

# GENERAL SCALE INTERPOLATION VIA CONTEXT-AWARE AUTOREGRESSIVE MODEL AND MULTIPLANAR CONSTRAINT

Shihong Deng, Jiaying Liu\*, Mading Li, Wenhan Yang and Zongming Guo

Institute of Computer Science and Technology, Peking University, Beijing, P. R. China, 100871

## ABSTRACT

In this paper, we propose a novel image interpolation algorithm suitable for general scale enlargement. Different from previous AR-based interpolation algorithms which employ predetermined reference configuration to predict pixel values, we consider the context information when building AR models. Optimal references are selected by incorporating nonlocal-based correlation coefficient and the indicator for local edge direction. Furthermore, the multiplanar constraint among similar patches is applied to enhance the correlation within the estimation window and serves as a kind of supplement to data fidelity term in AR model. The experimental results show that our method is effective in several enlargement scales and successfully alleviate the artifacts nearby edges and preserve their sharpness. The comparison experiments demonstrate that the proposed method can obtain desirable performance in terms of both objective and subjective results.

**Index Terms**— Autoregressive (AR), context modeling, interpolation, general scale, multiplanar

## 1. INTRODUCTION

Image interpolation refers to generating high resolution (HR) image utilizing the information of the low resolution (LR) image. The technique has been widely studied and become a hot research topic in image processing area. Its applications range from video communication, digital photograph enhancement to medical analysis.

The most essential task for image interpolation is to choose a proper model that can easily characterize the image texture and edge information. Accordingly, interpolation algorithm can be roughly classified into three categories: polynomial-based, edge-directed and learning-based methods. Polynomial-based methods, such as Bilinear, Bicubic [1] and Cubic Spline [2] estimate the missing pixels by convolving neighbouring pixels within a predetermined kernel. This kind of method is easy to implement and costs low computational resources, but pixels are treated identically ignoring the various local structures. Therefore, these methods lead to unpleasant results. Since the human visual system is sensitive to edges, many edge-directed interpolation algorithms are published. In studies [3] and [4], explicit edge information such as isophotes is calculated, and used as guidance for interpolation. To avoid the difficulty of detecting edge direction, AR model is developed using statistics to characterize edge structure. Li and Orchard [5] proposed a new edge directed interpolation method (NEDI). It exploited the geometric duality between LR covariance and HR covariance to estimate

HR pixels. Based on NEDI, Zhang and Wu [6] introduced a soft-decision adaptive interpolation algorithm (SAI). An extra AR model was added and HR pixels in a local window were estimated by solving a least-square problem. In [7], similarity modulated AR model was proposed to solve the inconformity in the local window.

Although AR-based interpolation algorithms can obtain good results, there still exists limitation that these algorithms can only be applied in  $2\times$  or  $2^m\times$  enlargement problem. With the rapid development of mobile phone and smart devices, resolution of video and image differs greatly between devices and the arbitrary scaling method is needed. To address this problem, Wu *et al.* [8] proposed an adaptive resolution up-conversion method providing support for general scale enlargement. Our previous work [9] proposed an adaptive general scale interpolation (AGSI) for the same reason. By reconstructing the AR model and adding a weighting map to the model, AGSI was proved to be a method of state-of-the-art performance. However, the aforementioned AR-based methods employ fixed model references such as pixels in diagonal-direction or cross-direction ignoring complicated and anisotropic dependencies in natural images. Some learning-based methods [10][11] can provide arbitrary scale enlargement, but their training processes are very time-consuming. Also, the results of learning-based methods significantly depend on sufficient similarity patterns in external database or the image itself.

In this paper, to overcome the drawbacks of existing AR-based interpolation algorithms, we proposed a novel general scale image interpolation algorithm. Firstly, we modify the fixed spatial reference configuration of traditional AR model by context-awareness. Using the *context-patch* distance, similar patches are collected to calculate the correlation between center position and candidate references. Collaborated with directional indicators which are designed to measure the direction of the edge, the AR model is reconstructed by optimal references. Then, multiplanar constraint among patches are introduced as supplement to data fidelity. Finally, as HR pixels and model parameters are unknown, structured total least-square is used to solve the problem in an iterative manner. Experimental results demonstrate that our method is successful in modelling the piecewise stationary assumption of natural images and achieves desirable performance from both objective and subjective perspectives.

The rest of the paper is organized as follows: Section 2 gives a brief review of AR model and shortly introduces the concept of context. Section 3 describes the novel general scale interpolation algorithm in detail. Experimental results and analysis are presented in Section 4. Finally, Section 5 concludes and remarks this paper.

## 2. CONTEXT AUTOREGRESSIVE (AR) MODEL

### 2.1. Autoregressive (AR) Model

In statistics and signal processing, an autoregressive (AR) model is a representation of a type of random process. Owing to its ability to

\*Corresponding author

This work was supported by National Natural Science Foundation of China under contract No. 61671025 and National Key Technology R&D Program of China under Grant 2015AA011605.

well formulate stochastic structure of sequential data, AR model is often utilized to model and predict various types of structural signals. Typically, a 2-D image signal  $X(m, n)$  can be modeled as an AR process as follows,

$$X(m, n) = \sum_{(i,j) \in \Omega} \varphi(i, j)X(m+i, n+j) + \sigma, \quad (1)$$

where  $\Omega$  and  $\varphi(i, j)$  represent adjacent neighbours of pixel  $X(m, n)$  and their corresponding weights (parameters of the AR model), respectively;  $\sigma$  is the estimation error. The formula attempts to characterize local structures around the unknown pixel and use the information to predict its value.

$2\times$  enlargement interpolation methods in [6][7] used two sets of AR model to predict missing HR pixels utilizing the geometric duality between HR and LR pixels. Later, [9] generalized the method to arbitrary scales by using HR pixels themselves as reference because variations of enlargement scale make it hard to maintain the geometric duality between HR and LR pixels. The model is formulated as follows

$$\min_{\mathbf{y}, \mathbf{a}, \mathbf{b}} \{ \alpha \|\mathbf{y}_c - \mathbf{A}\mathbf{y}\|^2 + \beta \|\mathbf{y}_c - \mathbf{B}\mathbf{y}\|^2 + \lambda \|\mathbf{x} - \mathbf{D}\mathbf{y}_c\|_2^2 \}, \quad (2)$$

where vector  $\mathbf{y}$  consists of pixels in a local windows  $W$  of the input HR image and  $\mathbf{y} = \begin{bmatrix} \mathbf{y}_c \\ \mathbf{y}_b \end{bmatrix}$ , vector  $\mathbf{y}_b$  refers to pixels on the boundaries of  $W$  while vector  $\mathbf{y}_c$  refers to others; matrix  $\mathbf{A}$  and  $\mathbf{B}$  are constructed by parameters  $\mathbf{a}$  and  $\mathbf{b}$  of corresponding AR model respectively; and matrix  $\mathbf{D}$  represents the Bicubic down-sampling process. Since  $\mathbf{y}$  and model parameters are both unknown, a structural total least-square based iterative process is used to solve the problem.

## 2.2. Context Modeling Prediction

Traditional piecewise AR models including the method mentioned above usually fix their model order and spatial reference empirically and the configuration remains the same throughout the estimation process. There are two drawbacks of the kind of AR model. First, the spatial reference selecting approach may not pick the optimal configuration. A predetermined reference configuration is very likely to be suboptimal because it may take adjacent but irrelevant pixels into consideration and miss further but relevant ones in model estimation, thus leads to estimation bias. Second, a unified reference configuration does not fit for all pixels which belong to different regions with different patterns and structures. To overcome these drawbacks, attempts are made to model the context awareness for AR model. In [12], Wu *et al.* proposed a method utilising correlation to choose both predictor reference and training set of AR model for every unknown pixel. The new technique is applied in lossless image coding and achieves superb performance. In this paper, we develop a more complex context modelling predictor for traditional AR model based on [12]. The detail will be presented in Section 3.1

## 3. THE GENERAL SCALE INTERPOLATION ALGORITHM

In this section, we present an algorithm for general scale image interpolation. First, a novel AR model based on context-aware modification is described, *context-patch* distance and a deliberately designed directional indicator are used to help determine the optimal model reference. Then, a multiplanar constraint exploiting nonlocal similarity is incorporated into AR model to enhance the correlation

between pixels. Finally, our generalized interpolation algorithm is performed in detail.

### 3.1. Context-Aware Adaptive Prediction

In order to better model local image structures, we extend the fixed spatial reference configuration of AR model to a changeable ones. By considering *context* of the missing pixels, we choose the proper reference configuration for them individually. In this way, we can make the AR prediction more accurate and stable.

#### 3.1.1. Context Based Patch Distance

Different with traditional AR model, we replace the fixed neighbouring references with a candidate reference set denoted as  $\Phi(y_i) = \{\phi_1(y_i), \phi_2(y_i), \dots, \phi_M(y_i)\}$ , where  $y_i$  is the pixel to be estimated.  $W_r$  represents the window containing all possible candidate references. The critical problem is how to determine the correlation levels between candidate positions and the central position where  $y_i$  lies. To solve this problem, we firstly collect a set of patches that are similar to  $W_r(y_i)$ :

$$T = \{y \mid \|d(W_r(y), W_r(y_i))\| \leq \tau_T\}, \quad (3)$$

where  $d(\cdot, \cdot)$  measure the distance between two patches and  $\tau_T$  is the threshold. Since most of the HR pixels are unknown, their values are approximated by the Bicubic method for initialization.

In order for more accurate patch matching, *context-patch* distance proposed in [13] is used. This method benefits both advantages of the small patch and the large patch by concatenate the regular content of a small patch with the compact representation of its large surroundings. First, similarity weights between central patch  $\mathbf{x}_i$  and its surrounding patches  $\mathbf{x}_j \in N(i)$  are measured:

$$w_{ij} = \exp \left\{ -\frac{\|\mathbf{x}_i - \mathbf{x}_j\|_2^2}{2\sigma^2} \right\}, \forall j \in N(i). \quad (4)$$

A larger  $w_{ij}$  indicates two patches are similar while a smaller one indicates the two are basically different. Then, these similarity weights are rearranged into a histogram  $\mathcal{H}$  of  $b$ -bins (*e.g.*  $b = 8$ ). In our practice, the size of small patch and its surrounding windows are set to  $5 \times 5$  and  $11 \times 11$ , respectively. After that, we define the distance function by combining the distance of small patch and corresponding histogram  $\mathcal{H}$  together

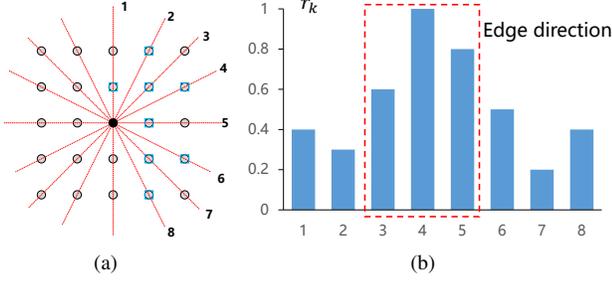
$$d(\mathbf{x}_i, \mathbf{x}_j) = \|\mathbf{x}_i - \mathbf{x}_j\|_2^2 + \|\mathcal{H}(\mathbf{x}_i) - \mathcal{H}(\mathbf{x}_j)\|_2^2. \quad (5)$$

With Eq. (3) and Eq. (5), we obtain the similar patch set  $T$ . Afterwards, using these similar patches which are supposed to have homogeneous patterns, correlation coefficients between central position and candidate references positions are calculated as follow

$$\rho_{\phi_m} = \frac{\sum_{y \in T} \phi_m(y)y - \frac{1}{|T|} \sum_{y \in T} y \sum_{y \in T} \phi_m(y)}{\sqrt{\sum_{y \in T} y^2 - \frac{1}{|T|} (\sum_{y \in T} y)^2} \sqrt{\sum_{y \in T} \phi_m(y)^2 - \frac{1}{|T|} (\sum_{y \in T} \phi_m(y))^2}}. \quad (6)$$

#### 3.1.2. Direction Guided Model Reference

Besides correlation coefficients, we also introduce a directional statistics indicator to help decide the model reference. For each pixel position in  $W_r$ , we consider the directions given by the lines joining the central position and themselves. Indexes of these directions are shown in Fig. 1(a).



**Fig. 1.** (a) 8 directions. Black dot is the central pixel. Pixels in blue square are nearest neighbour for central pixel. (b) Values of  $r_k$

Direction  $k$  is represented by a vector  $V = (h_k, v_k)$  where  $h_k$  and  $v_k$  are the horizontal and vertical distance between central position and its nearest neighbour in corresponding direction, respectively. Let pixel values in HR image be denoted as  $H(i, j)$ , differential values of neighbouring pixels at each direction are collected,

$$\Delta\mathbf{I}(h_k, v_k) = \left\{ \frac{H(i, j) - H(i + h_k, j + v_k)}{\sqrt{h_k^2 + v_k^2}} \right\}. \quad (7)$$

Let  $\mu_k$  and  $\sigma_k^2$  represent the mean and variance of  $\Delta\mathbf{I}(h_k, v_k)$ , respectively. The directional indicator for direction  $k$  is obtained as

$$r_k = \frac{1}{\log_2(\sigma_k^2 + \mu_k^2)}. \quad (8)$$

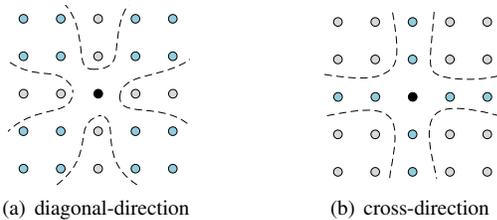
For normalization,  $r$  is scaled or shifted to make its value within  $(0, 1]$ . The indicator is significant, only when both variance and mean of  $\Delta\mathbf{I}(h_k, v_k)$  are small, which means that edge is along the direction and pixels in this direction are of large continuity. We can put more importance on pixels in directions with a large  $r$  value. Our ranking mechanism combines correlation coefficient and directional indicator together and is shown below

$$R_{\phi_m} = \frac{\rho_{\phi_m} \cdot r_k}{|\phi_m|}, \quad (9)$$

where  $k$  indicates the the direction  $\phi_m$  belongs to and  $|\phi_m|$  refers to the distance between  $\phi_m$  and central position. With Eq. (9), we can reorder the candidate reference set as follows

$$\Phi(y_i) = \{\phi_1(y_i), \phi_2(y_i), \dots, \phi_M(y_i), R_{\phi_1} \geq R_{\phi_2} \geq \dots R_{\phi_M}\}. \quad (10)$$

Since we have two sets of AR model, we divide  $W_r$  into two parts as shown in Fig.2. Correspondingly, candidate references are divided into two subsets  $\Phi^d$  and  $\Phi^c$ . Those candidates with high rankings are selected as reference in corresponding AR model. Thus,  $\mathbf{a} = \{a_{\phi_1^d}, \dots, a_{\phi_P^d}\}$ ,  $\mathbf{b} = \{b_{\phi_1^c}, \dots, b_{\phi_Q^c}\}$ .  $P$  and  $Q$  are the numbers of selected references in corresponding AR model and obtained by satisfy two conditions:  $R_{\phi_P^d}(R_{\phi_Q^c}) \geq \tau_R$  and if  $P(Q) < 4$ ,  $P(Q) = 4$ .



**Fig. 2.** Two sets of AR model. Black dot represents the central pixel while blue dots represents candidate references of AR model.

### 3.2. Multiplanar Constraint

When interpolation operation is performed in  $W$ , pixel values are constrained by their neighbouring reference pixels. Generally, there is a magnitude difference between  $W$  and  $W_r$ . Thus, the constraint imposed by AR parameters merely presents a small part of relations between pixels. Due to the abundant similarities in natural images, there exists plenty nonlocal correlations that can be used to get a more promising interpolation results. Also, the least-square problem for AR model makes it easy to integrate other constraints.

Here, we introduce the multiplanar constraint to utilize larger scale correlations within  $W$ . Similar patches, scales of which are approximately the same as  $W_r$ , are collected. The distance function here is defined as follows:

$$dis(\mathbf{y}_j, \mathbf{y}_i) = \|\mathbf{y}_j - \mathbf{y}_i\|_2^2 + \eta \|\nabla \mathbf{y}_j - \nabla \mathbf{y}_i\|_2^2 \quad (11)$$

where  $\nabla$  denotes the gradient operator and  $\eta$  is a parameter used to balance the contribution of two terms. Based on Eq. (11), a similar patch set compared with central patch is collected in  $W$

$$S = \left\{ \mathbf{x} \mid \exp \left\{ \frac{-dis(\mathbf{x}, \mathbf{x}_i)}{\alpha} \right\} \geq \tau_S, \mathbf{x} \in W \right\}, \quad (12)$$

where  $\mathbf{x}_i$  represents the central patch in  $W$  and  $\tau_S$  is a threshold and  $\alpha$  is the parameter to control the shape of the exponential function. When all the similar patches are stacked together, HR pixels in same position should be alike with each other. Thus, we construct the multiplanar constraint term as follow

$$\sum_{m \in S} \sum_{i \in P_m} \left( y_i^{P_m} - \bar{y}_i \right)^2, \quad (13)$$

where  $P_m$  and  $y_i^{P_m}$  denote the  $m$ th patch in  $S$  and the  $i$ th pixels in the corresponding patch, respectively;  $\bar{y}_i$  represents the average pixel values at the  $i$ th position of the patch.

### 3.3. The Generalized Interpolation Algorithm

By adding the weighting map  $\mathbf{W}$  to Eq. (2), combining with Eq. (13) we can get the complete version of objective function

$$\min_{\mathbf{y}, \mathbf{a}, \mathbf{b}} \{ \alpha \|\mathbf{W}(\mathbf{y}_c - \mathbf{A}\mathbf{y})\|^2 + \beta \|\mathbf{W}(\mathbf{y}_c - \mathbf{B}\mathbf{y})\|^2 + \lambda_1 \|\mathbf{x} - \mathbf{D}\mathbf{y}_c\|_2^2 + \lambda_2 \|\mathbf{E}\mathbf{y}\|_2^2 \}, \quad (14)$$

where  $\mathbf{A}$  and  $\mathbf{B}$  are constructed by parameter  $\mathbf{a}$  and  $\mathbf{b}$  which are defined in Section 3.1. Our previous works [7] and [9] analyze the reason and effect of using weighting map in AR model. And the weighting method in [9] is used here directly.

The objective function can be represented by a least-square problem as

$$\min_{\mathbf{y}, \mathbf{a}, \mathbf{b}} \|\mathbf{R}(\mathbf{y}, \mathbf{a}, \mathbf{b})\|^2, \quad (15)$$

$$\mathbf{R}(\mathbf{y}, \mathbf{a}, \mathbf{b}) = \begin{bmatrix} \sqrt{\alpha} \mathbf{W}(\mathbf{y}_c - \mathbf{A}\mathbf{y}) \\ \sqrt{\beta} \mathbf{W}(\mathbf{y}_c - \mathbf{B}\mathbf{y}) \\ \sqrt{\lambda_1} (\mathbf{x} - \mathbf{D}\mathbf{y}_c) \\ \sqrt{\lambda_2} \mathbf{E}\mathbf{y} \end{bmatrix}. \quad (16)$$

To solve the problem, the structure total least-square solution are used. Let  $\Delta\mathbf{y} = \begin{bmatrix} \Delta\mathbf{y}_c \\ \Delta\mathbf{y}_b \end{bmatrix}$ ,  $\Delta\mathbf{a}$  and  $\Delta\mathbf{b}$  be the small changes in  $\mathbf{y}$ ,  $\mathbf{a}$  and  $\mathbf{b}$ , respectively. The pixels on the boundaries are kept invariant for the purpose of better constraining the pixels in  $W$ . Thus,  $\Delta\mathbf{y}_b$  equals to zero.

After the matrix transformation and operation, we reduce the problem into a convenient representation

$$\min_{\Delta \mathbf{y}, \Delta \mathbf{a}, \Delta \mathbf{b}} \|R(\mathbf{y}, \mathbf{a}, \mathbf{b}) - \mathbf{C} \cdot \Delta R\|^2, \quad (17)$$

given

$$\mathbf{C} = \begin{bmatrix} \sqrt{\alpha} \mathbf{W}(-\mathbf{I} + \mathbf{A}_c) & \mathbf{F}_1 & \mathbf{0} \\ \sqrt{\beta} \mathbf{W}(-\mathbf{I} + \mathbf{B}_c) & \mathbf{0} & \mathbf{F}_2 \\ \sqrt{\lambda_1} \mathbf{D} & \mathbf{0} & \mathbf{0} \\ -\sqrt{\lambda_2} \mathbf{E} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$\Delta R = [\Delta \mathbf{y}_c, \Delta \mathbf{a}, \Delta \mathbf{b}]^T.$$

Hence, given the initial values of  $\mathbf{y}$ ,  $\mathbf{a}$  and  $\mathbf{b}$ , we can calculate  $\Delta R$  and use it to update  $\mathbf{y}$ ,  $\mathbf{a}$  and  $\mathbf{b}$  for next iteration. The initial values of  $\mathbf{a}$  and  $\mathbf{b}$  are set to  $(\frac{1}{P}, \frac{1}{P}, \dots, \frac{1}{P})$  and  $(\frac{1}{Q}, \frac{1}{Q}, \dots, \frac{1}{Q})$  in our implementation. Besides, coefficients  $\alpha$ ,  $\beta$ ,  $\lambda_1$  and  $\lambda_2$  are set to 0.2, 0.3, 0.5, 0.1 empirically. The iterative process is of high computational expense. We only apply the proposed method on high-frequency regions. Moreover, context selecting process is performed only when top three directional indicators introduced in Section 3.1.2 are adjacent to each other as shown in Fig.1(b), which means there is a significant edge in the local window and our algorithm works well for regions of this kind. Also, we output the central  $3 \times 3$  patch for each estimation. Although performance may drop slightly, it accelerates the processing speed 9 times and makes the proposed algorithm more effective.

#### 4. EXPERIMENTAL RESULTS

The proposed interpolation is implemented on MATLAB 8.6 platform. To evaluate the general performance of our method, we choose 1.5 and 1.7 as the enlargement scales and the proposed algorithm is compared with Bicubic, Wu's Work [8] and our previous work AGSI [9]. In addition, the proposed algorithm is compared with state-of-the-art dedicated  $2\times$  interpolation algorithms, such as SAI [6] and SAGA [4] and AGSI is also used as comparison. Our interpolation method was tested in a large amount of images. The testing images are selected from Kodak and USC-SIPI image databases.

For an enlargement scale  $s$ , we generate the LR image by down-sampling the original HR image with the scale of  $1/s$  first. And then different methods are applied to generate the HR images from the LR images. Peak Signal-to-Noise Ratio (PSNR) is used as the objective criterion to evaluate the experiment results.

**Table 1.** PSNR(dB) results of different methods at arbitrary scales

Images	Scale	Bicubic	Wu's	AGSI	Proposed
Cameraman	1.7	27.64	26.39	28.68	<b>28.79</b>
House		24.68	22.73	24.99	<b>25.18</b>
LightHouse		29.22	27.64	29.80	<b>29.84</b>
Monarch		34.27	31.72	35.20	<b>35.46</b>
Lena	1.5	37.67	35.61	<b>38.29</b>	38.28
Baboon		26.19	24.72	26.50	<b>26.69</b>
Bike		29.94	27.57	31.26	<b>31.47</b>
House		26.23	23.60	27.00	<b>27.13</b>
Average		29.48	27.51	30.22	<b>30.36</b>

Results of  $1.5\times$ ,  $1.7\times$  and  $2\times$  enlargement are shown in Table 1 and Table 2. From Table 1, we can conclude that the proposed interpolation algorithm achieves better performance in general scales interpolation compared with previous works and is suitable for general enlargement. Furthermore,  $2\times$  enlargement results shown in

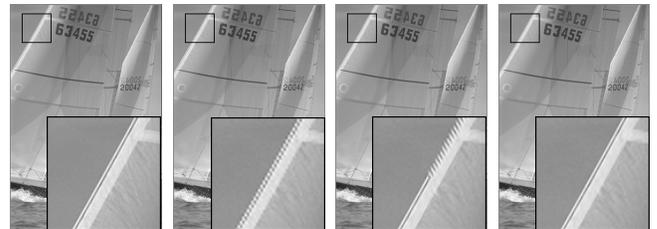
**Table 2.** PSNR(dB) results of different methods,  $2\times$  enlargement

Images	Bicubic	SAI	SAGA	AGSI	Proposed
Child	35.49	<b>35.63</b>	35.22	35.49	35.62
Lena	34.01	<b>34.76</b>	34.42	34.49	34.58
Cameraman	25.51	<b>25.99</b>	25.98	25.55	25.83
Airplane	29.40	29.62	29.02	29.87	<b>29.94</b>
Status	31.36	<b>31.78</b>	31.76	31.35	31.49
Sailboat	30.12	30.69	30.56	30.39	<b>30.80</b>
Bike	25.41	26.28	25.90	25.85	<b>26.07</b>
LightHouse	26.97	26.70	<b>27.23</b>	27.13	<b>27.23</b>
Average	29.67	29.84	29.80	29.69	<b>29.86</b>

Table 2 implies that our method is competitive with state-of-the-art interpolation method which are well designed for  $2\times$  enlargement.



(a)  $1.7\times$  enlargement on *Cameraman*



(b)  $2\times$  enlargement on *Sailboat*

**Fig. 3.** Subjective image quality comparison. From left to right: original image, Bicubic, AGSI, proposed method.

Specifically, the proposed method presents more desirable performance than other methods on long, sharp edges and large scale textures. Subjective image quality is demonstrated in Fig.3. In general scale (Fig.3(a)) and  $2\times$  enlargement (Fig.3(b)), Bicubic method produces evident zigzag artifacts, while AGSI maintains sharp edge in most places but fails on other parts with jags and ringing. Our method not only preserves the sharpness but also the continuity of edges, thus acquires better visual quality.

#### 5. CONCLUSION

In this paper, we present a novel general scale interpolation algorithm employing context-aware modification to AR model. By selecting the proper reference according to context information from a relatively large region, a more reliable AR model is built. It excludes some reference pixels that are irrelevant with prediction so as to reduce the noise. Meanwhile, it includes some closely related reference pixels that are ignored in traditional models due to their long distance from central pixels, increasing the model precision and stability. Also, we design the multiplanar constraint among similar patches as a kind of supplement to data fidelity, which helps to preserve more structural information. The experimental results show that the proposed algorithm works well for edges and achieves best performance in several enlargement scales.

## 6. REFERENCES

- [1] Robert G Keys, "Cubic convolution interpolation for digital image processing," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 29, no. 6, pp. 1153–1160, 1981.
- [2] Hsieh S Hou and H Andrews, "Cubic splines for image interpolation and digital filtering," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 26, no. 6, pp. 508–517, 1978.
- [3] Qing Wang and Rabab Kreidieh Ward, "A new orientation-adaptive interpolation method," *Image Processing, IEEE Transactions on*, vol. 16, no. 4, pp. 889–900, 2007.
- [4] Christine M Zwart and David H Frakes, "Segment adaptive gradient angle interpolation," *Image Processing, IEEE Transactions on*, vol. 22, no. 8, pp. 2960–2969, 2013.
- [5] Xin Li and Michael T Orchard, "New edge-directed interpolation," *Image Processing, IEEE Transactions on*, vol. 10, no. 10, pp. 1521–1527, 2001.
- [6] Xiangjun Zhang and Xiaolin Wu, "Image interpolation by adaptive 2-d autoregressive modeling and soft-decision estimation," *Image Processing, IEEE Transactions on*, vol. 17, no. 6, pp. 887–896, 2008.
- [7] Jie Ren, Jiaying Liu, Wei Bai, and Zongming Guo, "Similarity modulated block estimation for image interpolation," in *Image Processing (ICIP), 2011 18th IEEE International Conference on*. IEEE, 2011, pp. 1177–1180.
- [8] Xiaolin Wu, Mingkai Shao, and Xiangjun Zhang, "Improvement of h. 264 svc by model-based adaptive resolution upconversion," in *Image Processing, 2010 IEEE International Conference on*. IEEE, 2010, pp. 4205–4208.
- [9] Mading Li, Jiaying Liu, Jie Ren, and Zongming Guo, "Adaptive general scale interpolation based on weighted autoregressive models," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 25, no. 2, pp. 200–211, 2015.
- [10] Weisheng Dong, Lei Zhang, Rastislav Lukac, and Guangming Shi, "Sparse representation based image interpolation with nonlocal autoregressive modeling," *Image Processing, IEEE Transactions on*, vol. 22, no. 4, pp. 1382–1394, 2013.
- [11] Yaniv Romano, Matan Protter, and Michael Elad, "Single image interpolation via adaptive nonlocal sparsity-based modeling," *Image Processing, IEEE Transactions on*, vol. 23, no. 7, pp. 3085–3098, 2014.
- [12] Xiaolin Wu, Guangtao Zhai, Xiaokang Yang, and Wenjun Zhang, "Adaptive sequential prediction of multidimensional signals with applications to lossless image coding," *Image Processing, IEEE Transactions on*, vol. 20, no. 1, pp. 36–42, 2011.
- [13] Yaniv Romano and Michael Elad, "Con-patch: When a patch meets its context," *Image Processing, IEEE Transactions on*, vol. 25, no. 9, pp. 3967–3978, 2016.