A Low Complexity Energy Spreading Transform Coder

J. S. Byrnes

Abstract. We have shown how the Prometheus Orthonormal Set (PONS^{tm}), originally constructed to prove an uncertainty principle conjecture of H.S.Shapiro, can be effectively used to compress all common digital audio signals. This compression method is effective because of two fundamental properties, computational simplicity and energy spreading. Although there exist other transform coding methods, such as Walsh-Hadamard, which give compression while limiting the computational burden, we believe that the energy spreading feature of PONS is unique. We discuss the various advantages that result from these properties, show how the multidimensional analogue of PONS is constructed, and present an algorithm to decompose multidimensional data sets into smaller blocks with uniformly bounded energy. We then indicate the application of PONS to image processing.

§1 Introduction

We defined in [4] a "Walsh-like" complete orthonormal sequence for $L^2(0, 2\pi)$ which satisfies several important properties. Each function in this *Prometheus Orthonormal Set* (PONS^{tm}) is piecewise ± 1, can change sign only at points of the form $\frac{k}{2^n}(2\pi)$, $1 \leq k \leq 2^n - 1$, and is easily computable using a straightforward and fast recursive algorithm. In addition to these features shared with the Walsh functions, PONS:

- Is optimal with respect to the global uncertainty principle described in [4]
- Yields the *uniform* crest factor $\sqrt{2}$
- Spreads energy almost equally among all transform domain bins

We discuss the utilization of these properties below, beginning with applications to one-dimensional digital signals. Then the possible advantages over other signal processing methods are described. Finally, we consider the application of PONS to image processing, and give some examples of our preliminary results.

A Low Complexity Energy Spreading Transform Coder
J. S. Byrnes
pp. 1–13.

ISBN 0-12-xxxxxx-x

§2 One-dimensional signals

2.1 Audio processing

Our first implementation of PONS [5] yields high quality data compression for audio. For 16-bit 44.1, 22.05, and 11.025 kHz monaural signals we achieve 4 to 1 compression with virtually no audible difference in sound. The compression and decompression algorithms are comparable in execution speed and can operate in real-time on all modern PCs (e.g., 386 or higher IBM compatible, current Macintosh, UNIX workstation, etc.). These results are to be contrasted with a current standard, ADPCM, where the same compression ratio is achieved, but at much higher computational cost and with much lower quality.

2.2 Spread spectrum communications

A second application of PONS, in the context of one-dimensional signals, is in multi-user spread spectrum communications. The recently developed IS-95 standard for commercial code-division multiple access (CDMA) communications involves a two-stage direct sequence spreading process, first with a Walsh function and then with a longer pseudonoise (PN) sequence. Our research has shown that PONS offers two important advantages over the Walsh functions in this application:

- The minimal crest factor property of the PONS sequences provides much more uniform spreading of the signal's energy across the frequency band. This increases robustness with respect to channel effects, such as narrowband fading and interference, and also reduces spectral features that can be exploited by an adversary in military scenarios.
- Spreading with PONS sequences rather than Walsh sequences yields signals requiring lower short-term ("peak") power to maintain a specified average transmission power. This is advantageous in view of the significant power-control issues involved in multi-user spread spectrum (e.g., to address "near-far" problems). It also offers potential for reducing overall transmitter power requirements, which is highly desirable for battery-powered mobile units.

These advantages are illustrated in figures 1 and 2. Figure 1 compares frequency spectra of typical Walsh and PONS sequences. Figure 2 compares their autocorrelation structures.

Since PONS sequences can be as long as desired (any power of two), it is possible to use them in place of long PN sequences as spreading codes. We are currently investigating the potential value of direct sequence spreading based entirely on PONS sequences. One feature appears to be an improvement in short-time characteristics of the sequences. PN sequences can have arbitrarily long constant subintervals (i.e., consecutive terms all of which are 0 or all of

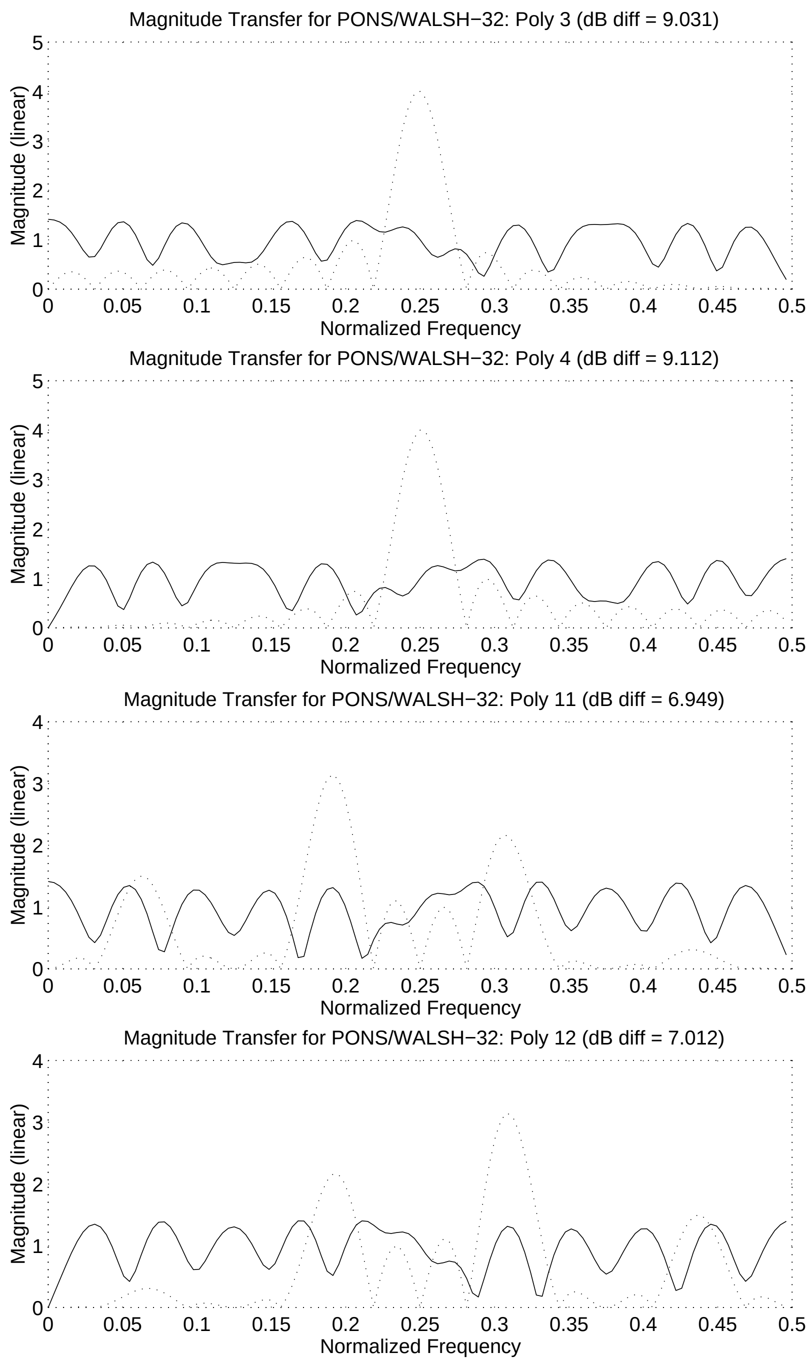

Figure 1. Magnitude spectra of representative 32-coefficient Walsh sequences (dotted curves) and PONS sequences (solid curves). Note the substantial reduction in peak dB for the PONS sequences.

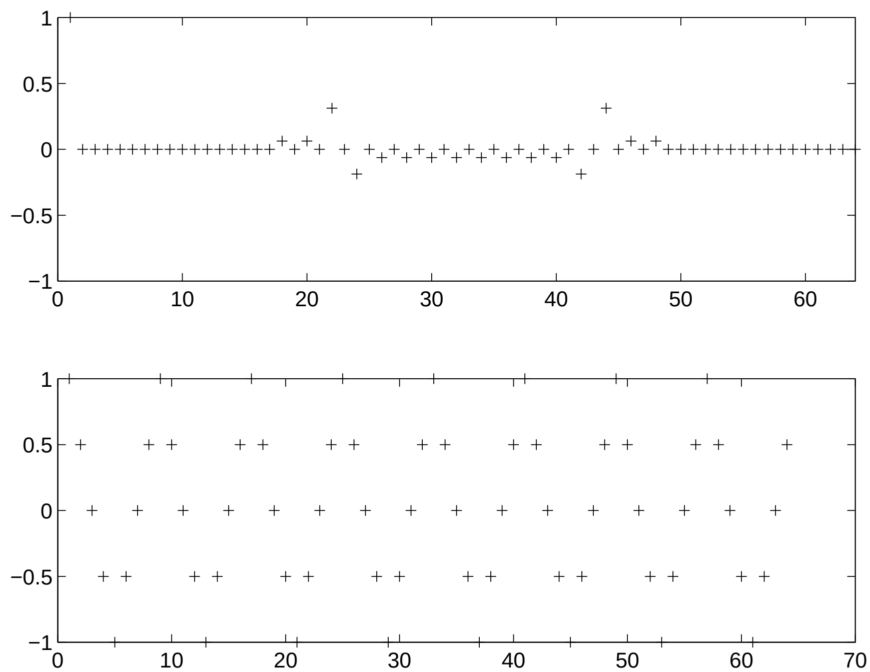

Figure 2. Autocorrelation functions of a representative 32-coefficient PONS sequence (top) and a representative Walsh sequence (bottom).

which are 1), whereas PONS sequences cannot have constant subintervals of length greater than 5.

§3 PONS advantages

The basic features of PONS, which have led to our initial success when applying it to one-dimensional signals and which are behind our optimism with regard to multidimensional applications, are:

- Very robust
- Easily extensible to any dimension
- Localizable (see section 4.4 and figures 5 and 6)
- Multi-level security
- Computational

    Extremely efficient computations

    All basic operations are integer adds and subtracts only

    Fast systolic algorithm

    Amenable to massively parallel processing

- Mathematical

    Elementary construction

    Quadrature mirror filter

    Low crest factor array

    Yields optimal uncertainty principle bounds

There are many potential benefits which may result from these features, and which we have begun to exploit in our initial one-dimensional applications. The benefits of this one-dimensional coder, when no quantization is used, include:

- No aliasing errors
- No amplitude or phase distortion
- Transform accomplished via adds/subtracts using integer (2's complement) arithmetic

- Exact reconstruction (no roundoff error due to coefficient roundoff or truncation error due to bit allocation based on source statistics)

- Near optimum coding gain for sufficiently wideband input (independent of input statistics or spectrum)

When quantization is employed, the coder has the following advantages over conventional coders:

- Quantizes all transform domain coefficients to the same precision

- Aliasing errors are canceled to the same precision in all transform domain subspaces

- Quantization errors appearing in the reconstruction are approximately Gaussian (as a linear sum of independent uniform errors)

- No bit allocation computation

- A real-time algorithm that has been implemented on general-purpose hardware

Among the many application areas that can take advantage of these benefits, the most promising are those that have requirements for:

- Wideband coding where the statistics of the source are either not known or not constrained to be anything specific

- Simple coding/decoding hardware

- Good performance over unreliable media (good transmission error properties)

- A simple progressive transmission scheme

§4 Image processing

In order to exploit the benefits indicated in §3 when processing multidimensional signals, we have generalized the one-dimensional basis given in [4] as well as derived a new tensor decomposition theorem. For simplicity we present only the two-dimensional versions of these two results here. Further details and proofs will appear elsewhere.

4.1 A two-dimensional PONS basis

Let the inner product $\langle A, B\rangle$ of the two matrices A, B (of the same dimension, say $a \times b$) be the sum of component-wise multiplication of the two matrices,

$$\langle A, B\rangle = \sum_{i=1}^{a}\sum_{h=1}^{b} a_{i,h} \cdot b_{i,h}.$$

Let $A_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, $A_2 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$, $A_3 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$, $A_4 = \begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix}$.
Observe that $\langle A_j, A_k\rangle = 4\delta_{jk}$. Thus, A_j, $1 \le j \le 4$, span the characteristic functions of the four sub-quartersquares, such as

0	1
0	0

(ignoring boundaries).

Theorem 1. *For any $n > 0$, there exist 2^{2n} matrices E_i, $1 \le i \le 2^{2n}$, with dimension $2^n \times 2^n$, with entries ± 1, and satisfying $\langle E_j, E_k\rangle = 2^{2n}\delta_{jk}$. Furthermore, these matrices satisfy properties corresponding to the uncertainty principle optimality, quadrature mirror filter, and low crest factor properties of their one-dimensional analog given in [4]. Finally, the sets of piecewise constant functions on the square, determined in the obvious way by these matrices, form a complete orthonormal set for $C([0,1] \times [0,1])$.*

We present the first step of the proof, since it gives a straightforward inductive definition of the E_i. A_i, $1 \le i \le 4$ above is the case $n = 1$. Suppose now that it is true for n. We must prove it for $n + 1$. So we have the E_i as stated in the theorem.

For each $j = 1, 2, \ldots, 2^{2n-2}$, form the $2^{n+1} \times 2^{n+1}$ matrices

$$F_1 = \begin{bmatrix} E_{4j-3} & E_{4j-2} \\ E_{4j-1} & E_{4j} \end{bmatrix}, \qquad F_2 = \begin{bmatrix} E_{4j-3} & E_{4j-2} \\ -E_{4j-1} & -E_{4j} \end{bmatrix},$$

$$F_3 = \begin{bmatrix} E_{4j-2} & E_{4j-3} \\ E_{4j} & E_{4j-1} \end{bmatrix}, \qquad F_4 = \begin{bmatrix} E_{4j-2} & E_{4j-3} \\ -E_{4j} & -E_{4j-1} \end{bmatrix},$$

$$F_9 = \begin{bmatrix} E_{4j-3} & -E_{4j-2} \\ E_{4j-2} & -E_{4j-3} \end{bmatrix}, \quad F_{10} = \begin{bmatrix} E_{4j-3} & -E_{4j-2} \\ -E_{4j-2} & E_{4j-3} \end{bmatrix},$$

$$F_{11} = \begin{bmatrix} -E_{4j-2} & E_{4j-3} \\ -E_{4j-3} & E_{4j-2} \end{bmatrix}, \quad F_{12} = \begin{bmatrix} -E_{4j-2} & E_{4j-3} \\ E_{4j-3} & -E_{4j-2} \end{bmatrix}.$$

F_5–F_8 are obtained from F_1–F_4 by exchanging E_{4j-3} with E_{4j-1} and E_{4j-2} with E_{4j}. F_{13}–F_{16} are the same as F_9–F_{12}, with E_{4j-3} replaced by E_{4j-1} and E_{4j-2} replaced by E_{4j}. Then there are $2^{2(n+1)}$ F_i's, and a straightforward calculation gives $\langle F_j, F_k \rangle = 2^{2(n+1)}\delta_{jk}$.

4.2 A two-dimensional tensor decomposition theorem

To take advantage of bases such as that constructed in 4.1, we have developed a method of decomposing a matrix into a sum of tensor products of smaller matrices. Once again we present only the result. The details and the generalization to higher dimensions, due mainly to B. Saffari and motivated by a conjecture of Izidor Gertner, will appear elsewhere.

Theorem 2. *Let $L(m,n)$ denote the set of all $m \times n$ real matrices, let a, b be divisors of m, n respectively, let $m = \mu \cdot a, n = \nu \cdot b$, let $A_j \in L(a,b)$, $(j = 1, 2, \ldots, a \cdot b)$ form an orthonormal basis of $L(a,b)$ under the inner product defined above, let $\otimes$ denote the standard tensor product, and let $M \in L(m,n)$. Then there exist $a \cdot b$ unique matrices $D_j \in L(\mu, \nu)$, $(j = 1, 2, \ldots, a \cdot b)$, such that*

$$M = \sum_{j=1}^{a \cdot b} D_j \otimes A_j.$$

4.3 Uniresolution and localization

Two of the unique and most important features of PONS that can be exploited for image processing are *uniresolution* and *localization*. Uniresolution means that when a decomposition of a high-resolution signal into a sum of low-resolution signals is performed, each of these low-resolution components contains essentially the same amount of energy. This is completely different from all classical signal processing techniques, and yields a possibly important advantage. Namely, since each low-resolution component contains significant energy from all portions of the signal, for applications such as automatic target detection/recognition and telebrowsing, just one or a few of these components might be sufficient. Thus, in certain instances, extremely high data compression is attainable even without employing any of the current sophisticated and often computationally intensive techniques. Moreover, PONS can be employed in conjunction with these other techniques if desired, thereby preserving their advantages while gaining the distinct properties of this new decomposition. A precise mathematical formulation of this uniresolution feature, developed mostly by H. S. Shapiro, will appear elsewhere.

Localization is the capability to reconstruct with high resolution any desired signal portion, while other parts of the signal remain at lower resolution. When applied to images, the result is a clear picture of the area that is wanted, while the background is recognizable but hazy. In telebrowsing, for example, it saves overall transmission and browsing time, and no redundant information need be stored or transmitted.

These properties combine to allow, with small computational effort involving only adds and subtracts, the decomposition of a high-resolution image into a collection of low-resolution images, which can then be processed individually or in parallel to determine where in the image a further search is necessary. Then the localization property can be exploited to exactly reconstruct the locations of interest. Due to the features of PONS discussed above and illustrated in figures $3-14$, this can be accomplished with minimal CPU time and computer memory. The figures require some explanation, which will be better understood if it is read in its entirety *before* they are examined.

Figure 3 is a low-resolution image of a particular scene. In fact, it is just 1 of the 256 low-resolution images that, when combined, yield the high-resolution image shown in figure 4. Thus, if figure 3 could be used to recognize the image, we would immediately have a compression ratio of 256 to 1. Even if recognition were not possible directly from figure 3, if one could determine where in the overall image there are regions of interest, and then focus on these regions while ignoring the remainder of the image, significant data compression would still be achieved.

However, when the reader first examines figure 3, it appears that nothing whatsoever can be recognized there. Certainly there are variations in grayscale occurring throughout, but just as certainly (or so it *appears*) there is no particular shape to be found. But, if one looks at figure 4 and then reexamines figure 3, the scene has appeared there as well! It might require slight changes in the distance or angle between the figure and the eye in order to see this, but it was indeed there all the time. A focus of our current research is to modify the PONS algorithm, so that the computer can see what is clearly in the very low-resolution image, but which the eye first misses.

The localization feature is illustrated in figures 5 and 6. In 5 the lower right portion of the image has been isolated by using only one thirty-second of the area of figure 4 for the reconstruction. The remainder of the image is compressed at 8 to 1, for the purpose of illustration. Thus, the compression ratio achieved in obtaining an exact reconstruction of a crucial portion of the image is almost 32 to 1. Similar remarks apply to figure 6.

Finally, robustness is illustrated in figure $7-14$. In figure 7, the 256-term expansion whose sum yields the exact reconstruction of figure 4 was scrambled by randomly permuting the coefficients. In spite of the fact that every single coefficient was thereby matched to the wrong basis element, the resulting image is still recognizable!

Lest the reader feel that the natural redundancy of the original image has contributed to this robustness, we present three additional scrambled and exactly reconstructed images in figures $8-13$. There is certainly no question as to which scrambled image belongs to which original. In fact, the scrambled images are somewhat recognizable on their own, even without seeing the full reconstructions. Robustness is further illustrated in figure 14, where the reproduced image is reasonably good despite the loss of half of the terms in the expansion (i.e., half of the data).

4.4 Applications

All of the benefits indicated in section §3 apply to the two-dimensional PONS algorithms as well. Our current image processing research deals with utilizing the distinctive properties of computational simplicity, robustness, localization, speed, and amenability to parallel computation in areas such as ATD/R, image compression and transmission over narrow bandwidth channels, and image characterization for the purpose of efficiently searching large databases.

Automatic target detection and recognition

Fundamental ATD/R problems require effective feature extraction, often in the presence of heavy clutter and noise, for their solution in terms of detection and recognition. One potential application of PONS here is to use its data compression and real-time localization capabilities to preprocess the image, in order to determine quickly which areas within the image require further investigation, and to localize further processing to these areas.

Another possibility is to compare a highly compressed (say one, or at most a few, PONS coefficients) version of the unknown image with a library of images stored in this same highly compressed manner. Because of the energy spreading feature, in which each individual coefficient contains essentially the same amount of hazy information about the entire image, in many situations a standard pattern-matching technique may be able to determine which, if any, of the stored images matches the unknown one.

Progressive image transmission

Progressive image transmission is a method of encoding, transmitting, and decoding digitized data representing an image in such a way that the main features of the image, for example outlines, may be displayed first at low resolution and subsequently refined to higher and higher resolution. In progressive transmission, an image is encoded by an electronic analog to multiple scans of the same image at differing resolutions.

At the receiving end, progressive image decoding results in an initial approximate reconstruction of the image, followed by successively better images whose fidelity is gradually built up from succeeding scans. In one embodiment of this general scheme, the receiver or viewer can abort the transmission sequence at a less than perfect resolution or can decide on the basis of a partial image to proceed with further transmission and reconstruction. The objective is to show significant but broad features of an image at an early stage of transmission so that a viewer can interactively respond. There is a kind of coarse data compression resulting from this method of transmitting images which arises from the fact that in appropriate applications only a small part of the total data needs to be sent if a low resolution image is desired.

The order in which the image data are selected, transmitted, and presented to the end user may be dynamically prioritized in response to image content and immediate user interest. Such transmission can result in a display which has a non-uniform resolution. Regions containing visually or operationally significant information may be rendered at a much higher resolution, with refinement deferred for areas of uniform intensity or lesser importance.

Of the many progressive image schemes in existence, the technique in most widespread use in transform domain progressive image transmission is based on the Discrete Cosine Transform (DCT). DCT progressive image coding is incorporated in the recently approved JPEG standard. These methods of implementing progressive image transmission, however, have the following disadvantages:

- The computational burden of the inverse progressive image transmission is high, since it requires the computation of the Inverse DCT (in addition to the forward DCT at the transmission end). To satisfy these computational requirements, the DCT is often implemented using specialized VLSI circuits.

- Using a given JPEG encoding of an image, it is not possible to reconstruct efficiently an arbitrary subimage of high interest with full resolution while other low interest areas remain at low resolution. Contrast this with figures 5 and 6, produced using PONS.

- When progressive image transmission is applied to telebrowsing of archived images, highly compressed data are kept in on-line storage and used to browse the data efficiently to determine potentially useful data sets for further processing. Once the selection is made, the original data are obtained from off-line storage. Here the browse quality data and the corresponding original data contain redundant information, causing a fraction of this information to be transmitted twice, thereby reducing performance.

Progressive image transmission employing PONS is being designed to overcome these and other deficiencies. Localization and robustness are the two most important features in this regard. They are demonstrated in figures $5-14$. The reason behind the robustness of this coding technique, particularly clear in figure 14, is that PONS provides a "balanced" representation of typical images, i.e., one in which the energy is spread evenly among all the transform coefficients. In contrast, Fourier and wavelet representations of typical images concentrate energy in low-frequency terms. If any of these terms are lost in transmission, the image will be essentially destroyed.

Medical image processing

PONS offers several important potential benefits for the processing of medical images. Such images will typically be large (as much as 1 gigabyte for a single pap smear slide at 40× magnification). On the other hand, much of the key diagnostic information will be contained in small portions of the image. The localization feature will facilitate quick telemedicine access to the significant portions while minimizing the load on the network. Similar savings can be imagined in medical image storage. Further, in the context of a preprocessor which highlights anomalies in the image, such a facility will aid diagnosis.

Image database search

As databases of images grow, the problem of searching them to locate desired images becomes more and more intractable. A basic constraint is the amount of data, typically on the order of a megabyte of 8-bit numbers for each black-and-white image. Thus, identifying and retrieving a specific image can require processing involving an enormous number of pixels. With current databases consisting of tens of thousands of images, an efficient means of image characterization for search purposes is essential. Employing properly designed image snippets, which enable significant data reduction, is one promising approach.

To achieve both uniqueness and data reduction, a procedure is required which employs an image snippet that is both a highly compressed version of the original image and also contains information from the entire image. These are exactly the properties that characterize the uniresolution feature discussed in section 4.3.

Another desirable property, for example in telebrowsing, is to be able to examine in full detail a portion of an image without having to manipulate the large amount of data required to view it in its entirety. The localization feature can be immediately employed to perform this required task.

§5 Conclusion

Any digital signal of any dimension can be simply and uniquely expressed in terms of PONS. This fact, together with its unique localization capability and its robustness, indicate considerable potential for this new digital coding and transmission tool. Further research, blending the mathematical ideas with the engineering implementation, will determine whether this promise can be brought to fruition.

References

[1] J. Benedetto and A. Teolis, "Wavelet auditory models and data compression," *Applied and Computational Harmonic Analysis*, vol. 1, pp. 3–28, 1993.

[2] J. Benedetto and A. Teolis, Nonlinear method and apparatus for coding and decoding acoustic signals with data compression and noise suppression us-

ing cochlear filters, wavelet analysis, and irregular sampling reconstruction. U.S. Patent 5388182, 1995.

[3] J. Benedetto, J.S. Byrnes, D.J. Newman, H.S. Shapiro, and A. Teolis, Wavelet auditory models and irregular sampling. Tech. Report AR-92-2, Prometheus Inc., 1992.

[4] J.S. Byrnes, "Quadrature mirror filters, low crest factor arrays, functions achieving optimal uncertainty principle bounds, and complete orthonormal sequences – a unified approach," *Applied and Computational Harmonic Analysis*, vol. 2, pp. 261–266, 1994.

[5] James S. Byrnes, Michael A. Ramalho, Gerald Ostheimer and Izidor Gertner, Discrete one dimensional signal processing method and apparatus using energy spreading coding. U.S. Patent pending.

[6] B. Le Floch, M. Alard, and C. Berrou, "Coded Orthogonal Frequency Division Multiplex," *Proc. of the IEEE*, vol. 83, no. 6, Special Issue on Digital Television, pp. 982–996, 1995.

[7] H.S. Shapiro, Extremal problems for polynomials and power series. Sc.M. Thesis, Massachusetts Institute of Technology, 1951.

[8] R.N.J. Veldhuis, "Bit rates in audio source coding," *IEEE Journal on Selected Areas in Communications*, vol. 10(1), 1992.

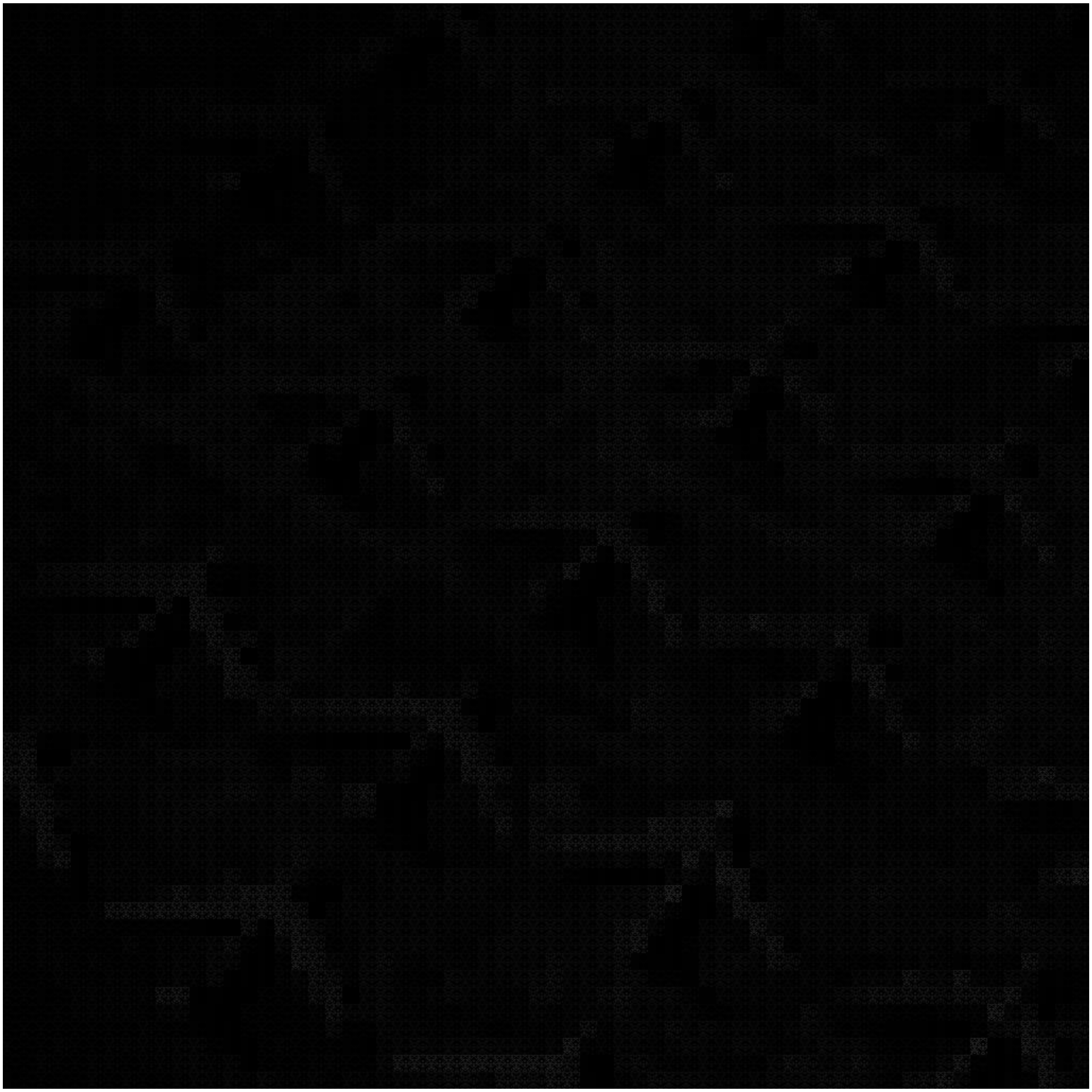

Figure 3. Test image I, 256:1 compression.

Figure 4. Test image I, full reconstruction.

Figure 5. Test image I, localized.

Figure 6. Test image II, localized. Full resolution features in a low resolution (compressed) background.

Figure 7. Test image I, scrambled.

Figure 8. Test image III, scrambled.

Figure 9. Test image IV, scrambled.

Figure 10. Test image II, scrambled.

Figure 11. Test image II, full reconstruction.

Figure 12. Test image III, full reconstruction.

Figure 13. Test image IV, full reconstruction.

Figure 14. Test image III reconstructed using half of the terms (128 out of 256).