A Hybrid Wavelet Compression Scheme for Material Textures [♦]

D. Ramasubramanian, Laveen N. Kanal
LNK Corporation Inc., 6811 Kenilworth Avenue #306, Riverdale MD 20737
{daram, kanal}@lnk.com

Russ Moulton, William Journigan
JRM Technologies, Inc., PB-3866, Fredericksburg VA 22407
russ@jrmtech.com

Abstract: In this paper, we present a novel compression scheme based on the wavelets for compressing material textures and terrain data. Our new hybrid technique is efficient, attractive and appropriate for material textures as demonstrated by the results. Further, the proposed scheme, which is amenable to real time implementation, is found to be superior to the existing state-of-art hardware-based compression schemes such as DIRECTX-S3TC and FXT1.

1. Introduction

Compression techniques have appeared in the literature for more than three decades. Essentially, most of the techniques exploit correlation that exists among the pixels, in both the spatial and spectral domains. Methods based on the statistics of the images, such as Huffman coding and arithmetic coding, achieve lossless compression. However, the order of compression is very small. On the other hand, methods based on the transformation of images to the spectral domain result in lossy compression but with larger compression ratios [1].

Material-based textures pose a unique problem for compression. Unlike other texture images, index value at one particular location refers to the physical property associated with the material. These materials give rise to specific signatures. Obviously, compression errors can wreck havoc on sensor signature renderings. Errors in height field elevation or color-based texture might be tolerable – *but small errors in material index mean a completely different material is selected for signature synthesis.*

One of the well-known methods to compress images is the international standard proposed by the Joint Photographic Expert Group (JPEG). However, one major drawback of this method is the unacceptable amount of artifacts and distortion in the decompressed images for a very high order of compression. It is also possible that some of the important features or targets in the images may be lost or significantly affected in the course of compression. In particular, JPEG techniques are found to be unsuitable for compressing the material-based textures. Furthermore, in applications involving the analysis of images at different resolution levels, the JPEG-based compression method achieves poor performance.

Over the past several years, wavelets and other similar multi-resolution techniques have been successfully exploited to compress images. Of particular interest, both in terms of

♦ This research work was supported by navy under the SBIR Phase I contract N6835-00-C-0299.

0-7695-1110-4/01 $10.00 © 2001 IEEE

efficiency and accuracy, are the wavelet-based schemes and various hybrid forms. Methods for effectively compacting the data stream by advanced entropy coding have emerged in the recent studies, as also have integerizable and reversible transformations. The multi-resolution property of the wavelet transform is highly relevant for both Level-of-Detail (LOD) terrain geometry and associated mip-mapping of textures as well as the progressive transmission of images.

The problem addressed in this paper is to develop a powerful wavelet-based scheme in order to compress the terrain data and texture images. Some of the challenges include that it should handle material based textures obtained from different types of sensors as well as color based. In addition, the proposed solution should be computationally efficient and attractive in order to develop hardware engines for the real-time applications.

2 Eigen Wavelets for Compression

The computation of the discrete wavelet transform can be thought of as a two channel decomposition of a signal and iterating the same procedure on the lower band. This procedure is commonly referred to as "sub-band coding". In general, the decomposition could consist of N channels using bandpass filters covering the entire frequency range and each band down sampled by a factor of N. If the set of functions used constitute an orthonormal basis then the signal can be reconstructed by combining the up sampled and inverse filtered signals. In general, the significant coefficients that represent the important features of an image tend to be small relative to the original values. It is this characteristic that allows compression of wavelet encoded data with little loss of information. By eliminating these small coefficients, a reduced representation of the original image can be obtained that keeps much of the information of the original image, and upon reconstruction, these deleted coefficients introduce only small errors.

In addition to this elimination of high-frequency detail, coefficients can also be quantized. This implies statistically examining the set of coefficients in a wavelet representation of a signal, quantizing these coefficients into some number of bins, and storing a representation of which bin a particular coefficient came from. Upon reconstruction of the original signal, the centers of these bins are used as coefficients, rather than the original coefficients.

Let us now assume that data matrix $A \in \Re^{N \times M}$ is formed by horizontally stacking successive spectral responses represented by column vectors. The number of pixels in the scanline is M and N is the number of spectral channels. We assume that A is full rank. This assumption is usually true, and most certainly holds when there is additive noise in the data. We also assume that $M \geq N$. Since the matrix is full rank, A can be represented as the product

$$A = B \cdot C,$$

for an appropriate basis set $B \in \Re^{N \times N}$ and coefficient matrix $C \in \Re^{N \times M}$. These matrices can be evaluated using the Singular Value Decomposition (SVD). We note the following:

1. In the case of lossless data compression, the entries of data matrix A are integers.

2. We must be able to represent B and C with finite precision in order to reconstruct A without loss. In order to achieve compression, the rows of C must show a significant decay in magnitude.
3. B and C must be amenable to efficient encoding.

In theory, we can construct finite precision representations of B and C that satisfy condition 2 and condition 3 holds when the basis vectors in B are appropriately arranged. However, we can expect that the finite precision representations will themselves contain too much information, thereby violating condition 4. Besides, SVD computation is not inexpensive. This prompts us to consider an alternative formulation based on prediction [2].

Assume now that only the top R basis vectors and their corresponding coefficients are used in representing A. Clearly, this reduced rank representation is imperfect. Assume also that the difference between A and its rank R representation, referred to as the residual matrix, is transmitted to the receiver. The basis and coefficient matrices are $B \in \Re^{N \times R}$ and $C \in \Re^{R \times M}$ respectively, and residual matrix is $\Delta \in \Re^{N \times M}$. Now, the data matrix A is represented as $A = B \cdot C + \Delta$.

The product $B \cdot C$ is a linear prediction of A, where the linear basis is dependent on the data. Residual is the corresponding prediction error. Limited rank eigen analysis of a data matrix produces three outputs: the basis vector set, the set of coefficients and a data matrix of residuals. For a reversible system, we place a constraint of finite precision on the various matrices. Data matrix A is integral and so is the residual array. Constraining B and C to integer matrices is not effective for data compression. In fact, the basis vectors (which are rows of B) are orthonormal. Therefore, we scale B prior to quantization, i.e. we use a fixed point representation of B. Typically, 8 bits of precision are sufficient for B.

The coefficient matrix C is more difficult to quantize. Each coefficient entry is formulated as the dot product of two vectors. These vectors are the stripped data matrix row and the appropriate basis function, where stripped data matrix refers to the residual matrix after extracting the higher-rank projections. The product of these terms has the same precision as B, for an integer data array. However, coefficients are usually much larger in range than entries of B and hence the quantization rule used for B is inapplicable to C. We use an empirical rule for quantizing C, based on its average range and an index.

2.1 Residual Processing

The prediction residual matrix is, by construction, integer valued. Since eigen analysis provides a good approximation of A, the residual matrix has small entries. Much like a grayscale image, it displays strong local continuity and small long range dependence. Inspired by the success of wavelet transform based image compression, we used a 4-level integerized *lifting* wavelet transform [3] for compressing images.

Although most wavelet transform coefficients are small, their possible range is large. If each integer in this range is treated as a symbol for entropy encoding, the cardinality of the symbol set is large. This is not suitable for efficient encoding if most of the symbols do not occur, or occur rarely, from the standpoint of practicality. Since efficient entropy coding means that the symbol statistics must be transmitted to the receiver, a large cardinality bloats the size of the header information.

In order to simplify the entropy coding process, we use the 3 component representation for the transform coefficient first proposed by Said and Pearlman[4], called the *magnitude-set variable-length-integer* (MS-VLI). This representation splits the integer into a *magnitude* index, a *sign* bit and a *mantissa* or *variable length index.* The magnitude represents a bin that contains the raw integer. The sign bit and mantissa, whose length is determined from the magnitude, are transmitted as raw bits. The adverse impact of not entropy coding the mantissa is mild, since the distributions of mantissa bits within a magnitude bin is relatively flat. Since the bins are 2^k wide, k bits efficiently represent the location within this bin.

3. Compression Experiments

In our experiments, we had employed a 2D wavelet codec engine using the EigenWavelet (2,2) basis function. Figures 1 and 2 show the results of the compression experiments on material texture, and terrain image data respectively. The compression scheme, compression ratio (CR) and Mean Square Error (E) are indicated in each case. The images in Figure 1 represent the synthesized signature outputs from the compressed-restored material textures. The input material texture is obtained from the NVESD Paint the Night Fort Hunter-Liggett database, and the 3 sensor views shown represent FLIR at 13:00, FLIR at 23:00 and radar respectively. The terrain elevation data in Figure 2 is also from the NVESD database, which is a 4km x 4km block of 16-bit data at 1 metre spatial resolution. For the purpose of display, compressed-restored terrain data is graphically rendered from the same ground-view in JRM's OpenGL database viewer, and is also shown as a relief map gray-scale image.

Our algorithm design provides the user with the option of quantization levels to increase the compression ratio, with gracefully degrading compression. The numbers appended to LNK_EWT in the figures refer to the level of wavelet transform decompositions performed, and the amount of quantization applied to the wavelet transform coefficients. With suitable quantization parameters, one can achieve a variable rate of compression ranging from 2 to 30. It is clear that the choice of the parameters is primarily dependent on the type of the texture and its characteristics. Further, a near-maximum amount of compression can be achieved without much degradation in the visual quality of the reconstructed data. As expected, the sensor and terrain visualization views degrade with increasing wavelet transform quantization. In order to minimize errors, we have also employed a material index buffering scheme that significantly reduces the impact of material texture compression on sensor signature synthesis. The details will appear elsewhere.

4. Summary

We have developed a fast and scaleable wavelet-based hybrid algorithm for compressing texture and terrain images. Experimental results, as applied to the material textures and terrain data, demonstrate that the performance of our algorithm is superior to many of the existing schemes.

5. References

1. K. Sayood, *Introduction to Data Compression*, Morgan Kaufmann Publishers Inc. San Francisco, CA, 1996.

2. B. Yang, "Projection Approximation Subspace Tracking", IEEE Trans. Signal Proc., 43(1), pp.95-107, 1995.

3. W. Sweldons, "The Lifting Scheme: A Custom-design Construction of Second Generation Wavelets, SIAM Jl. Math. Anal., 29(2), 1997.

4. A. Said and W.A. Perlman, "An Image Multiresolution Representation for Lossless and Lossy Compression", SPIE Symp. On Visual Communications and Image Processing, Cambridge, MA, 1993.

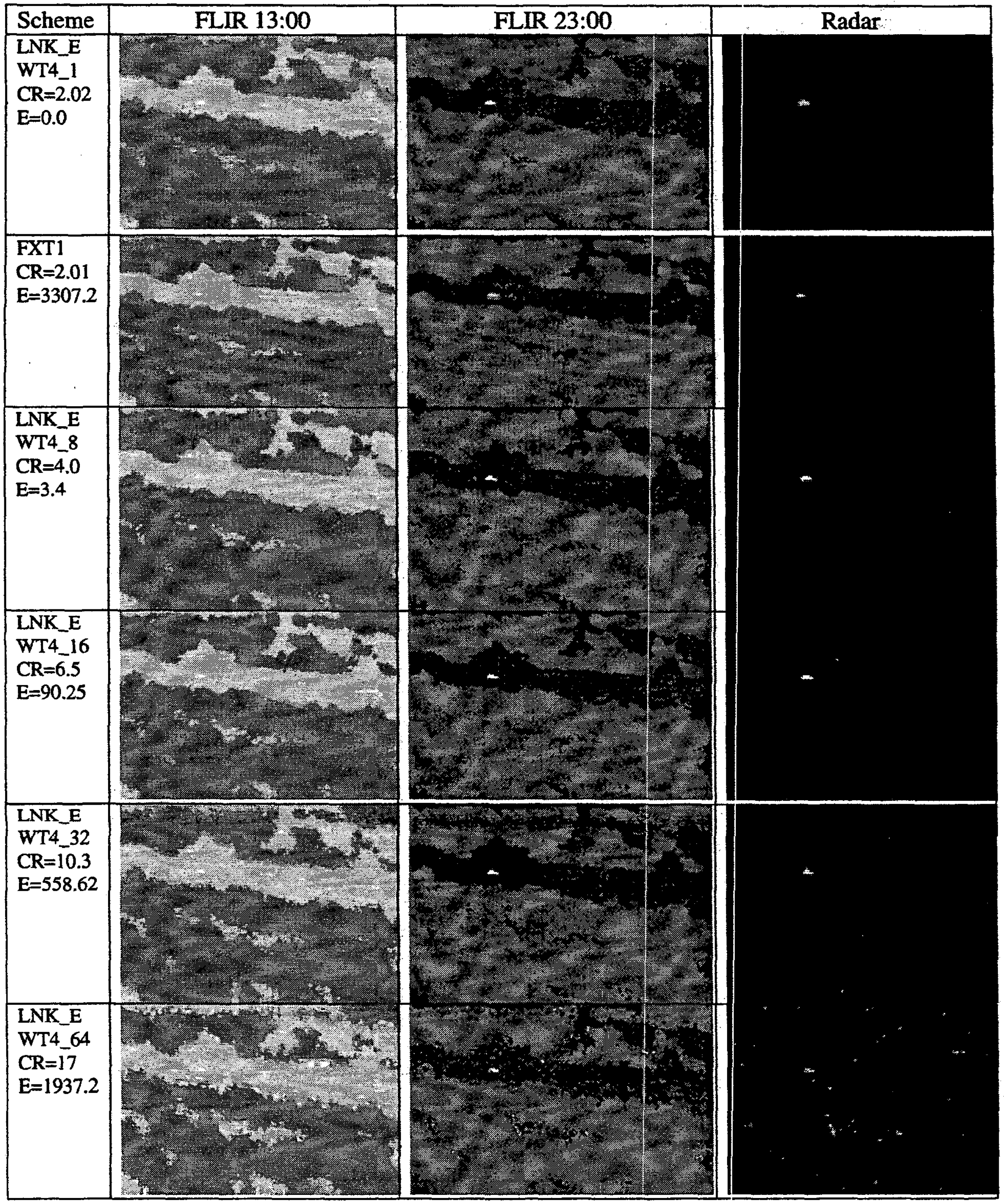


Figure 1. Sensor Signature Images after Compression of Material Texture

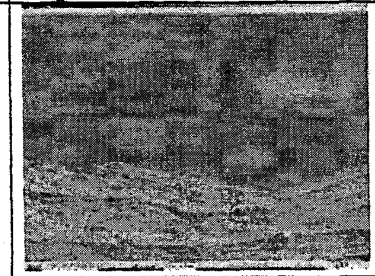
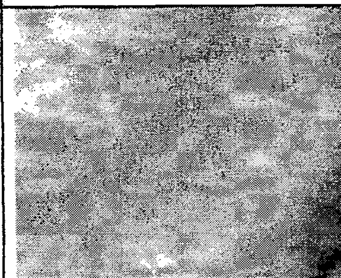
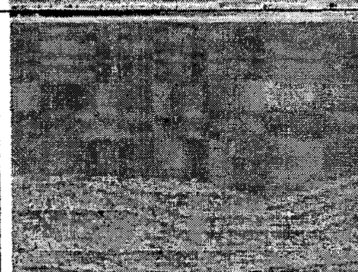
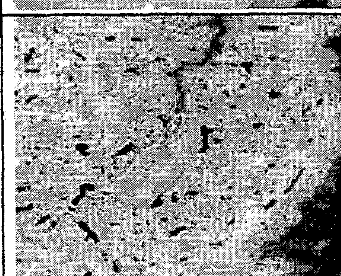
Scheme	OpenGL Ground Rendering	Relief View
LNK_EWT 4_1 CR=8.6 E=0.0 (No Loss)		
LNK_EWT 4_32 CR=937 E=58.96		
LNK_EWT 4_64 CR=1214 E=293.10		
LNK_EWT 4_128 CR=1639 E=1281.99		
FXT1 CR=4.02 E=327 E 6		

Figure 2. Compressed-Restored Terrain, NVESD PTN Ft. Hunter-Liggett 1 Metre, 16-bit Terrain