Nonlocal Edge-directed Interpolation

Xinfeng Zhang[1], Siwei Ma[2], Yongbing Zhang[3], Li Zhang[2], Wen Gao[2]

[1] Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
[2] School of Electronics Engineering and Computer Science, Peking University
[3] Department of Computer Science, Harbin Institute of Technology, Harbin, China
{xfzhang, swma, ybzhang, zhanglili, wgao}@jdl.ac.cn

Abstract. In this paper, we proposed a new edge-directed image interpolation algorithm which can preserve the edge features and natural appearance of images efficiently. In the proposed scheme, we first get a close-form solution of the optimal interpolation coefficients under the sense of minimal mean square error by exploiting autoregressive model (AR) and the geometric duality between the low-resolution and high-resolution images .Then the coefficients of the Nonlocal Edge-directed interpolation (NLEDI) are derived with structure similarity in images, which are solutions of weighted least square equations. The new image interpolation approach preserves spatial coherence of the interpolated images better than the existing methods and it outperforms the other methods in terms of objective and subjective image quality.

Keywords: image interpolation, nonlocal edge-directed interpolation, AR

1   Introduction

Image interpolation is one of the most prevalent techniques in digital image processing with a variety of applications. One obvious application of image interpolation is the reproduction of images captured by digital cameras for high quality prints. Another important application is that it can make Standard-definition video frames fit to the High-Definition DTV (HDTV) receiver. Moreover, in consumer electronics area, image interpolation is beneficial and necessary in computer vision, surveillance, medical imaging, remote sensing and other fields.

In recent studies on image interpolation, it is agreed that for many applications, the main emphasis should be on the perceptual quality of images. Sharpness of edges and freedom from artifacts are two critical factors in the perceived quality of the interpolated images. Therefore, the main purpose of image interpolation is to recover sharp edges and textures, while suppressing pixel blocking (known as jaggies) and other visual artifacts. Classical techniques, such as pixel republication, bilinear or bicubic interpolation have the problem of blurred edges or artifacts around edges, to which the human visual system are highly sensitive. To tackle this problem, several methods [1]-[8] have been proposed to improve the subjective quality. Adaptive interpolation techniques [1]-[4] can spatially adapt the interpolation coefficients to better match the local structures around the edges. *Wang* and *Ward* proposed to make use of the gradient to get the edge direction and got the best correlation pixels with

the interpolated pixel [5]. *Zhang* and *Wu* adaptively fused two interpolation results in two mutually orthogonal directions using the statistics of a local window [6]. *Li* and *Orchard* [7] proposed a Wiener-filtering like interpolation scheme utilizing the covariance of the low-resolution (LR) image to estimate the high-resolution (HR) image covariance, which represents the edge direction information to some extent. Since this method needs to compute the covariance matrix in a local window for each interpolated pixel, it may introduce some artifacts due to the local structures changes and, hence, the incorrect estimation of covariance. *Asuni* and *Giachetti* proposed INEDI (Improved New Edge-Directed Interpolation) to improve the algorithm of *Li* by adopting circular windows and adaptively selecting the size of windows [8]. However, in the region of fast luminance change, the windows are too small to get a stable solution.

In this paper, we propose a new edge-directed method for image interpolation (NLEDI). The proposed method can depress the artifacts caused by the difference of the geometry configuration in local window, and make the solution stable. The optimal interpolation coefficients are derived in the weighted least square sense. By further taking advantage of the idea of nonlocal means filter [9], we obtain different weights for samples with different structures from the interpolated position.

The rest of the paper is organized as follows. Section 2 presents the detail description of our proposed NLEDI. Section 3 gives the experimental results, in objective and subjective quality respectively. Finally this paper is concluded in Section 4.

2 Nonlocal Edge-directed Interpolation method

2.1 AR interpolation model

In statistics and signal processing, AR model is a type of random process which is often used to model and predict various types of natural phenomena. AR model can be formulated as (1),

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t \quad . \tag{1}$$

where p is the order of the AR model, $\varphi_1, \cdots, \varphi_p$ are the parameters of the model, c is a constant and ε_t is white noise. Since it can adaptively represent local features of signals, many research works introduced it into the image interpolation to predict the unknown pixels. Without loss of generality, we assume that the LR image $X_{i,j}$ of sized $M \times N$ directly comes from the HR image Y of sized $2M \times 2N$, i.e., $Y_{2i,2j} = X_{i,j}$. The task of image interpolation is how to interpolate the interlacing lattice $Y_{2i+1,2j+1}$ from the lattice $Y_{2i,2j} = X_{i,j}$. Therefore, we can formulate missing pixel in HR images using a fourth-order AR model as follows (refer to Fig.1 (a)):

$$\hat{Y}_{2i+1,2j+1} = \sum_{k=0}^{1} \sum_{l=0}^{1} \alpha_{2k+l} Y_{2(i+k),2(j+l)} + n_{2i+1,2j+1} \quad .$$

(2)

where $\hat{Y}_{2i+1,2j+1}$ is the estimation of the missing pixel $Y_{2i+1,2j+1}$, $Y_{2(i+k),2(j+l)}$ are the four 8-connected neighbors and $n_{2i+1,2j+1}$ is the noise.
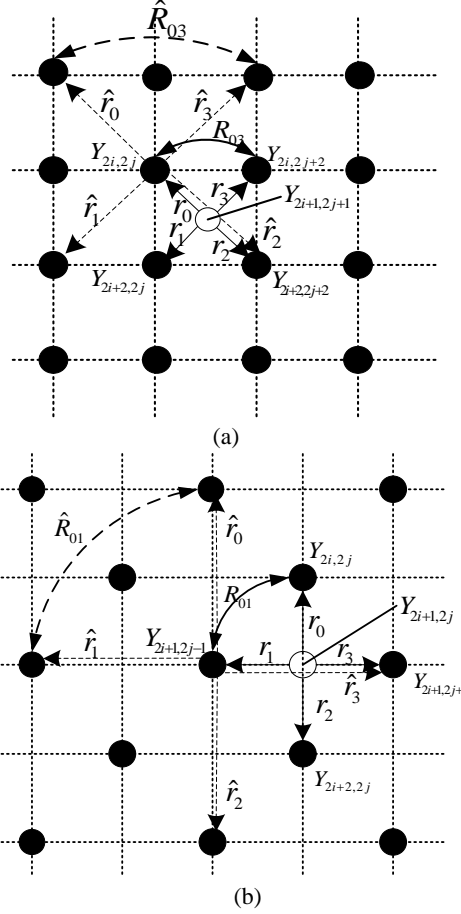

(a)

(b)

Fig. 1 The spatial configuration of image interpolation with AR model. The solid dots are pixels in LR or already interpolated. The circles are to be interpolated HR pixels. (a) The spatial configuration when interpolating $Y_{2i+1,2j+1}$ from $Y_{2i,2j}$. (b) The spatial configuration when interpolating $Y_{i,j}$ ($i+j$ is odd) from $Y_{i,j}$ ($i+j$ is even).

With equation (2), we can get a cost function defined as

$$e^2 = (Y_{2i+1,2j+1} - \hat{Y}_{2i+1,2j+1})^2 = (Y_{2i+1,2j+1} - \boldsymbol{\alpha}^T \mathbf{Y})^2 \quad .$$

(3)

where $\boldsymbol{\alpha} = [\alpha_0 \quad \alpha_1 \quad \alpha_2 \quad \alpha_3]^T$ is the interpolation coefficients in vector form, $\mathbf{Y} = [Y_{2i,2j} \quad Y_{2i,2j+2} \quad Y_{2i+2,2j} \quad Y_{2i+2,2j+2}]^T$ is the vector of four 8-connected neighboring

pixels. For simplicity, we just consider the noise-free case. Under the assumption that the natural image can be modeled as a locally stationary Gaussian process., we can get the optimal interpolation coefficients under minimum mean square error sense as

$$\boldsymbol{\alpha} = \mathbf{R}^{-1}\mathbf{r} \quad , \tag{4}$$

where $\mathbf{R} = [R_{k,l}] = E(\mathbf{Y}^T\mathbf{Y})$, and $\mathbf{r} = [r_k] = E(Y_{2i+1,2j+1}\mathbf{Y})$, ($0 \le k,l \le 3$) are the local covariances at the high resolution. Utilizing the similarity between the high-resolution covariances and the low-resolution covariances which couple the pair of pixels at the different resolution but along the same orientation, we can estimate covariances for HR from the pixels in LR as

$$\hat{\mathbf{R}} = \frac{1}{W^2}C^TC, \hat{\mathbf{r}} = \frac{1}{W^2}C^T\vec{\mathbf{y}} \quad , \tag{5}$$

where $\vec{\mathbf{y}} = [y_1...y_k...y_{W^2}]^T$ is the vector containing the $W \times W$ pixels inside a local window and C is a $W^2 \times 4$ matrix whose k-th row vector is the four 8-connected neighbors of y_k along the diagonal direction. According to (4) and (5), we get the interpolation coefficients.

$$\boldsymbol{\alpha} = (C^TC)^{-1}(C^T\vec{\mathbf{y}}) \quad . \tag{6}$$

By substituting (6) into (2), the interpolated value of $Y_{2i+1,2j+1}$ can be obtained. When interpolating the interlacing lattice $Y_{i,j}(i+j=odd)$ from the lattice $Y_{i,j}(i+j=even)$, the interpolation coefficients are obtained in the same way. From Fig.1 (a) and (b), we can see that they are isomorphic, if we rotate Fig.1 (a) $\pi/4$ clockwise and then down to a scaling factor of $2^{1/2}$. A detail interpretation can be found in [7].

2.2   Nonlocal Edge-directed Interpolation

In this section, we extend the AR interpretation to derive the nonlocal edge-directed interpolation algorithm by analyzing the AR interpolation coefficients. We take the interpolation of $Y_{2i+1,2j+1}$ as an example and the pixels $Y_{i,j}(i+j=odd)$ can be interpolated in the same way.

2.2.1   Covariances Estimation for HR
From (5), we can take low-resolution pixels with the same geometry configuration to estimate the high-resolution covariances. In a general situation, the covariance estimation $\hat{\mathbf{R}}$, $\hat{\mathbf{r}}$ and the coefficient $\vec{\boldsymbol{\alpha}}$ should be as follows.

$$\hat{\mathbf{R}}=\hat{E}(\mathbf{YY}^T)=\begin{bmatrix} [\vec{p}_1.*C(:,1)]^T C(:,1) & [\vec{p}_1.*C(:,1)]^T C(:,2) & [\vec{p}_1.*C(:,1)]^T C(:,3) & [\vec{p}_1.*C(:,1)]^T C(:,4) \\ [\vec{p}_1.*C(:,2)]^T C(:,1) & [\vec{p}_1.*C(:,2)]^T C(:,2) & [\vec{p}_1.*C(:,2)]^T C(:,3) & [\vec{p}_1.*C(:,2)]^T C(:,4) \\ [\vec{p}_1.*C(:,3)]^T C(:,1) & [\vec{p}_1.*C(:,3)]^T C(:,2) & [\vec{p}_1.*C(:,3)]^T C(:,3) & [\vec{p}_1.*C(:,3)]^T C(:,4) \\ [\vec{p}_1.*C(:,4)]^T C(:,1) & [\vec{p}_1.*C(:,4)]^T C(:,2) & [\vec{p}_1.*C(:,4)]^T C(:,3) & [\vec{p}_1.*C(:,4)]^T C(:,4) \end{bmatrix} \qquad (7)$$
$$=(\mathbf{P}.*C)^T C$$

$$\hat{\mathbf{r}}=\hat{E}(Y_{2i+1,2j+1}\mathbf{Y})=\left[[\vec{p}_1.*C(:,1)]^T \vec{\mathbf{y}} \quad [\vec{p}_1.*C(:,2)]^T \vec{\mathbf{y}} \quad [\vec{p}_1.*C(:,3)]^T \vec{\mathbf{y}} \quad [\vec{p}_1.*C(:,4)]^T \vec{\mathbf{y}}\right]^T \qquad (8)$$
$$=(\mathbf{P}.*C)^T \vec{\mathbf{y}}$$

$$\boldsymbol{\alpha} = [(\mathbf{P}.*C)^T C]^{-1}[(\mathbf{P}.*C)^T \vec{\mathbf{y}}] \qquad (9)$$

where $\mathbf{P}=[\vec{p}_1 \quad \vec{p}_1 \quad \vec{p}_1 \quad \vec{p}_1]$ and $\vec{p}_1$ is a vector whose components are composed of the probability of every sample of $\mathbf{Y}$ in C, $\vec{\mathbf{y}}=[y_1...y_k...y_{W^2}]^T$ are the samples of $Y_{2i,2j}$ and $C(:,i)$ represents the i-th column of C. The operator '.*' represents element-by-element multiplication of two matrix. Therefore, we can see that the prior estimation of high-resolution covariance is a special case that all the samples have the same probability, which is not always reasonable. The samples used to calculate coefficients in a local window should have the similar geometric structure with the region centered in the interpolated pixel. Otherwise, the geometric duality can not be satisfied in a local window where the geometric structure is different. In smooth or large scale edge region, where the samples in a small local window have equal probabilities may be reasonable, because all of them can reflect similar structures (i.e. edge direction). However, in the close-grained region with more different structures and different samples reflect different image structures, even opposite structure, i.e. Fig.2.

Fig. 2 Scheme of NLEDI strategy. Samples with similarity structure are assigned large probabilities, i.e. p1, p2; samples with disparity structure are assigned small probabilities, i.e. p3

From Fig 2, we can see that there are three kinds of samples in the local window *A*. First kind is non-edge sample. They do not include edge direction information. The second kind is edge on $45°$ direction (black edge) and the last one is edge on $135°$ (white edge). Although the interpolated pixel in *A* is on $45°$ direction, the edge information included in the covariance may vanish as a result of all samples with equal probabilities. Therefore, we proposed that samples with different structure should have different probabilities. If we can assign appropriate probabilities to different samples, it can not only improve the estimation of the high-resolution covariances, but also break the limit of a local window. This is because in a natural image, similar structure may have recurrence not only in local area but also other areas [10]. In next section, we set up a probability model just as [9].

2.2.2 Probability modeling

Reviewing the interpolation coefficients in (9), we can transform the equation in the following form,

$$\boldsymbol{\alpha} = [(\mathbf{P}_\bullet * C)^T C]^{-1}[(\mathbf{P}_\bullet * C)^T \vec{\mathbf{y}}] \Leftrightarrow (\mathbf{P}_\bullet * C)^T C\boldsymbol{\alpha} = (\mathbf{P}_\bullet * C)^T \vec{\mathbf{y}} \Leftrightarrow C\boldsymbol{\alpha} = \vec{\mathbf{y}} \quad . \tag{10}$$

From Equation (10), we can see that the AR interpolation coefficient $\boldsymbol{\alpha}$ should fits all the sample pixels in LR lattice. Therefore, weighted least square can be used to find an approximated solution to the overdetermined systems.

$$\boldsymbol{\alpha} = \min_{\boldsymbol{\alpha}} \left\| \vec{p}_1 {}_\bullet *(C\boldsymbol{\alpha} - \vec{\mathbf{y}}) \right\| \quad . \tag{11}$$

Where $\vec{p}_1$ is a probability vector just as prior section. Equation (11) expresses that $\boldsymbol{\alpha}$ minimizes the weighted sum of square errors between the true values and the interpolated values. Compared with structure in the region to be interpolated, the more similar the sample structure is, the better the geometry duality is. Therefore, the probability should reflect the similarity of structures between the samples and the interpolated region (we call it as *center region* below).

Since we introduce probability into the AR model, the samples should not necessary be limited in a small local window. More samples can be selected even in the whole images. Small probability can be assigned to the samples which have high disparities with the *center region*. Referring to Fig.2, the sample in E has different structure from the center region B, so the probability p_3 approximates to zero. In order to reflect the structure similarity, we compare the structures of two windows (we call them as *similarity windows* below, sized $m \times n$) centered the sample and the interpolated region respectively. The probability of the *j*-th sample of the *i*-th *center region* can be given by

$$p_{i,j} = \frac{1}{Z(i)} e^{-\frac{\left\| (W_{sj} - W_{ci})_\bullet * K \right\|_2}{h}} \quad . \tag{12}$$

$$Z(i) = \sum_{j} e^{-\frac{\left\| (W_{sj} - W_{ci}) \cdot * K \right\|_{2}}{h}} \qquad (13)$$

$$K(x, y) = \frac{1}{C} e^{-(\sqrt{(x-\frac{m}{2})^2 + (y-\frac{n}{2})^2})}, \, x \in [1, m], y \in [1, n] \qquad (14)$$

where W_{sj} and W_{ci} are matrices whose elements are pixel values in *similarity windows*, h is a constant to control the decay of the exponential function , C is a normalization constant. $K(x, y)$ is a weight value for each pixel in similarity window, which decrease with the distance from the center of *similarity window*. This is similar with the well-known Non-local means filter, which uses probability to represent the similarity between two pixels [9]. We use probability to measure the similarity between two samples. Furthermore, we also consider the distant effects on the structure similarity. Therefore, the difference between two similarity windows multiplies a weight matrix *K,* which decays with the Euclidean distance.

With introducing probability, samples are not limited in a small local window. We can select more samples in a larger window, even in the whole image. Thus, the solution of (9) will be more stable. However, In order to reduce the complexity of the algorithm, we also estimate covariances in a local window and only use the proposed interpolation method in edge areas. At the same time, in order to exclude the case that the matrix $(\mathbf{P} \cdot * C)^T C$ can not be invertible, AR coefficients are calculated only under the condition that its determinant is larger than a given threshold. In non-edge areas and regions where the determinant is too small, conventional *bilinear* interpolation method is good enough for pleasing result.

3 Experimental results

To verify the performance of the proposed NLEDI, extensive experiments were conducted with comparison to its predecessors, including two typical AR interpolation methods. They are the new edge-directed interpolation (NEDI) in [7], improved NEDI (INEDI) in [8] and the classical interpolation method *bilinear.*

In order to compare objective quality, we firstly downsample HR images every other pixel both in horizontal and vertical directions to get the input LR images, and then utilize different methods to improve the resolution up to original size. In order to validate our method, several images in different resolution are used for testing. The resolution of test images is as follows, Cameraman and Parrot (256x256), Lena and Baboon (512x512), Plane and Bike (768x510). Fig.3 depicts all these images. Table 1 tabulates the PSNR results. On all instances, the proposed NLEDI algorithm consistently ranks the first among all methods in terms of PSNR performance. The objective qualities, which result from NEDI, vary severely for different images. However, INEDI and our proposed method can be adaptive according to the content of images.

Table 1. PSNR values (dB) obtained on 2x enlarged images with different methods.

	bilinear	NEDI	INEDI	proposed
Cameraman	25.5137	25.4395	25.6623	25.9197
Parrot	32.4210	32.2567	32.2571	32.6752
Lena	33.4353	33.7021	34.0126	34.3756
Baboon	21.8706	21.7548	21.8815	22.0315
Plane	30.1370	28.7327	30.5982	30.8878
Bike	25.6324	25.6865	26.1371	26.3787

In Fig. 4 and Fig. 5, we compared results of our proposed method with *bilinear*, NEDI [7] and INEDI [8]. In Fig.4, We interpolate the original image into 2 times size. Fig.4.(a) is the original LR image (size of 200×200). The *bilinear* interpolation blurred the edges and produces many jags on edges in (b). NEDI, INEDI and our proposed methods can preserve edges very well in Fig.4 (c), (d) and (e) respectively. However, NEDI and INEDI also produce artifacts on edge in the red box, because the edge too thin and the AR coefficients with local structure can not reflect the edge significantly. The INEDI depress the artifacts in structure varying area through reducing the local window size. While our proposed method can also get visual pleasing results through the assignment of the different probabilities to different samples in local windows. In Fig.5, we interpolate the original image into 2 times size. We can see that *bilinear*, NEDI and INEDI all blur the lines in Fig.5 (b), (c) and (d), respectively. Because NEDI and INEDI assign equal probability to all samples and the samples on edge in local windows are not enough to preserve edge information in variances. Although INEDI can resize the window, it also induces to unstable solution when amount of samples are too small. Our proposed method gives visual pleasing results in Fig.5 (e) because we can use nonlocal samples with different probabilities.

In prior experiment, we also take account of the complexity. In Table 2, we list the running time for each method and the measurement is *second (s)*. The situation of our experiment is as follows, *matlab 7.0, intel Core2 Duo CPU and 3.25GB memory*. All the three AR interpolation methods have much higher complexity than *bilinear*. As a result of the probabilities calculation, our proposed method has higher complexity than NEDI, but has lower complexity than INEDI.

Table 2. Running time on 2x enlarged images with different methods.

	bilinear	NEDI	INEDI	proposed
Cameraman	0.0310	3.2500	46.5310	23.4060
Parrot	0.0320	3.2810	51.4840	24.6880
Lena	0.1410	14.3750	217.0780	114.3280
Baboon	0.1560	15.5470	300.3440	195.6720
Plane	0.2190	20.0310	209.6720	115.0940
Bike	0.2190	23.0310	392.5310	252.4540

4 Conclusion

In this paper, we present an analysis of AR-based interpolation methods and point out its defects. In particular, we introduce probabilities to each sample and make samples not limited in local area of the interpolated position. We propose a method that uses image structure similarity as probabilities of samples. The experimental results show that the proposed algorithm can not only enhance the objective quality of the interpolated images, but also improve the subjective quality significantly. However, the high complexity is still a problem, and we will try to tackle this problem in our further research work.

References

1. S. Battiato, G. Gallo, F. Stanco, "A locally-adaptive zooming algorithm for digital images," Image and Vision Computing, Vol. 20, No. 11, pp. 805-812, September 2002 1, 2
2. K. Jensen and D. Anastassiou, "Subpixel edge localization and the interpolation of still images," IEEE Trans. Image Processing, vol. 4, pp. 285-295, Mar. 1995.
3. S. Carrato, G. Ramponi, and S. Marsi, "A simple edge-sensitive image interpolation filter," in Proc. IEEE Int. Conf. Image Processing, vol. 3, 1996, pp. 711-714
4. B. S. Morse and D. Schwartzwald, "Isophote-based interpolation," in Proc. IEEE Int. Conf. Image Processing, vol. 3, 1998, pp. 227-231
5. Q. Wang and R. Ward, "A new edge-directed image expansion scheme," in Proc. IEEE Int. Conf. Image Processing, vol. 3, 2001, pp. 899-902
6. L. Zhang and X.L. Wu, "An edge-guided image interpolation algorithm via directional filtering and data fusion," IEEE Trans. Image Processing 15 (2006), pp. 2226–2238
7. X. Li and M. T. Orchard "New edge-directed interpolation," IEEE Trans. Image Process , vol. 10, pp. 1521, Oct. 2001
8. N. Asuni and A. Giachetti. "Accuracy improvements and artifacts removal in edge based image interpolation." In Proc. 3rd Int. Conf. Computer Vision Theory and Applications (VISAPP'08), 2008
9. A Buades,., B.Coll, , and J. M.Morel, (2005) "A non-local algorithm for image denoising," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2, pp 60–65 June.
10. M. Protter, M. Elad, H. Takeda, and P. Milanfar, "Generalizing the non-local-means to super-resolution reconstruction," IEEE Trans. Image Processing, Vol. 18, pp.36-51,Jan. 2009.
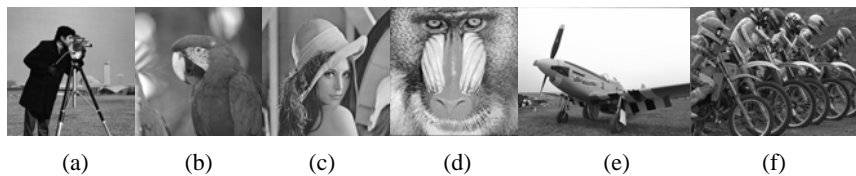
(a)   (b)   (c)   (d)   (e)   (f)

Fig. 3 test images for objective quality. (a) Cameraman, (b) Parrot, (c) Lena, (d) Baboon, (e) Plane, (f) Bike

Fig. 4 (a) original image, (b) *bilinear* interpolation result, (c) NEDI result, (d) INEDI result, (e) proposed method result
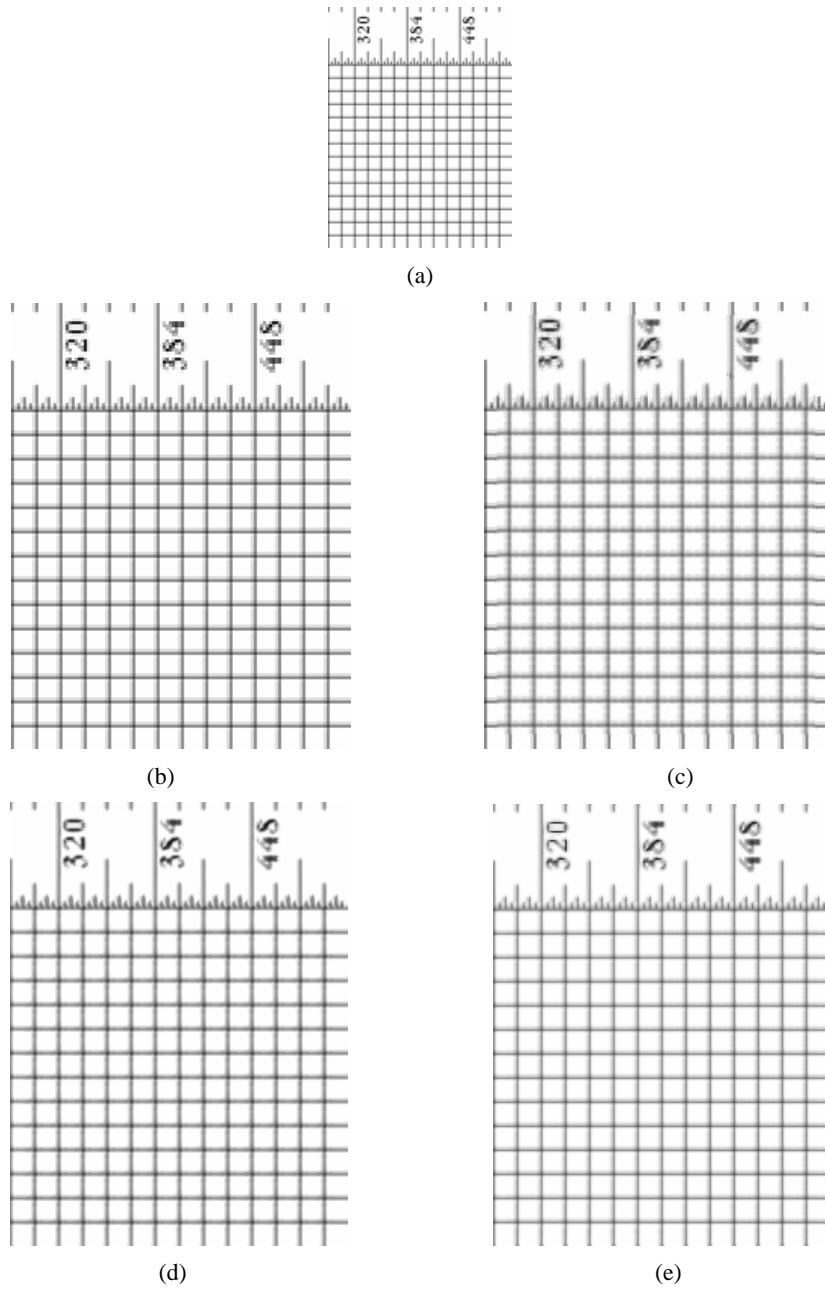
Fig. 5 (a) original image, (b) *bilinear* interpolation result, (c) NEDI result, (d) INEDI result, (e) proposed method