STRUCTURE PRESERVING IMAGE INTERPOLATION VIA ADAPTIVE 2D AUTOREGRESSIVE MODELING

Xiangjun Zhang and Xiaolin Wu

Department of Electrical & Computer Engineering, McMaster University
Hamilton, ON, Canada L8S 4K1; zhangxj@grads.ece.mcmaster.ca/xwu@ece.mcmaster.ca

ABSTRACT

The performance of image interpolation depends on an image model that can adapt to nonstationary statistics of natural images when estimating the missing pixels. However, the construction of such an adaptive model needs the knowledge of every pixels that are absent. We resolve this dilemma by a new piecewise 2D autoregressive technique that builds the model and estimates the missing pixels *jointly*. This task is formulated as a non-linear optimization problem. Although computationally demanding, the new non-linear approach produces superior results than current methods in both PSNR and subjective visual quality. Moreover, in quest for a practical solution, we break the non-linear optimization problem into two subproblems of linear least-squares estimation. This linear approach proves very effective in our experiments.

Index Terms— Image interpolation, autoregressive process, optimization, soft decision.

1. INTRODUCTION

With ever increasing computation power in image and video processing, more sophisticated adaptive image interpolation methods were proposed in recent years. To preserve directional information, Li and Orchard proposed a technique of estimating the covariance of high resolution image from the covariance of the low resolution image, and a Wiener-filtering like interpolation scheme based on the estimated covariance [1]. Muresan and Parks [2] cast the method of [1] into the light of adaptive optimal recovery, and proposed a general approach to image interpolation.

The reproduction quality of any image interpolation algorithm primarily depends on its adaptability to varying pixel structures across an image, which is the central theme of this paper. In fact, modeling of non-stationarity of image signals is a common challenge facing many image processing tasks, such as compression, restoration, denoising, and enhancement. We had a measured success in this regard in a research on predictive lossless image compression [3]. In that work a natural image is modeled as a piecewise 2D autoregressive process. The model parameters are estimated on the fly for each pixel using sample statistics of a local window, assuming that the image is piecewise stationary. In this work we take a similar approach to image interpolation. An obvious difference is in that the sample set for parameter estimation has to be causal to the current pixel for compression, but does not need to be for interpolation, which is to the advantage of the latter. On the other hand, for image interpolation the fit of the model to true sample statistics is made far more difficult by the fact that only a low-resolution version of the original can be observed.

Now the problem on hand is one of chicken and egg. Correct interpolation of missing pixels relies on a good model, whereas the model parameters can be reliably estimated only if the missing pixels are known. The main innovation of this work is a new image interpolation technique that combines the two tasks of estimating the model parameters and interpolating the missing pixels with the estimated model. The joint estimation aims to achieve the maximum possible statistical agreement between estimated model parameters and the interpolated pixels, constrained by known low resolution pixels. This idea is formulated into a non-linear optimization problem with model parameters and missing pixels both as variables. By solving the optimization problem, we obtained high resolution images of quality superior to the best of the existing techniques.

2. PIECEWISE STATIONARY AUTOREGRESSIVE MODEL

For the purpose of adaptive image interpolation, we model the image as a piecewise autoregressive (PAR) process:

$$X(i,j) = \sum_{(m,n)\in T} a(m,n)X(i+m,j+n) + v_{i,j} \quad (1)$$

where T is the spatial template for the regression operation. The term $v_{i,j}$ is a random perturbation independent of spatial location (i,j) and the image signal, and it accounts for both fractal-like fine details of image signal and measurement noise. The validity of the PAR model henges on a mechanism that adjusts the model parameters $a(m,n)$ to local pixel structures. The fact that any semantically meaningful image constructs, such as edges and surface textures, are formed by spatially coherent contiguous pixels, suggests piecewise statistical stationarity of the image signal. In other words, in the setting of the PAR model, the parameters $a_{m,n}$ remain constant in a small locality, although they may and often do vary significantly in different segments of a scene.

The validity of the PAR model with locally adaptive parameters is corroborated by the success of this modeling technique in lossless image compression. Among all known lossless image coding methods, including CALIC, TMW [4], and invertible integer wavelets, those that employ the PAR model with adjusted parameters on a pixel-by-pixel basis have delivered the lowest lossless bit rates [3,5]. In the principle of Kolmogorov complexity, the true model of a stochastic process is the one that yields the minimum description length. Thus we have strong empirical evidence to support the appropriateness and usefulness of the PAR model for natural images.

3. NON-LINEAR OPTIMIZATION APPROACH TO IMAGE INTERPOLATION

Let I_h be the high resolution (HR) image to be estimated by interpolating the low resolution (LR) image I_l observed. The LR image I_l is a down sampled version of the HR image I_h by a factor of two. Let $x_i \in I_l$ and $y_i \in I_h$ be the pixels of images I_l and I_h respectively. We write the neighbors of pixel location i in the HR image as y_{i-t}, $t = 1, 2, \cdots$. Since $x_i \in I_l$ implies $x_i \in I_h$, we also write an HR pixel $y_i \in I_h$ as x_i (likewise, y_{i-t} as x_{i-t}) when it is in the LR image, $y_i \in I_l$, as well.

For operational reasons to be self-evident shortly, we interpolate the missing HR pixels in two passes. In the first pass, we interpolate those HR pixels $y_i \in I_h$ whose four 8-connected neighbors are known LR pixels $x_{i-t} \in I_l$, $t = 1, 2, 3, 4$. This configuration is depicted in Fig. 1(a). Using the PAR image model introduced in the preceding section, we pose the interpolation problem as one of non-linear optimization:

$$\min_{\boldsymbol{a},\boldsymbol{y}} \left\{ \sum_{i \in W} \left\| y_i - \sum_{t=1}^{4} a_t x_{i-t} \right\| + \sum_{i \in W} \left\| x_i - \sum_{t=1}^{4} a_t y_{i-t} \right\| \right\}. \tag{2}$$

Fig. 1(b) depicts the sample relationships involved in (2).

Now let us explain the rationale of (2). To apply the PAR model to interpolation, we face the challenge of having to estimate the model parameters $\boldsymbol{a} \in \Re^4$ from an incomplete data set (three-quarters of the pixels are missing). On one hand, the interpolation needs a good model that fits the true HR data. On the other hand, the model parameters can be reliably estimated only if the missing HR pixels are known. These two interdependent tasks of estimating the model parameters and interpolating the HR pixels with the estimated model present a chicken-and-egg dilemma. We circumvent the dilemma by treating model parameters $\boldsymbol{a}$ and the missing HR pixels $\boldsymbol{y}$ both as variables in the optimal estimation

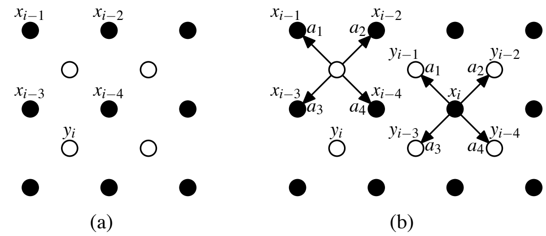


Fig. 1. (a) Spatial configuration in first pass. (b)PAR model parameters $\boldsymbol{a} = (a_1, a_2, a_3, a_4)$ in relationship to spatial correlations of pixels.

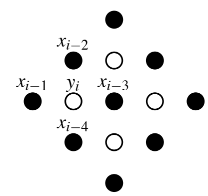


Fig. 2. Spatial configuration in second pass.

scheme of (2). This allows us to estimate $\boldsymbol{a}$ and $\boldsymbol{y}$ jointly under the constraint of the known LR image I_l. The optimization objective is to obtain the best statistical fit between the estimated model and the interpolated pixels. For this reason (2) requires the PAR model of the same parameters $\boldsymbol{a} \in \Re^4$ to not only fit $y_i \in I_h$ with $\{x_{i-t} \in I_l\}_{1 \le t \le 4}$, but also fit $x_i \in I_l$ with $\{y_{i-t} \in I_h\}_{1 \le t \le 4}$. If the image signal is autoregressive and stationary in the small local window W, which is fortunately true or a good enough approximation for most natural images, then the proposed approach is well grounded.

Once the missing HR pixels in the first pass are interpolated, we obtain half of the HR pixels. The remaining missing HR pixels are to be interpolated in the second pass. The interpolation problem in the second pass is essentially the same as the one just discussed. The only difference is that we interpolate the missing HR pixels $y_i \in I_h$ using their four 4-connected neighbors, which are either known in I_l or estimated in the first pass. The problem has the same formulation of non-linear optimization as in (2), if we simply rotate the spatial configuration of Fig. 1(a) by 45 degrees (see Fig. 2).

Autoregressive data fitting is a common technique in many image interpolation techniques [1, 2]. The distinctive advantage of the proposed non-linear optimization approach over the existing linear regression methods is that the former builds the adaptive autoregressive model using LR and estimated HR data in conjunction, whereas the latter methods train the model with LR data only.

However, with the new approach, the problem solution no longer has a closed form as in least-squares estimation. An iterative method, such as gradient descent, is needed to solve (2). And the solution may not be globally optimal since the objective function is not convex. Extra cares should be taken to expedite the convergence of the iterative method and to ensure a good solution. A high quality initial solution $(\boldsymbol{a}_0, \boldsymbol{y}_0)$ can make the gradient descent method converge quickly to a good solution. One possibility is to use the fast bicubic interpolation to compute $\boldsymbol{y}_0$ and then compute $\boldsymbol{a}_0$ from $\boldsymbol{y}_0$ using linear least-squares estimation.

In order to avoid poor locally optimal solutions in solving (2), one should add constraint(s) or a regularization term to (2) whenever possible. This is to incorporate into (2) useful domain knowledge about the HR pixel to be interpolated. For instance, in the presence of edges, which can be detected from the LR image, we can impose a smoothness constraint in the edge direction.

4. LINEAR LEAST-SQUARE SOFT-DECISION SOLUTIONS

The computational cost of the above non-linear optimal image interpolation approach is high. In quest for a more practical solution, we develop approximate algorithms in this section. The difficulty in solving (2) lies in its non-linearity. But the objective function of (2) will reduce to linear in its variables, if the PAR model parameters $\boldsymbol{a}$ can be estimated ahead of time. Guided by this observation, we work out an efficient solution by breaking the optimization problem of (2) into two linear least-square subproblems.

First, we determine the model parameters $\boldsymbol{a}$ by the following linear least-squares estimation:

$$\hat{\boldsymbol{a}} = \arg\min_{\boldsymbol{a}} \left\{ \sum_{i \in W} \left\| x_i - \sum_{1 \le t \le 4} a_t x_{i-t} \right\|_2 \right\} \tag{3}$$

where x_{i-t} are the four 8-connected neighbors of the location i in I_l as labeled in Fig. 3.

By plugging estimates $\hat{\boldsymbol{a}}$ into (2), we finally estimate the missing HR pixels by solving another linear least-square problem

$$\begin{aligned} \hat{\boldsymbol{y}} = \arg\min_{\boldsymbol{y}} \Bigg\{ & \sum_{i \in W} \left\| y_i - \sum_{1 \le t \le 4} \hat{a}_t x_{i-t} \right\|_2 \\ & + \sum_{i \in W} \left\| x_i - \sum_{1 \le t \le 4} \hat{a}_t y_{i-t} \right\|_2 \Bigg\} \end{aligned} \tag{4}$$

In the above derivations to simplify the algorithm we assumed that the second order statistics of LR image I_l and HR image I_h are sufficiently similar in the local window W. The same assumption was made by the interpolation method of Li and Orchard [1]. They justified this assumption for the case where an edge exists in the local window W. But the approach of [1] is fundamentally different from ours. In [1] once the parameters of the interpolation filter are determined similarly to (3), each missing pixel is interpolated independently

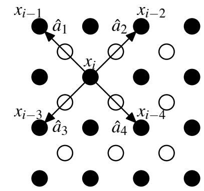


Fig. 3. Sample relations in estimating model parameters $\boldsymbol{a}$.

from others, which can be characterized as hard-decision estimation. In contrast, we adopt a soft-decision estimation approach. Rather than making one estimate at a time in isolation, the objective function (4) accounts for the mutual influences among estimates of neighboring missing pixels. These estimates are jointly optimized in the least squares sense in a local window W so that the PAR model fits all pixels in W, regardless from I_l or I_h.

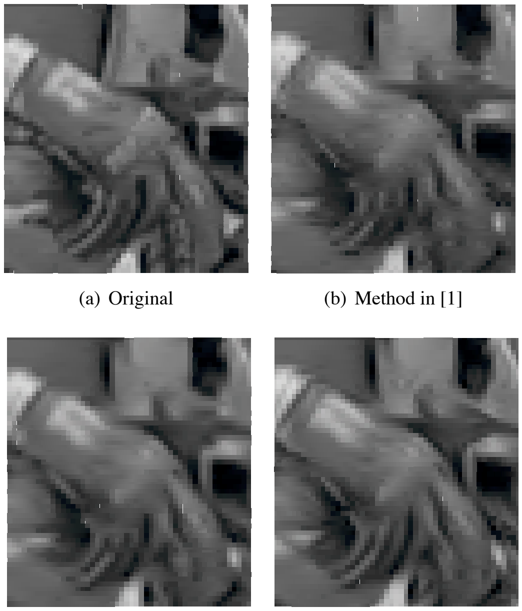
Experiments show that the edge-based interpolation method of [1] is prone to artifacts on spatial features of high curvature, for which the second order statistics may differ from LR to HR images. In such cases the soft-decision estimation strategy makes the proposed interpolation approach considerably more robust. Also, several variants are possible in the framework of soft-decision estimation to make the interpolation more adaptive.

5. EXPERIMENTAL RESULTS AND REMARKS

The two proposed image interpolation methods, the one of non-linear optimization and the one of linear least squares (LLS) soft-decision approach, were implemented and compared with two other methods: cubic convolution interpolation [6], and edge oriented interpolation [1]. Experiments were conducted on a number of test images often used in the literature, such as Lena (from the JPEG test set) and Bike (from the Kodak test set). As expected, the visual differences of the tested algorithm were exhibited in areas of high frequency contents such as edges and fine textures. Fig. 4 and Fig. 5 are parts of the Lena and Bike images, both original and the interpolated versions by the tested methods. It appears that the edges reconstructed by the two new methods are sharper with much less ringing and aliasing than other methods, contributing to superior visual quality. The difference between the non-linear optimization and the least-square soft-decision methods is small in both visual quality and PSNR. Given its much lower complexity, the latter method should be preferred in practice.

The non-linear optimization method has a very interesting property that can be beneficial for some applications. It tends to enhance or exaggerate fine details. One can see this by observing the bottom contour of the eye and the eye lashes in Lena images reconstructed by different methods.

The PSNR results for the test images are presented in Ta-



(a) Original (b) Method in [1]
(c) Proposed LLS method (d) Proposed nonlinear method

Fig. 4. Parts of original and reconstructed Bike images.

ble 1, which shows clear advantage of the new methods.

Image	Bicubic	Method [7]	Method [1]	Proposed LLS
Lena	33.85	32.99	33.79	34.58
Parrot	32.94	32.45	32.63	33.31
Bike	25.40	25.26	25.83	27.00
Flight	29.11	30.50	30.83	31.56

Table 1. The PSNR(dB) results

6. REFERENCES

[1] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Trans. Image Processing*, pp. 1521–1527, Oct. 2001.

[2] D. D. Muresan and T. W. Parks, "Adaptively quadratic (aqua) image interpolation," *IEEE Trans. on Image Processing*, vol. 13, pp. 690–698, May 2004.

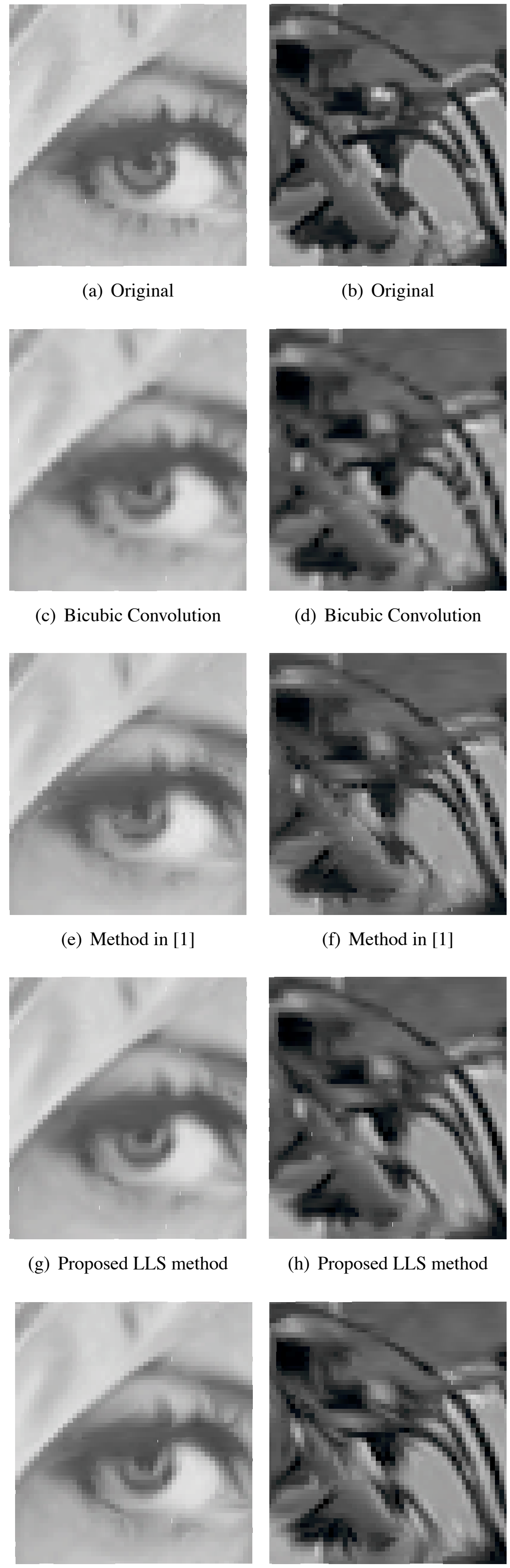
[3] X. Wu, K. U. Barthel, and W. Zhang, "Piecewise 2d autoregression for predictive image coding." in *IEEE ICIP'98*, vol. 3, Oct. 1998, pp. 901–904.

[4] B. Meyer and P. Tischer, "Tmw – a new method for lossless image compression," in *Proceed. of PCS'97*, Oct. 1997.

[5] I. Matsuda, H. Mori, J. Maeda, and S. Itoh, "Design and evaluation of minimum-rate predictors for lossless image coding," *Trans. IEICE (DII)*, vol. J85-D-II, pp. 448–456, Mar. 2002.

[6] R. G. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans. Acoustic, Speech and Signal Processing*, vol. 29, pp. 1153–1160, 1981.

[7] 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.



(a) Original (b) Original
(c) Bicubic Convolution (d) Bicubic Convolution
(e) Method in [1] (f) Method in [1]
(g) Proposed LLS method (h) Proposed LLS method
(i) Proposed nonlinear method (j) Proposed nonlinear method

Fig. 5. Parts of original and reconstructed Lena and Bike images.