IMAGE INPAINTING BASED ON GEOMETRICAL MODELING OF COMPLEX WAVELET COEFFICIENTS

Gang Hua, Michael T. Orchard

Department of Electrical and Computer Engineering
Rice University
Houston, TX 77005 USA

ABSTRACT

The restoration of missing regions in images (inpainting) is mathematically an interpolation problem and has many important applications. This paper proposes a novel iterative inpainting algorithm based on the interpolation of the Complex Wavelet Transform (CWT) coefficients with simple geometrical models on the magnitude and phase of the coefficients. The geometrical models describe the directionality and uniformity of the CWT magnitudes and the linearity of the CWT phases around edges and within texture areas. Both piecewise smooth signals and structured textures can be interpolated accurately with the proposed models. Motivated by the iterative reconstruction of an image from its CWT magnitude or phase, we propose an inpainting algorithm with iterative magnitude and phase estimation and CWT reconstruction. Simulation results show that the proposed algorithm achieves high PSNR and appealing visual quality with low computation complexity.

Index Terms— Wavelet transforms, Image restoration

1. INTRODUCTION

The restoration of missing regions in images has many important applications, such as the removal of scratches in old paintings [1], the predictive coding of images and videos, and the recovery of damaged image/video blocks due to errors in transmission or storage [2]. This problem has been known as *inpainting* among museum restoration artists. Inpainting algorithms estimate or interpolate a missing region of an image from information provided by surrounding regions based on some assumed model for images. This paper proposes a novel inpainting algorithm that interpolates piecewise smooth regions, edges, and patterned textures in images based on models placed on the magnitudes and phases of a Complex Wavelet Transform (CWT) representation of the unknown image.

It is clear that two types of image information need to be interpolated by any reasonable image model for inpainting. Within smooth (or piecewise smooth) regions, gray levels of the missing region should be smoothly interpolated based on surrounding gray levels. Many linear methods (polynomial interpolation, bandlimited interpolation, etc.) perform this processing well. But when surrounding pixel values indicate that some spatial structure (an edge, edges, or a patterned texture) passes through the missing region, a second type of interpolation is needed. In such cases, it is perhaps more clear to view the structure itself as being interpolated, rather than the pixel values. For example, inpainting a region containing a sharp edge involves first smoothly interpolating the contour defined by the edge, and then smoothly interpolating the pixel values on either side of the edge. Similarly, inpainting a region surrounded by patterned texture involves replicating the surrounding structure smoothly through the missing region. Because this second type of interpolation involves estimating the locations of structure features, nonlinear processing approaches are needed.

Existing approaches to inpainting implicitly define the missing region as the solution to a nonlinear optimization problem. For example, [3, 4] use variational methods, computing the missing region as the solution to a nonlinear PDE used to propagate information from the surrounding areas. These variational approaches work well on piece-wise smooth images but poorly on textures. The approaches of [5, 6, 7] define the missing region as the solution to an optimization problem seeking to maximize the sparsity of the image's linear expansion with respect to a specified dictionary of images. These approaches are very complex, and their performance depends heavily on the choice and design of various dictionaries.

This paper proposes a much more direct and unified approach to interpolating both gray levels and spatial structures, using the CWT representation of the image. The CWT represents an image with a redundant collection of coefficients generated by a bank of bandpass, analytic filters. The filters are selected so that, with redundancy of two in each direction, each band offers an unaliased representation of signal energy in a particular frequency range. The central idea of our approach comes from the observation: the missing region of an image is correctly interpolated from surrounding regions if the missing (or partly missing) CWT coefficients corresponding to that region are correctly interpolated from surrounding regions of coefficients. To interpolate the CWT coefficients in each band, we separately interpolate their magnitudes and their phases. I.e. using the CWT, we translate the inpainting problem into many smaller impainting problems of each band's magnitude and phase fields. CWT magnitudes represent the local band energy, and are typically very smooth in the highest energy bands associated with edges or patterned textures. Thus, although any approach can be used to interpolate the magnitude fields, we find that very simple directional smoothing of these fields gives very good results. CWT phases are only significant for coefficients with large magnitudes, and, for such coefficients, the phases represent the location of the band's energy. Using linear-phase CWT filters, we find the unwrapped phase fields associated with edges and patterned texture are approximated well by linear interpolation models. It should be noted that, since our very simple linear interpolation models are applied to parameters (magnitudes and phases) that are nonlinearly related to the image pixel values, they do not correspond to linear modeling assumptions on the image itself.

The paper is organized as follows. In section 2, we discuss the CWT and its important properties related to inpainting. In section 3, we propose a new iterative inpainting algorithm based on simple geometrical models of CWT magnitudes and phases. Section 4 gives some simulation results and section 5 concludes the paper and discusses possible future work.

2. THE COMPLEX WAVELET TRANSFORM

The CWT (see [8] and the reference therein) is a multi-resolution representation of images. In its magnitude and phase form, the CWT decomposes an image $\boldsymbol{f}(x, y)$ into a set of magnitudes $\boldsymbol{\rho}(x, y; k)$ and phases $\boldsymbol{\theta}(x, y; k)$, where k is in an index set Φ of scales and orientations. The CWT magnitude represents a smoothed measurement of the local signal energy for the designated frequency band, and the CWT phase indicates the location of that energy relative to the position of the each coefficient.

$$\left\{\boldsymbol{\rho}(x, y; k) e^{j\,\boldsymbol{\theta}(x,y;k)} : k \in \Phi\right\} = \mathrm{CWT}(\boldsymbol{f})$$

The frequency responses of the CWT filterbank on the first three levels are shown in Fig. 1 [1]. With carefully designed filters and redundancy, the CWT is nearly alias free, its magnitude response is a smooth measure of local energy, and its phase response is exactly or approximately linear. In higher dimensions, the CWT is approximately shift and rotational invariant.

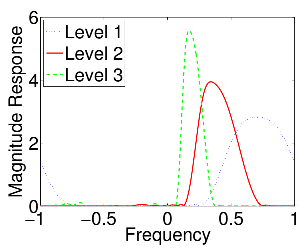
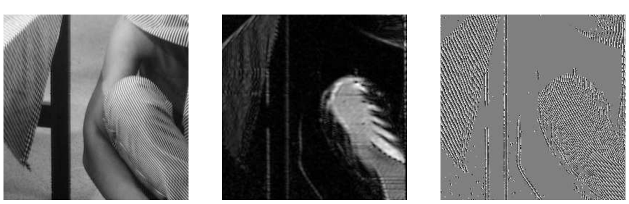


Fig. 1. The frequency response of the CWT

Two properties of the CWT are critical to our development of the inpainting algorithm of this paper. First, the CWT magnitude and phase exhibit strong geometrical regularity around edges and within texture areas. As illustrated in Fig. 2, the CWT magnitude is a smooth ridge-like function around edges (e.g., the arm and the chair leg). If an edge is symmetric, the CWT phase is approximately linear. Structured textures (e.g., the pants and the table cloth) can usually be decomposed into localized directional narrow band 2D components by the CWT and each component has uniform or patterned magnitude and approximately linear phase.



(a) Barbara (b) CWT magnitude (c) CWT phase

Fig. 2. The geometrical regularity of the CWT magnitude and phase

The second property of the CWT critical to our algorithm development is the fact that images can be reconstructed from CWT magnitude or phase. The CWT can be considered as a localized Fourier transform. It is well known that an image can be reconstructed from its localized Fourier magnitude or phase iteratively with POCS-like (projection onto convex set) iterations [10, 11, 12]. Similarly, an image can be reconstructed from its CWT magnitude or phase under certain conditions (e.g., Fig. 3). An algorithm to reconstruct from CWT magnitude is given below and the algorithm for reconstruction from CWT phases is similar.

[1] All the figures and results in this paper are generated with an implementation of the CWT following [9].

```
Iterative Reconstruction from the magnitude
Given initial f̂ and the magnitude ρ.
(1) Compute the CWT: (ρ̃, θ̃) = CWT(f̂).
(2) Let (ρ̂, θ̂) = (ρ, θ̃).
(3) Compute the inverse CWT: f̂ = ICWT(ρ̂, θ̂).
(4) Goto (1).
```

The reconstruction process is shown in Fig. 3. The convergence is very slow because some regions in images are not suitable for either magnitude only or phase only representation. For example, the magnitude contains little information about the texture structure and the phase has no significance in smooth regions. The convergence becomes very fast if the suitable representation is chosen (e.g., use magnitude for smooth areas and phase for textures).

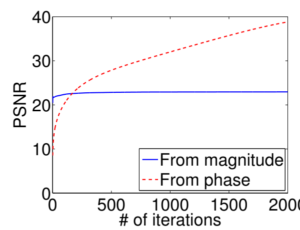


Fig. 3. Iterative reconstruction of Fig. 2 (a) from only its CWT magnitude and phase (the reconstruction from magnitude cannot converge to the right image because the magnitude has little information about the texture structures).

3. THE PROPOSED INPAINTING ALGORITHM

Suppose in an image $\boldsymbol{f}$, the region $\boldsymbol{f_a}$ is known and the region $\boldsymbol{f_b}$ is missing. The missing $\boldsymbol{f_b}$ has to be estimated from the information available in $\boldsymbol{f_a}$ with some assumed image model.

$$\boldsymbol{f} = \begin{bmatrix} \boldsymbol{f_a} \\ \boldsymbol{f_b} \end{bmatrix}$$

In the CWT domain, there are roughly corresponding missing regions of the magnitude ($\boldsymbol{\rho}$) and phase ($\boldsymbol{\theta}$) in each band. Therefore, the original inpainting problem of estimating $\boldsymbol{f_b}$ is translated into the problem of estimating $(\boldsymbol{\rho_b}, \boldsymbol{\theta_b})$ in each band.

$$\boldsymbol{\rho} = \begin{bmatrix} \boldsymbol{\rho_a} \\ \boldsymbol{\rho_b} \end{bmatrix}, \qquad \boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta_a} \\ \boldsymbol{\theta_b} \end{bmatrix}$$

As discussed above, $\boldsymbol{\rho}$ and $\boldsymbol{\theta}$ have strong geometrical regularity around edges and within textures. In section 3.1, we propose simple models to describe the geometrical regularity and estimate the missing magnitude and phase:

$$\widehat{\boldsymbol{\rho}} = \begin{bmatrix} \boldsymbol{\rho_a} \\ \widehat{\boldsymbol{\rho}}_{\boldsymbol{b}} \end{bmatrix}, \qquad \widehat{\boldsymbol{\theta}} = \begin{bmatrix} \boldsymbol{\theta_a} \\ \widehat{\boldsymbol{\theta}}_{\boldsymbol{b}} \end{bmatrix}$$

Here, we assume that the models hold around and within the missing region and the image with a set of magnitude and phase satisfying the models is a reasonable estimate of the original image.

If the estimation is perfect ($\widehat{\boldsymbol{\rho}} = \boldsymbol{\rho}$ and $\widehat{\boldsymbol{\theta}} = \boldsymbol{\theta}$), the image $\boldsymbol{f}$ can be recovered trivially with the inverse CWT ($\boldsymbol{f} = \mathrm{ICWT}(\boldsymbol{\rho}, \boldsymbol{\theta})$). In reality, there are errors in the estimation. Then, due to the redundancy of the CWT, the image with exactly the estimated magnitude ($\widehat{\boldsymbol{\rho}}$) and phase ($\widehat{\boldsymbol{\theta}}$) may not exist. The image given by the inverse CWT ($\widetilde{\boldsymbol{f}} = \mathrm{ICWT}(\widehat{\boldsymbol{\rho}}, \widehat{\boldsymbol{\theta}})$) may be unsatisfactory, because its magnitude and phase fields ($(\widetilde{\boldsymbol{\rho}}, \widetilde{\boldsymbol{\phi}}) = \mathrm{CWT}(\widetilde{\boldsymbol{f}})$) do not necessarily satisfy the modeling constraints that were applied to estimate $\widehat{\boldsymbol{\rho}}$ and $\widehat{\boldsymbol{\theta}}$. Inspired by image reconstruction from its CWT magnitude and phase, in section 3.2, we propose an iterative estimation algorithm to enforce the geometrical models on the magnitude and phase and converge to an image satisfying the proposed models.

3.1. Geometrical Models for CWT Magnitude and Phase

We propose a directional model for the CWT magnitude in the missing region. Suppose a block of m by m magnitudes is missing in one CWT band and denote the i-th column of the missing magnitudes as $\boldsymbol{\rho_i}$. The columns $\boldsymbol{\rho_i}$ are modeled as different shifts of a common function $\boldsymbol{p}$:

$$\boldsymbol{\rho_i} = D(\boldsymbol{p}, \tau_i)$$

where $D(\boldsymbol{p}, \tau)$ is the shift of $\boldsymbol{p}$ with amount τ. We assume that the variable τ_i changes smoothly with column number i (it can be linear, quadratic or more complex functions of i). We found that a linear model of τ_i is adequate for inpainting the missing blocks with size 16 by 16 in real images [2], because within such a small region image structures are close to be straight. Therefore, in this paper, we simply choose τ_i to be linear in i. Correspondingly, the phases in the missing region is simply modeled as a 2D linear function.

The magnitude and phase estimation with the proposed models is explained by an example shown in Fig. 4. For clarity and space reasons, only the interpolation strategy is explained and the implementation details are neglected. To interpolate the magnitude in Fig. 4 (a) from (b), we take the columns of the magnitudes just to the left ($\boldsymbol{\rho}_l$) and right ($\boldsymbol{\rho}_r$) of the missing region and determine their relative shift ($\widehat{\boldsymbol{\tau}}$) by maximizing the cross correlation of $D(\boldsymbol{\rho}_l, \tau)$ and $\boldsymbol{\rho}_r$. We assume that the columns of magnitudes in the missing region ($\boldsymbol{\rho}_i, l < i < r$) are shifts of $\boldsymbol{\rho}_l$ and $\boldsymbol{\rho}_r$ and the shift τ_i changes linearly with i. Therefore, we can use the shift operator D to estimate all the columns of magnitudes in the missing region as below:

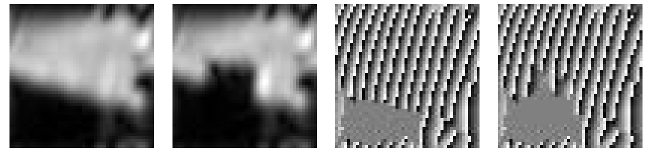
$$\widehat{\tau}_i = \frac{i-l}{r-l}\widehat{\tau}$$
$$\widehat{\boldsymbol{\rho}}_i = \frac{r-i}{r-l}\mathrm{D}(\boldsymbol{\rho}_l, \widehat{\tau}_i) + \frac{i-l}{r-l}\mathrm{D}(\boldsymbol{\rho}_r, -\widehat{\tau} + \widehat{\tau}_i)$$

For the estimation of the missing phases, we fit the unwrapped CWT phase to a linear 2D plane with the current estimate of the magnitude as weights. If the linear model fits the phases in the surrounding areas very well, we use it to predict the phases in the missing region. Otherwise, we keep the current estimate of the phase and rely on the magnitude estimate. For the case of Fig. 4, the phase in the missing region can be recovered accurately.

3.2. The Proposed Iterative Inpainting Algorithm

To address the inpainting problem, we propose an iterative algorithm to construct an image estimate $\widehat{\boldsymbol{f}}$ with magnitude and phase complying with the models proposed above. It is interesting to note that

[2]The block size for image and video coding is usually no larger than 16 by 16.



(a) (b) (c) (d)

Fig. 4. The CWT magnitude and phase of the true image ((a) and (c)) and of the image with missing block ((b) and (d)) (the phases associated with very small magnitudes are set to 0).

image reconstruction from the CWT magnitude or phase is exactly the case when only $\widehat{\boldsymbol{\rho}}$ or $\widehat{\boldsymbol{\theta}}$ is accurate. Therefore, when either $\boldsymbol{\rho}$ or $\boldsymbol{\theta}$ can be estimated accurately, we could use that iterative algorithm to reconstruct the missing block. When both $\widehat{\boldsymbol{\rho}}$ and $\widehat{\boldsymbol{\theta}}$ are very noisy, by replacing step (2) of the reconstruction algorithm in section 2 with the magnitude and phase estimation proposed in section 3.1, we have an iterative inpainting algorithm. Even if none of $\widehat{\boldsymbol{\rho}}$ and $\widehat{\boldsymbol{\theta}}$ is accurate enough to reconstruct a good estimate at the beginning, by repeatedly enforcing the geometrical models of the magnitude and phase, we expect that the algorithm will keep refining them and converge to an image complying with the proposed models.

The algorithm is, however, not robust enough against the magnitude and phase estimation errors caused by the initial estimate $\widehat{\boldsymbol{f}}_{\boldsymbol{0}}$ (calculated by averaging the available neighboring pixels). When the initial estimate contains some strong spurious edges, the estimation may produce significant errors and the algorithm may converge to a poor estimate of the missing block. One way of removing the spurious edges in the initial estimate is to use an iterated denoising method similar to [6] with the CWT based denoiser in [9] to improve the quality of the initial estimate. This improved initial estimate works very well with the proposed iterative inpainting algorithm. A better way is to combine the iterated denosing and iterative estimation together. We choose a sequence of decreasing thresholds and hard threshold the interpolated magnitude with the i-th threshold in the sequence at the i-th iteration. The successive hard thresholding operation will remove the influence of the spurious edges in the initial estimate. The entire iterative algorithm is presented below:

The Proposed Iterative Inpainting Algorithm

Given parameter $T_0 > T_1$ and Δ

(1) Set $n = 1$ and $T = T_0$.

(2) Compute the initial image estimate $\widehat{\boldsymbol{f}}_{\boldsymbol{0}}$.

(3) Compute the CWT: $(\widetilde{\boldsymbol{\rho}}, \widetilde{\boldsymbol{\theta}}) = \mathrm{CWT}(\widehat{\boldsymbol{f}}_{\boldsymbol{n-1}})$

(4) Interpolate $(\widetilde{\boldsymbol{\rho}}, \widetilde{\boldsymbol{\theta}})$ to get $(\widehat{\boldsymbol{\rho}}, \widehat{\boldsymbol{\theta}})$

(5) Hard threshold $\widehat{\boldsymbol{\rho}}$ with T.

(6) Compute the inverse CWT: $\widetilde{\boldsymbol{f}} = \mathrm{ICWT}(\widehat{\boldsymbol{\rho}}, \widehat{\boldsymbol{\theta}})$.

(7) Compute new estimate $\widehat{\boldsymbol{f}}_{\boldsymbol{n}} = [\boldsymbol{f_a}^T, \widetilde{\boldsymbol{f}}_{\boldsymbol{b}}^T]^T$.

(8) Set $n = n+1$, $T = T - \Delta$ and goto (3) while $T > T_1$.

(9) Output the final inpainting result

4. SIMULATION RESULTS

The simulation results of the proposed algorithm are shown in Fig. 5 and the results of the iterated denoising method in [5, 6] with 16 by 16 DCT are also shown for comparison. To finish all the shown examples, it takes about 50 seconds with our algorithm, while the C code of the iterated denoising method takes about 6 minutes on the same computer. The image blocks are all from Lena and Barbara with 16 by 16 missing blocks. We set $\Delta = 2$, $T_1 = 6$, and T_0 to 4 times the standard deviation of the pixels surrounding the missing block. The proposed algorithm generates inpainting results with high PSNR and good visual quality. It may appear to the readers that the proposed linear models are too simple for real life images. However, they are very effective when applied on the CWT magnitude and phase and combined with the proposed iterative procedure.

5. CONCLUSION AND FUTURE WORK

In this paper, we proposed an new iterative inpainting algorithm with simple geometrical models of the CWT magnitude and phase. The proposed algorithm gives inpainting results with both high PSNR and appealing visual quality for piecewise smooth signals, textures and their mixtures. In the future, more sophisticated models of CWT magnitude and phase could be developed to deal with more complicated image structures.

6. REFERENCES

[1] M. Bertalmío, L. A. Vese, G. Sapiro, and S. Osher, "Simultaneous structure and texture image inpainting," *IEEE Trans. Image Proc.*, vol. 12, no. 8, pp. 882–889, Aug. 2003.

[2] W. Zeng and B. Liu, "Geometric-structure-based error concealment with novel applications in block-based low bit rate coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, pp. 648–665, June 1999.

[3] C. Ballester, M. Bertalmío, V. Caselles, G. Sapiro, and J. Verdera, "Filling-in by joint interpolation of vector fields and gray levels," *IEEE Trans. Image Proc.*, vol. 10, no. 8, pp. 1200–1211, Aug. 2001.

[4] T. F. Chan and J. Shen, "Mathematical models of local non-texture inpaintings," *SIAM J. Appl. Math.*, vol. 62, no. 3, pp. 1019–1043, 2001.

[5] Onur G. Guleryuz, "Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising-part i: theory," *IEEE Trans. Image Proc.*, vol. 15, no. 3, pp. 539–554, March 2006.

[6] Onur G. Guleryuz, "Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising-part ii: adaptive algorithms," *IEEE Trans. Image Proc.*, vol. 15, no. 3, pp. 555–571, March 2006.

[7] M. Elad, J.-L Starck, D. Donoho, and P. Querre, "Simultaneous cartoon and texture image inpainting using morphological component analysis (mca)," *Journal on Applied and Computational Harmonic Analysis ACHA*, vol. 19, pp. 340–358, 2005.

[8] I. W. Selesnick, R. G. Baraniuk, and N. Kingsbury, "The dual-tree complex wavelet transform - a coherent framework for multiscale signal and image processing," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 123–151, Nov. 2005.

[9] Gang Hua, "Noncoherent image denoising," M.S. thesis, Rice University, 2005.

[10] J. Behar, M. Porat, and Y.Y. Zeevi, "Image reconstruction from localized phase," *IEEE Trans. Signal Proc.*, vol. 40, no. 4, pp. 736–743, Apr. 1992.

[11] S. Urieli, M. Porat, and N. Cohen, "Image reconstruction from localized phase," *IEEE Trans. Image Proc.*, vol. 7, no. 6, pp. 838–853, June 1998.

[12] G. Michael and M. Porat, "Image reconstruction from localized fourier magnitude," in *Proc. Int. Conf. on Image Proc. (ICIP)*, 2001, vol. 1, pp. 213–216.

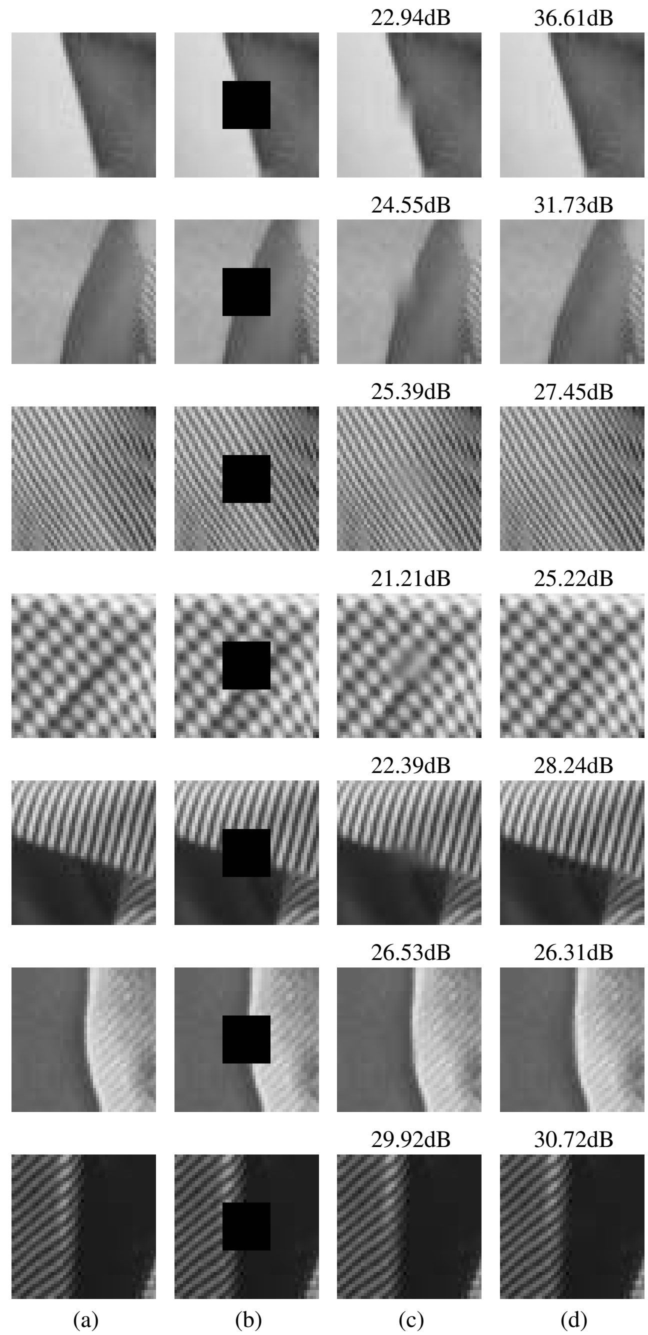


Fig. 5. Simulation results: (a) clean images, (b) missing blocks, (c) results of [5, 6], and (d) our results (the dB numbers in (c) and (d) are the PSNR of the missing blocks)