Lapped Nonlinear Interpolative Vector Quantization and Image Super-Resolution

David G. Sheppard, Kannan Panchapakesan, Ali Bilgin,
Bobby R. Hunt, Michael W. Marcellin

Department of Electrical and Computer Engineering
University of Arizona
Tucson, AZ 85721

Abstract

This paper presents an improved version of an algorithm designed to perform image restoration via *nonlinear interpolative vector quantization* (NLIVQ). The improvement results from using lapped blocks during the encoding process. The algorithm is trained on original and diffraction-limited image pairs. The discrete cosine transform is used in the codebook design process to control complexity. Simulation results are presented which demonstrate improvements over the non-lapped algorithm in both observed image quality and peak signal-to-noise ratio. In addition, the nonlinearity of the algorithm is shown to produce super-resolution in the restored images.

1 Introduction

Vector quantization (VQ) maps consecutive, usually non-overlapping, segments of input data to their best matching entry in a codebook of reproduction vectors [4]. VQ is generally considered a data compression technique. However, VQ algorithms have been presented which perform other signal processing tasks concurrently with compression. These span the range from speech processing tasks such as speaker recognition and noise suppression, to image processing tasks like half-toning, edge detection, enhancement, classification, reconstruction, and interpolation [2].

In earlier work [7, 8], a novel algorithm was presented for image super-resolution based on *nonlinear interpolative vector quantization* (NLIVQ) [3]. This algorithm addressed the classical problem of removing the blur caused by a diffraction-limited optical system [1]. Such a system acts as a low pass filter with an absolute spatial cutoff frequency proportional to its exit pupil diameter, and completely suppresses spatial frequency components of the original scene outside the system passband [5]. Image super-resolution encompasses correction of the filtering in the passband and some recovery of spatial frequency components outside the passband [6].

An improved version of the algorithm is presented in this work. As before, the algorithm is trained on original and diffraction-limited image pairs which are assumed to be representative of the class of images of interest. The DCT is used to process the image blocks in order to manage codebook complexity. Improvements result from lapping the blocks during encoding. This suppresses many of the artifacts present in images processed with earlier versions of the algorithm and produces super-resolved images which are qualitatively and quantitatively better.

The following sections present a brief review of the algorithm design process, the improved lapped algorithm, and simulation results which demonstrate the improvements in image super-resolution as compared with earlier non-lapped versions of the algorithm.

2 Nonlinear Interpolative VQ and Image Restoration

In this section, the basic theory behind the algorithm and its design are discussed. The task at hand is to design an operator which takes as its input a blurred image block and produces the unblurred original block. This is done by training the algorithm with a large number of blurred and unblurred images. Let $\{F^i, G^i\}_{i=1}^n$ be a sequence of image pairs, where F^i and G^i are the original and diffraction-limited $N \times N$ images, respectively. Decompose each image pair of the sequence into $M \times M$ blocks which will serve as the VQ training data. Let f^{ik} and g^{ik} be block k from F^i and G^i, respectively. Assume that the encoder $\mathbf{E}$, decoder $\mathbf{D}$, and the associated codebook $\mathbf{C}$, are given for a VQ that minimizes the distortion

$$D = E\left[d\left(g^{ik}, \tilde{g}^{ik}\right)\right]. \tag{1}$$

The process for choosing the quantized block $\tilde{g}^{ik}$ can be written as

$$\tilde{g}^{ik} = \mathbf{D}\left(\mathbf{E}\left(g^{ik}\right)\right) = \arg\min_{c_l \in \mathbf{C}} \; d\left(g^{ik}, c_l\right), \tag{2}$$

where c_l refers to entry l of $\mathbf{C}$.

Define the nonlinear VQ restoration algorithm as a new decoder $\mathbf{D}^*$, and its associated codebook $\mathbf{C}^*$, which minimizes the conditional expectation

$$D = E\left[\; d\left(f^{ik}, \tilde{f}^{ik}\right)^2 \;\middle|\; \mathbf{E}\left(g^{ik}\right) = l \;\right], \tag{3}$$

where $\mathbf{E}$ returns the index of the matching codebook entry. For a given set of training data, let $B_l = \left\{f^{ik} : \mathbf{E}\left(g^{ik}\right) = l\right\}$. Define entry l of $\mathbf{C}^*$ as the centroid of B_l, or

$$c_l^* = \left(\frac{1}{|B_l|}\right) \sum_{f^{ik} \in B_l} f^{ik}. \tag{4}$$

Finally, the nonlinear VQ restoration algorithm is given by

$$\tilde{f}^{ik} = \mathbf{D}^*\left(\mathbf{E}\left(g^{ik}\right)\right) = c^*_{\mathbf{E}(g^{ik})}, \tag{5}$$

where $\tilde{f}^{ik}$ is the restored image block.

It is important to note that the blurred imagery must be oversampled sufficiently to avoid aliasing if the algorithm achieves super-resolution.

3 Codebook Design and Lapped Encoding

The encoder codebook $\mathbf{C}$ is designed using a technique based on the discrete cosine transform (DCT). The DCT-based scheme, which is non-iterative, allows much larger codebooks than are practical with the Lloyd algorithm. The procedure for designing the DCT-based encoder is summarized for $M \times M$ blocks in the following steps:

1. Compute the DCT $\hat{g}^{ik}$ of each input block, g^{ik}.
2. For an encoding rate of R bits/pixel, allocate $L = RM^2$ bits among the transform coefficients to minimize the mean-squared error distortion of the quantized DCT blocks.
3. If l_{mn} is the number of bits allocated to the (m, n) DCT coefficient, design the (scalar) Lloyd-Max quantizer having $2^{l_{mn}}$ levels for that coefficient. The coefficient is assumed to be Laplacian distributed.
4. Define the fixed-length vector quantizer encoder $\mathbf{E}$ as the concatenation of the binary codes for the (scalar quantized) transform coefficients. This concatenation (or its decimal equivalent) is the codeword index.

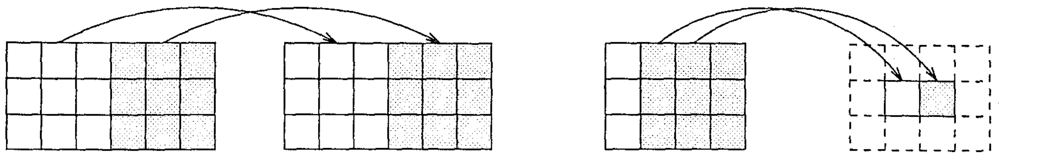
The next step is to compute the codebook $\mathbf{C}^*$ for the nonlinear VQ decoder. This follows directly from the encoder design in deterministic fashion and can be summarized in the following steps

1. For each input block g^{ik} derived from the set of N diffraction-limited images, as defined above, compute the index produced by the encoder $\mathbf{E}(g^{ik}) = q$.
2. Add the block f^{ik}, as defined above, to the running sum for codeword c_q^* and increment the counter s_q^* for that codeword.
3. After all blocks in the training set have been processed according to steps (1) and (2), compute each codeword in C^* as the average of each running sum according to $c_q^* = \frac{c_q^*}{s_q^*}$.

Restoration of one image block requires the calculation of the DCT, the scalar quantization of the DCT coefficients, and a table lookup. The computational complexity of these calculations grows linearly with the number of pixels in the image block (M^2) and is roughly independent of the encoding rate (R).

The lapped encoding used in the improved algorithm does not require a new codebook design procedure. The difference is that lapped blocks in the blurred image are mapped to a sub-block in the restored image. For example, 3×3 blocks in the blurred image may map to a single pixel (the center pixel of the restored block) in the output image. The blocks have a two column overlap in this case. For 4×4 blocks, the output may be a 2×2 or 1×1 sub-block from the output block produced by the decoder codebook. This is depicted graphically in Figure 1. Only the 3×3 block size was used in this work. The improved results indicate that the larger errors in the output blocks are near the block edges, the source of the blocking artifacts seen in the non-lapped algorithm output.

The simulation results described below are obtained by applying the algorithm to mean-removed image blocks. Estimation of the mean of the restored block is dealt with as a separate problem. This allows all of the bits available to be used in representing the AC information of the block, resulting in better performance. Restoration of the block mean is done with a Wiener filter process.



(a) Non-lapped encoder

(b) Lapped encoder

Figure 1: Lapped encoder contrasted with the non-lapped version for 3×3 blocks

The parameters for the results below are: 1) 3×3 blocks; 2) 12 bits/mean-removed block; 3) a training set of 70 (512×512) image pairs of aerial views of urban areas; 4) optical cutoff frequency equal to half the folding frequency; and 5) no noise in the blurred images. Figure 2 displays crops of an "original" test image (outside the training set), the blurred image produced from the original, and non-lapped and lapped restorations. This image is similar in edge content to many of the images in the training set. Note that near the edge of the lapped restoration the pixels for which there is insufficient support for the mask have been set to zero.

In general, peak signal to noise ratio (PSNR) values of images processed by the algorithm improved by 1.5 to 2.5 dB in the non-lapped case. The lapped algorithm produces improvements in the 2.5 to 4.5 dB range. This quantitative improvement in the images is matched by a significant improvement in visual quality. This is true for images both in and out of the training set.

Super-resolution is usually defined in terms of the recovery of spatial frequency components and the improved performance in this regard is shown in Figure 3, where the log_{10} of the Fourier transform magnitudes of the images from Figure 2 are displayed. It is evident that the stronger features in the original spectrum have reappeared in the non-lapped and lapped restoration spectra. The effect is more pronounced in the lapped case.

4 Conclusion

An improved algorithm for image super-resolution based on nonlinear interpolative vector quantization was presented. The NLIVQ training process determines the important statistical properties of the data and accomplishes the design of a nonlinear restoration algorithm. A DCT encoder was employed to manage the codebook complexity and avoid iterative training. The improvements resulting from using lapped blocks in the decoder can be seen in the suppression of artifacts present in earlier results. Both quantitative and qualitative improvements were obtained in addition to a increased super-resolution of the processed images.

References

[1] H. C. Andrews and B. R. Hunt, *Digital Image Restoration*, Prentice-Hall, 1977.

[2] P. C. Cosman, K. L. Oehler, E. A. Riskin and R. M. Gray, "Using vector quantization for image processing," *Proc. IEEE*, Vol. 81, pp. 1325–41, 1993.

[3] A. Gersho, "Optimal nonlinear interpolative vector quantization," *IEEE Trans. Comm.*, Vol. 38, pp. 1285–87, 1990.

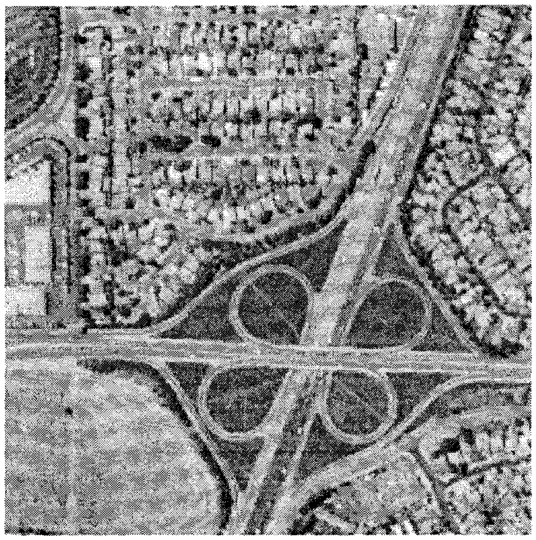
[4] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer, 1992.

[5] J. W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, 1968.

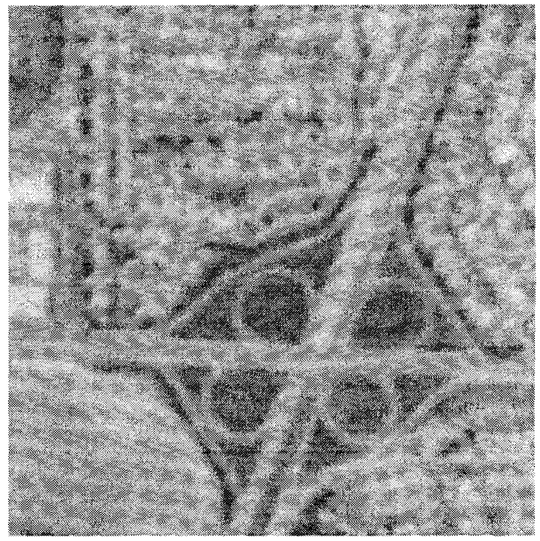
[6] B. R. Hunt, "Super-resolution of images: algorithms, principles, performance," Int. J. Imaging Sys. and Tech., Vol. 6, pp. 297–304, 1995.

[7] D. G. Sheppard, A. Bilgin, M. S. Nadar, B. R. Hunt, M. W. Marcellin, "A vector quantizer for image restoration," *IEEE International Conference on Image Processing*, Lausanne, Switzerland, Sep. 16-19, 1996, pp. 439-41.

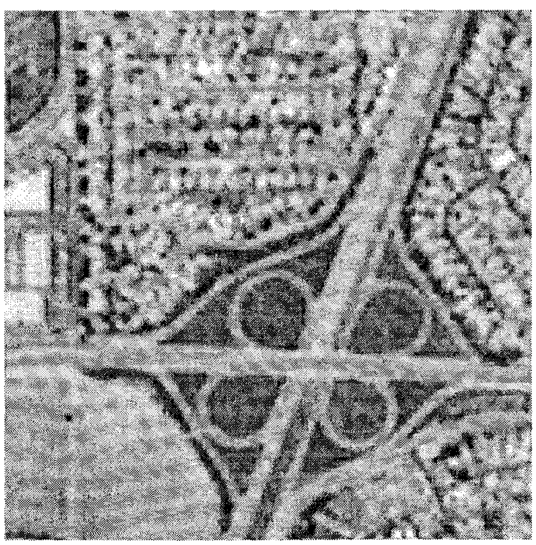
[8] D. G. Sheppard, A. Bilgin, M. S. Nadar, B. R. Hunt, M. W. Marcellin, "A vector quantizer for image restoration," *IEEE Transactions on Image Processing*, to appear.



(a) Original image



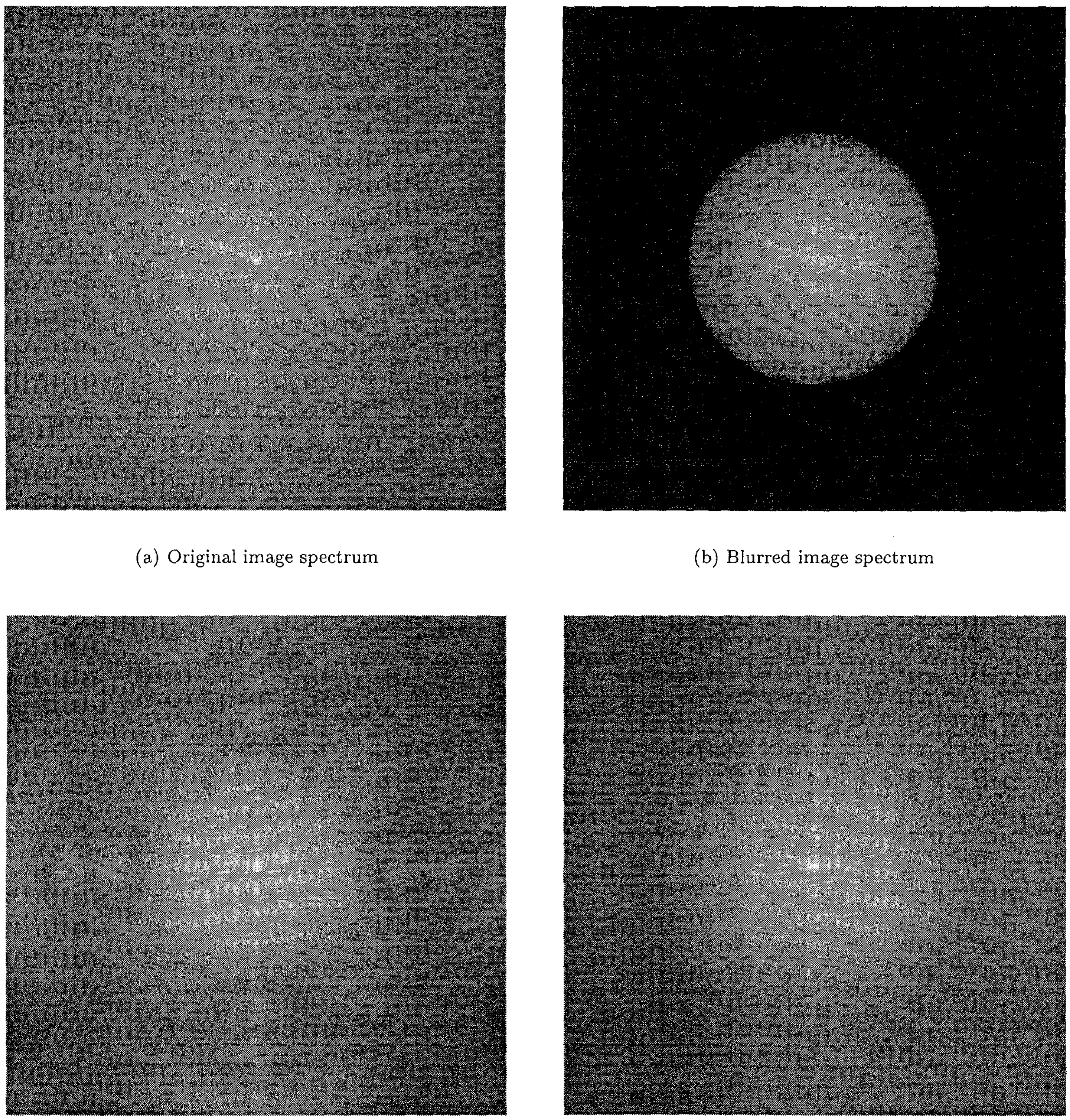
(b) Blurred image (PSNR = 21.36 dB)



(c) Non-lapped restoration (PSNR = 23.85 dB)

(d) Lapped restoration (PSNR = 25.77 dB)

Figure 2: Crops from the images used to test the algorithms



(a) Original image spectrum

(b) Blurred image spectrum

(c) Non-lapped restoration spectrum

(d) Lapped restoration spectrum

Figure 3: Spectra of the images used to test the algorithms.