum_{i=1}^{m}\sum_{j}^{n} f_{l,k}^2(i,j)}\,. \tag{3}$$

The coefficient of noise is small and relatively stable in NSCT domain and the value of c_{v} is relatively small. The threshold $T_\gamma = 0.03\sqrt{m \cdot n}$ is set to differentiate between the noise and weak edges. The pixel will be labeled as noise if its value is lower than T_γ and be modified as the mean of neighborhood pixels.

For every pixel, interpolation should be performed along the edge directions in noise-free image to preserve the edges, so detect the information of edge direction is a necessary step. After the NSCT, contours and textures are mostly located in high frequency sub-bands and will be captured in responding directional sub-bands. Figure 1 shows the NSCT representation of the "Peppers" image. Taking a neighborhood with size of 3×3 for the pixel in the high frequency sub-band we can calculate the coefficient of variation. The maximum and second-maximum are $c_{\mathrm{v,m}}$ and $c_{\mathrm{v,s}}$, the responding directions are d_{m} and d_{s}, respectively. By the Newton interpolation along d_{m} and d_{s} we can obtain the f_{m} and f_{s}, respectively. The final value of the pixel is calculated as

$$f = \frac{cv_{\mathrm{s}}}{c_{\mathrm{v,s}} + c_{\mathrm{v,m}}} f_{\mathrm{s}} + \frac{cv_{\mathrm{m}}}{c_{\mathrm{v,s}} + c_{\mathrm{v,m}}} f_{\mathrm{m}}\,. \tag{4}$$

2.3. Dim target enhancement in low frequency sub-band

There are almost no noises in low frequency sub-band. But the reconstructed image is darker than LR image because of energy dispersion. To improve the visual effect, image enhancement is performed in low frequency sub-band. In addition, those dim-targets have smooth change of brightness from its neighborhood region, therefore they provide no clear edges. To enhance the dim-targets and make the background uniform, the smoothing process should be performed in the faultless areas and the sharpening process should be performed in the local dim areas. Therefore, a gray level transformation in the spatial domain is introduced to enhance the global contrast of reconstructed image effectively. Because human vision is sensitive to contrast, maximum contrast can be used to estimate the threshold between the background and dim-target. If this threshold is t, the absolute contrast $c(t)$ and best threshold T can be obtained as

$$c(t) = \min(f(u) - t, f(l) - t)\,, \tag{5}$$

$$T = \underset{t}{\operatorname{Arg\,max}}(c(t))\,, \tag{6}$$

where $f(u)$ is the mean of those pixels whose values are greater than t while $f(l)$ is the mean of those pixels whose values are less than t.

Contrast stretching uses a nonlinear function

$$f = \operatorname{sign}(f(i,j)) \cdot (|f(i,j)| - f_{\min}) \cdot \left(\sin\left(\frac{\pi}{2} \cdot \frac{f(i,j) - f_{\min}}{T - f_{\min}}\right)\right) + f_{\min}$$

for $f_{\min} < |f(i,j)| < T$ and

$$f = \operatorname{sign}(f(i,j)) \cdot (|f(i,j)| - f_{\max}) \cdot \left(\cos\left(\frac{\pi}{2} \cdot \frac{T - |f(i,j)|}{f_{\max} - T}\right)\right) + f_{\max} \tag{7}$$

for $T < |f(i,j)| < f_{\max}$.

2.4. The arithmetic flow

The arithmetic flow consists of two following steps:

1. The LR image is transformed by NSCT. The LR image is decomposed with three stages by dmaxflat NSDFB and 9-7 NSP of the algorithm proposed in this paper, and then eight directions are decomposed in each high-frequency sub-band.

2. Remove the noise of high sub-band using equations (1)–(3) and get new high-frequency coefficient with coefficient of variation weighted interpolation described by equation (4).

3. Bilinear interpolation is used in low-frequency image. Then the coefficients are enhanced by equation (8) and estimated are low-frequency coefficients.

4. High resolution image is obtained by inverse NSCT with estimated coefficients that are obtained in steps 2 and 3.

3. Experimental results

The synthetic image and natural image are tested with Bicubic, algorithm proposed in [9], algorithm proposed in [10], and the above method. The processed images are in Figs. 2 and 3.

3.1. Experiment on synthetic image

A synthetic image is obtained by double down-sampling and adding the Gaussian noise. Figure 2 illustrates five images. Figure 2a is the part of "boat". Figs. 2b to 2d are the HR images reconstructed by the Bicubic, the algorithm proposed in [9], and

the algorithm proposed in [10],respectively. Figure 2e is the HR image reconstructed by the method proposed in this paper.

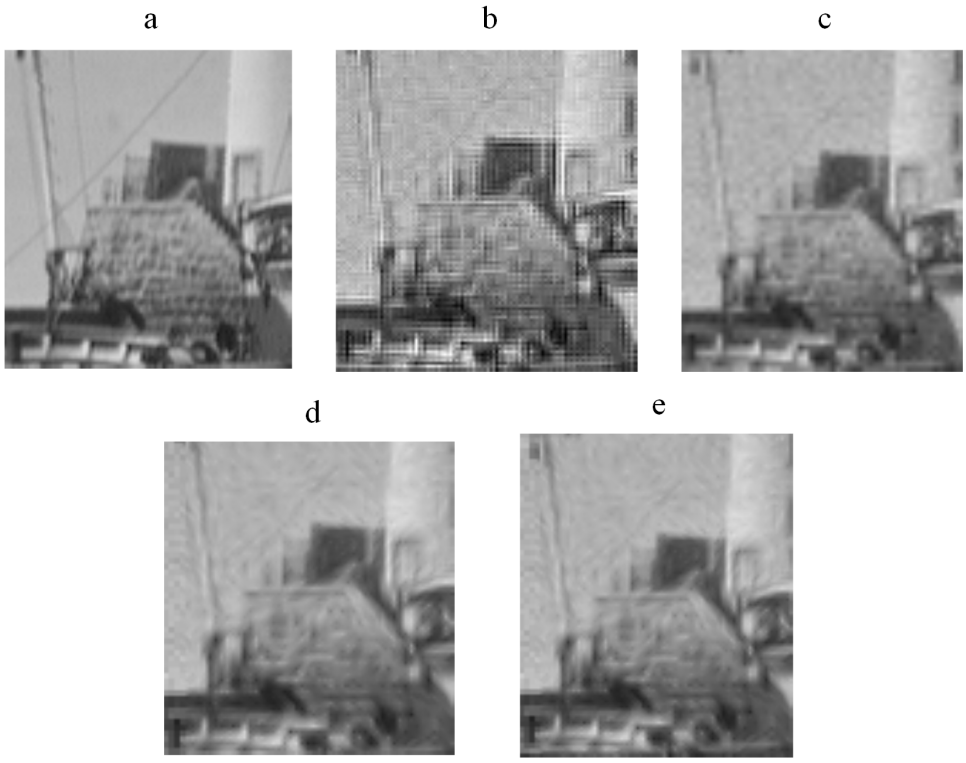

Fig. 2. Super-resolution results of part of boat: a–origin, b–Bicubic, c–processing by [9], d–processing by [10], e–processing by proposed algorithm

3.2. Experiment on natural image

We selected a natural image and tested it using various algorithm. There exist only global translation and rotation in the test image. Experimental results show the efficiency of the proposed algorithm in dim-target enhancement with low contrast.

4. Discussion

The simulation results show that the aliasing phenomenon is serious in image reconstructed by Bicubic. In Fig. 1a, the blocking artifacts are obvious on the mask and roof. SSIM, PSNR and entropy are tested in the LR image and reconstructed image and the results are showed in Tables 1–4. The proposed algorithm is more successful in noise suppression, edge enhancement and gets highest PSNR. The value of SSIM showed in Table 2 illustrates that the proposed algorithm can preserve the

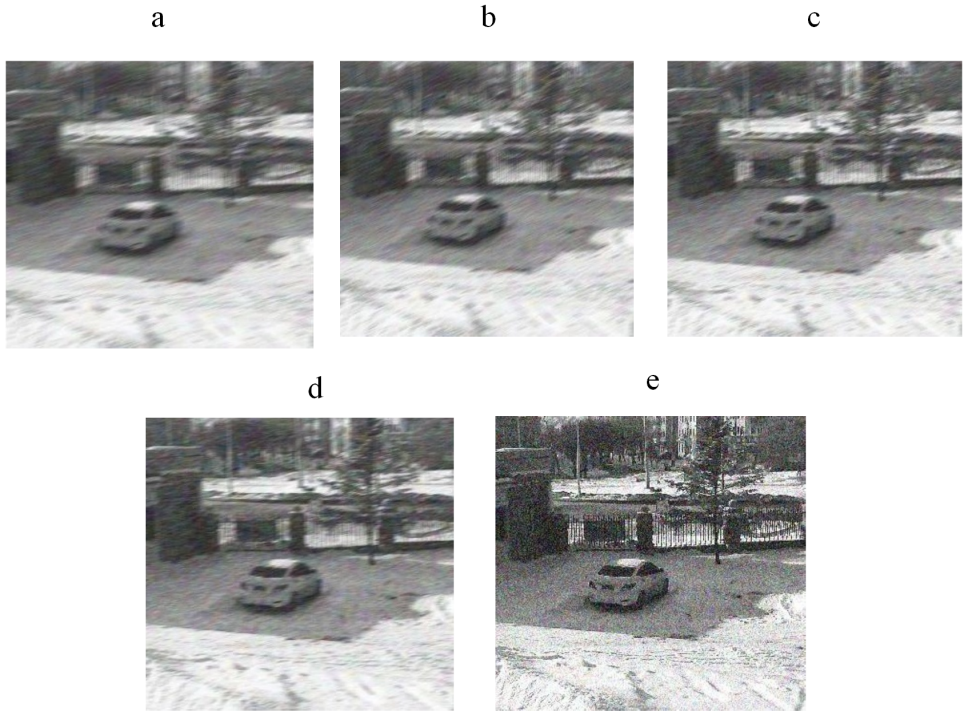

Fig. 3. Super-resolution results of natural image: a–origin, b–Bicubic, c–processing by [9], d–processing by [10], e–processing by proposed algorithm

structure of image effectively. Table 3 shows the corresponding result.

Compared with the bilinear interpolation, the PSNR of the proposed algorithm is improved by 2.4–16.7 %, the average PSNR increase is 5.8 %. Compared with the method of ref. [9], the PSNR of the proposed algorithm is improved by 0.1–3.8 %, and the average PSNR increase is 1.4 %. Compared with the method of ref. [10], the PSNR of the proposed algorithm is improved by 0.1–1.5 %, and the average PSNR increase is 0.35 %.

Table 1. Comparison of the PSNR obtained by different algorithms

image	Bicubic	ref. [9]	ref. [10]	proposed method
airport	29.257	30.784	30.785	30.808
baboon	26.868	28.140	28.424	28.613
barb	26.748	30.852	31.602	31.221
boat	28.676	29.644	29.987	29.988
couple	32.468	33.456	34.652	34.744
girl	31.458	31.782	32.413	32.440
moom surface	24.763	25.154	25.447	25.755
man	30.476	31.735	31.878	31.996
pixel ruler	29.367	30.452	30.553	30.576
testpat	31.782	32.418	32.568	32.669

Table 2. Comparison of the SSIM obtained by different algorithms

image	Bicubic	ref. [9]	ref. [10]	proposed method
airport	0.734	0.754	0.787	0.788
barb	0.759	0.812	0.816	0.828
boat	0.766	0.823	0.832	0.833
baboon	0.721	0.793	0.814	0.826
couple	0.785	0.812	0.825	0.842
girl	0.791	0.845	0.855	0.856
moom surface	0.719	0.763	0.784	0.797
man	0.782	0.803	0.826	0.832
pixel ruler	0.795	0.821	0.832	0.845
testpat	0.771	0.824	0.826	0.844

Compared with the bilinear interpolation, the SSIM of the proposed algorithm is improved by 6.2–14.1 %, the average SSIM increase is 8.8 %. Compared with the method [9], the SSIM of the proposed algorithm is improved by 1.2–4.5 %, the average SSIM increase is 3.0 %. Compared with the method [10], the SSIM of the proposed algorithm is improved by 0.1–2.2 %, the average SSIM increase is 2.1 %.

Table 3. Comparison of the entropy obtained by different algorithms

image	LR image	Bicubic	ref. [9]	ref. [10]	proposed method
airport	3.2569	3.2653	3.3219	3.3323	3.3732
baboon	2.3735	2.4159	2.5748	2.6137	2.6490
barb	2.4888	2.5132	2.6966	2.7345	2.7443
boat	2.6324	2.7032	2.7828	2.8344	2.8432
couple	2.9843	3.0352	3.1274	3.1435	3.2523
girl	2.6369	2.6697	2.7132	2.7312	2.7627
moom surface	3.2153	3.2408	3.3448	3.3562	3.3712
pixel ruler	2.9837	3.0221	3.0435	3.0521	3.0693
testpat	2.8732	2.9072	2.9546	2.9723	2.9952

Compared with the bilinear interpolation, the entropy of the proposed algorithm is improved by 1.6–9.5 %, the average entropy increase is 4.9 %. Compared with the method described in ref. [9], the entropy of the proposed algorithm is improved by 0.8–4.0 %, the average entropy increase is 1.9 %. Compared with the method described in [10], the entropy of the proposed algorithm is improved by 0.4–3.5, the average entropy increase is 1.0 %.

Table 4. Comparison of the MSE obtained by different algorithms

image	Bicubic	ref. [9]	ref. [10]	proposed method
airport	2.4654	1.9874	1.8456	1.8327
baboon	2.2853	1.8823	1.5436	1.4425
barb	2.1357	1.6349	1.4537	1.4218
boat	2.0458	1.4917	1.3627	1.2807
couple	1.9079	1.7543	1.6789	1.6325
girl	1.8894	1.7547	1.6756	1.6479
moom surface	2.5864	2.1464	1.9864	1.8768
man	2.0254	1.9854	1.8546	1.8103
pixel ruler	2.4712	2.0957	1.9954	1.8628
testpat	2.0569	1.9875	1.8539	1.8212

Compared with bilinear interpolation and method [9], the MSE of the algorithm is greatly improved. Compared with the method [10], the MSE of the proposed algorithm is improved by 0.7–6.6 %, the average MSE increase is 3.6 %.

5. Conclusion

The results show that the proposed algorithm provides clear edges, clear texture and overcomes the adverse effects caused by uneven illumination to a certain extent. Also, the effectiveness of the proposed method was confirmed in various simulation conditions.

References

[1] S. Farsiu, M. D. Robinson, M. Elad, P. Milanfar: *Fast and robust multiframe super resolution.* IEEE Trans. Image Processing *13* (2004), No. 10, 1327–1344.

[2] H. Su, Y. Wu, J. Zhou: *Super-resolution without dense flow.* IEEE Trans. Image Processing *21* (2012), No. 4, 1782–1795.

[3] A. K. Katsaggelos, R. Molina, J. Mateos: *Super resolution of images and video.* Synthesis Lectures on Image, Video, and Multimedia Processing. Morgan and Claypool e-Books, 2007.

[4] B. Maihe, D. Barchiesi, M. D. Plumbley: *INK-SVD, learning incoherent dictionaries for sparse representations.* Proc. IEEE IC Acoustics, Speech and Signal Processing, 25–30 March 2012, Kyoto, Japan, 3573–3576.

[5] W. T. Freeman, T. R. Jones, E. C. Pasztor: *Example-based super-resolution.* IEEE Comp. Graph. Appl. *22* (2002), No. 2, 56–65.

[6] M. N. Do, M. Vetterli: *The contourlet transform: An efficient directional multiresolution image representation.* IEEE Trans. Image Processing *14* (2005), No. 12, 2091–2106.

[7] A. L. da Cunha, J. Zhou, M. N. Do: *The nonsubsampled contourlet transform: Theory, design, and applications.* IEEE Trans. Image Processing *15* (2006), No. 10, 3089 to 3101.

[8] Wei Zhang, Zhongcheng Fan: *A no-reference contourlet-decomposition-based image quality assessment method for super-resolution reconstruction.* Proc. SPIE 9273, Optoelectronic Imaging and Multimedia Technology III, 9–11 Oct. 2014, Beijing, China. Vol. 9273.

[9] Haiyuan Zhang, Nian Cai, Nan Zhang: *Using weighted parabolic interpolation to zoom images based on an error-amended sharp edge algorithm.* Computer Engineering and Applications. *7* (2011), No. 25, 194–197.

[10] Jiantao Wang, Yun Zhao, Qing Gao: *Multi-scale wavelet super-resolution image reconstruction based on nonuniform sampling multi-frame images.* Proc. 16th National Symposium on Remote Sensing, 24 Nov. 2008, Beijing, China, Vol. 7123.

Received November 16, 2016